2. Advanced Counting

When properly applied, the (double) counting argument can lead to more subtle results than those discussed in the previous chapter.

2.1 Bounds on intersection size

How many r-element subsets of an n-element set can we choose under the restriction that no two of them share more than k elements? Intuitively, the smaller k is, the fewer sets we can choose. This intuition can be made precise as follows. (We address the optimality of this bound in Exercise 2.5.)

Lemma 2.1 (Corrádi 1969)**.** *Let $A_1, \ldots, A_N$ be r-element sets and X be their union. If $|A_i \cap A_j| \leq k$ for all $i \neq j$, then*

$$|X| \geq \frac{r^2 N}{r + (N-1)k}. \tag{2.1}$$

Proof. Just count. By (1.11), we have for each $i = 1, \ldots, N$,

$$\sum_{x \in A_i} d(x) = \sum_{j=1}^{N} |A_i \cap A_j| = |A_i| + \sum_{j \neq i} |A_i \cap A_j| \leq r + (N-1)k\,. \tag{2.2}$$

Summing over all sets A_i and using Jensen's inequality (1.15) we get

$$\sum_{i=1}^{N} \sum_{x \in A_i} d(x) = \sum_{x \in X} d(x)^2 \geq \frac{1}{n}\Big(\sum_{x \in X} d(x)\Big)^2 = \frac{1}{n}\Big(\sum_{i=1}^{n} |A_i|\Big)^2 = \frac{(Nr)^2}{n}\,.$$

Using (2.2) we obtain $(Nr)^2 \leq N \cdot |X|\,(r + (N-1)k)$, which gives the desired lower bound on $|X|$. $\square$

S. Jukna, *Extremal Combinatorics*, Texts in Theoretical Computer Science. An EATCS Series, DOI 10.1007/978-3-642-17364-6_2,

Given a family of sets $A_1, \ldots, A_N$, their *average size* is

$$\frac{1}{N} \sum_{i=1}^{N} |A_i|.$$

The following lemma says that, if the average size of sets is large, then some two of them must share many elements.

Lemma 2.2. *Let X be a set of n elements, and let $A_1, \ldots, A_N$ be subsets of X of average size at least n/w. If $N \geq 2w^2$, then there exist $i \neq j$ such that*

$$|A_i \cap A_j| \geq \frac{n}{2w^2}. \tag{2.3}$$

Proof. Again, let us just count. On the one hand, using Jensen's inequality (1.15) and equality (1.10), we obtain that

$$\sum_{x \in X} d(x)^2 \geq \frac{1}{n} \Big(\sum_{x \in X} d(x) \Big)^2 = \frac{1}{n} \Big(\sum_{i=1}^{N} |A_i| \Big)^2 \geq \frac{nN^2}{w^2} \,.$$

On the other hand, assuming that (2.3) is false and using (1.11) and (1.12) we would obtain

$$\begin{aligned} \sum_{x \in X} d(x)^2 &= \sum_{i=1}^{N} \sum_{j=1}^{N} |A_i \cap A_j| = \sum_{i} |A_i| + \sum_{i \neq j} |A_i \cap A_j| \\ &< nN + \frac{nN(N-1)}{2w^2} = \frac{nN^2}{2w^2} \Big(1 + \frac{2w^2}{N} - \frac{1}{N} \Big) \leq \frac{nN^2}{w^2}, \end{aligned}$$

a contradiction. □

Lemma 2.2 is a very special (but still illustrative) case of the following more general result.

Lemma 2.3 (Erdős 1964b)**.** *Let X be a set of n elements $x_1, \ldots, x_n$, and let $A_1, \ldots, A_N$ be N subsets of X of average size at least n/w. If $N \geq 2kw^k$, then there exist $A_{i_1}, \ldots, A_{i_k}$ such that $|A_{i_1} \cap \cdots \cap A_{i_k}| \geq n/(2w^k)$.*

The proof is a generalization of the one above and we leave it as an exercise (see Exercises 2.8 and 2.9).

2.2 Graphs with no 4-cycles

Let H be a fixed graph. A graph is H-free if it does not contain H as a subgraph. (Recall that a *subgraph* is obtained by deleting edges and vertices.) A typical question in graph theory is the following one:

How many edges can a H-free graph with n vertices have?

That is, one is interested in the maximum number $\mathrm{ex}(n, H)$ of edges in a H-free graph on n vertices. The graph H itself is then called a "forbidden subgraph."

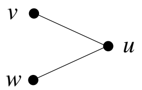
Let us consider the case when forbidden subgraphs are *cycles*. Recall that a cycle C_k of length k (or a k-cycle) is a sequence $v_0, v_1, \ldots, v_k$ such that $v_k = v_0$ and each subsequent pair v_i and v_{i+1} is joined by an edge.

If $H = C_3$, a triangle, then $\mathrm{ex}(n, C_3) \geq n^2/4$ for every even $n \geq 2$: a complete bipartite $r \times r$ graph $K_{r,r}$ with $r = n/2$ has no triangles but has $r^2 = n^2/4$ edges. We will show later that this is already optimal: any n-vertex graph with more than $n^2/4$ edges must contain a triangle (see Theorem 4.7). Interestingly, $\mathrm{ex}(n, C_4)$ is much smaller, smaller than $n^{3/2}$.

Theorem 2.4 (Reiman 1958)**.** *If $G = (V, E)$ on n vertices has no 4-cycles, then*

$$|E| \leq \frac{n}{4}(1 + \sqrt{4n - 3}) .$$

Proof. Let $G = (V, E)$ be a C_4-free graph with vertex-set $V = \{1, \ldots, n\}$, and $d_1, d_2, \ldots, d_n$ be the degrees of its vertices. We now count in two ways the number of elements in the following set S. The set S consists of all (ordered) pairs $(u, \{v, w\})$ such that $v \neq w$ and u is adjacent to both v and w in G. That is, we count all occurrences of "cherries"



in G. For each vertex u, we have $\binom{d_u}{2}$ possibilities to choose a 2-element subset of its d_u neighbors. Thus, summing over u, we find $|S| = \sum_{u=1}^{n} \binom{d_u}{2}$. On the other hand, the C_4-freeness of G implies that no pair of vertices $v \neq w$ can have more than one common neighbor. Thus, summing over all pairs we obtain that $|S| \leq \binom{n}{2}$. Altogether this gives

$$\sum_{i=1}^{n} \binom{d_i}{2} \leq \binom{n}{2}$$

or

$$\sum_{i=1}^{n} d_i^2 \leq n(n-1) + \sum_{i=1}^{n} d_i . \tag{2.4}$$

Now, we use the Cauchy–Schwarz inequality

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right)$$

with $x_i = d_i$ and $y_i = 1$, and obtain

$$\left(\sum_{i=1}^{n} d_i\right)^2 \le n\sum_{i=1}^{n} d_i^2$$

and hence by (2.4)

$$\left(\sum_{i=1}^{n} d_i\right)^2 \le n^2(n-1) + n\sum_{i=1}^{n} d_i\,.$$

Euler's theorem gives $\sum_{i=1}^{n} d_i = 2|E|$. Invoking this fact, we obtain

$$4|E|^2 \le n^2(n-1) + 2n|E|$$

or

$$|E|^2 - \frac{n}{2}|E| - \frac{n^2(n-1)}{4} \le 0\,.$$

Solving the corresponding quadratic equation yields the desired upper bound on $|E|$. □

Example 2.5 (Construction of dense C_4-free graphs). The following construction shows that the bound of Theorem 2.4 is optimal up to a constant factor.

Let p be a prime number and take $V = (\mathbb{Z}_p \setminus \{0\}) \times \mathbb{Z}_p$, that is, vertices are pairs (a, b) of elements of a finite field with $a \ne 0$. We define a graph G on these vertices, where (a, b) and (c, d) are joined by an edge iff $ac = b + d$ (all operations modulo p). For each vertex (a, b), there are $p - 1$ solutions of the equation $ax = b + y$: pick any $x \in \mathbb{Z}_p \setminus \{0\}$, and y is uniquely determined. Thus, G is a $(p - 1)$-regular graph on $n = p(p - 1)$ vertices (some edges are loops). The number of edges in it is $n(p - 1)/2 = \Omega(n^{3/2})$.

To verify that the graph is C_4-free, take any two its vertices (a, b) and (c, d). The unique solution (x, y) of the system

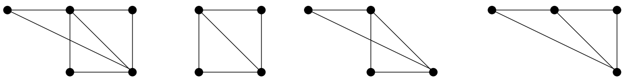
$$\begin{cases} ax = b + y \\ cx = d + y \end{cases} \qquad \text{is given by} \qquad \begin{aligned} x &= (b - d)(a - c)^{-1} \\ 2y &= x(a + c) - b - d \end{aligned}$$

which is only defined when $a \ne c$, and has $x \ne 0$ only when $b \ne d$. Hence, if $a \ne c$ and $b \ne d$, then the vertices (a, b) and (c, d) have precisely one common neighbor, and have no common neighbors at all, if $a = c$ or $b = d$.

2.3 Graphs with no induced 4-cycles

Recall that an *induced subgraph* is obtained by deleting vertices together with all the edges incident to them (see Fig. 2.1).

Theorem 2.4 says that a graph cannot have many edges, unless it contains C_4 as a (not necessarily induced) subgraph. But what about graphs that



G

Fig. 2.1 Graph G contains several copies of C_4 as a subgraph, but none of them as an *induced* subgraph.

do not contain C_4 as an *induced* subgraph? Let us call such graphs *weakly C_4-free.*

Note that such graphs can already have many more edges. In particular, the complete graph K_n is weakly C_4-free: in any 4-cycle there are edges in K_n between non-neighboring vertices of C_4. Interestingly, any(!) dense enough weakly C_4-free graph must contain large complete subgraphs.

Let $\omega(G)$ denote the maximum number of vertices in a complete subgraph of G. In particular, $\omega(G) \leq 3$ for every C_4-free graph. In contrast, for *weakly* C_4-free graphs we have the following result, due to Gyárfás, Hubenko and Solymosi (2002).

Theorem 2.6. *If an n-vertex graph $G = (V, E)$ is weakly C_4-free, then*

$$\omega(G) \geq 0.4 \frac{|E|^2}{n^3} \, .$$

The proof of Theorem 2.6 is based on a simple fact, relating the average degree with the minimum degree, as well as on two facts concerning independent sets in weakly C_4-free graphs.

For a graph $G = (V, E)$, let $e(G) = |E|$ denote the number of its edges, $d_{\min}(G)$ the smallest degree of its vertices, and $d_{\text{ave}}(G) = 2e(G)/|V|$ the average degree. Note that, by Euler's theorem, $d_{\text{ave}}(G)$ is indeed the sum of all degrees divided by the total number of vertices.

Proposition 2.7. *Every graph G has an induced subgraph H with*

$$d_{\text{ave}}(H) \geq d_{\text{ave}}(G) \qquad \textit{and} \qquad d_{\min}(H) \geq \frac{1}{2}\, d_{\text{ave}}(G) \, .$$

Proof. We remove vertices one-by-one. To avoid the danger of ending up with the empty graph, let us remove a vertex $v \in V$ if this does not decrease the average degree $d_{\text{ave}}(G)$. Thus, we should have

$$d_{\text{ave}}(G - v) = \frac{2(e(G) - d(v))}{|V| - 1} \geq d_{\text{ave}}(G) = \frac{2e(G)}{|V|}$$

which is equivalent to $d(v) \leq d_{\text{ave}}(G)/2$. So, when we stick, each vertex in the resulting graph H has minimum degree at least $d_{\text{ave}}(G)/2$. $\square$

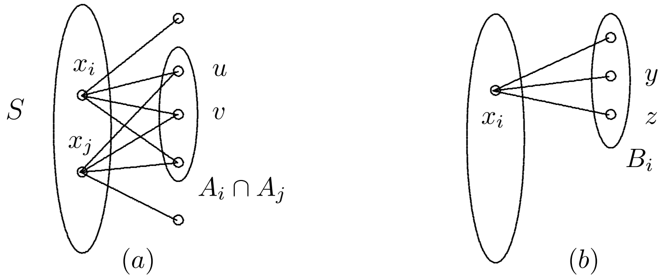


Fig. 2.2 (a) If u and v were non-adjacent, we would have an induced 4-cycle $\{x_i, x_j, u, v\}$. (b) If y and z were non-adjacent, then $(S \setminus \{x_i\}) \cup \{y, z\}$ would be a larger independent set.

Recall that a set of vertices in a graph is *independent* if no two of its vertices are adjacent. Let $\alpha(G)$ denote the largest number of vertices in such a set.

Proposition 2.8. *For every weakly C_4-free graph G on n vertices, we have*

$$\omega(G) \geq \frac{n}{\binom{\alpha(G)+1}{2}}.$$

Proof. Fix an independent set $S = \{x_1, \ldots, x_\alpha\}$ with $\alpha = \alpha(G)$. Let A_i be the set of neighbors of x_i in G, and B_i the set of vertices whose only neighbor in S is x_i. Consider the family $\mathcal{F}$ consisting of all α sets $\{x_i\} \cup B_i$ and $\binom{\alpha}{2}$ sets $A_i \cap A_j$. We claim that:

(i) each member of $\mathcal{F}$ forms a clique in G, and
(ii) the members of $\mathcal{F}$ cover all vertices of G.

The sets $A_i \cap A_j$ are cliques because G is weakly C_4-free: Any two vertices $u \neq v \in A_i \cap A_j$ must be joined by an edge, for otherwise $\{x_i, x_j, u, v\}$ would form a copy of C_4 as an induced subgraph. The sets $\{x_i\} \cup B_i$ are cliques because S is a maximal independent set: Otherwise we could replace x_i in S by any two vertices from B_i. By the same reason (S being a *maximal* independent set), the members of $\mathcal{F}$ must cover all vertices of G: If some vertex v were not covered, then $S \cup \{v\}$ would be a larger independent set.

Claims (i) and (ii), together with the averaging principle, imply that

$$\omega(G) \geq \frac{n}{|\mathcal{F}|} = \frac{n}{\alpha + \binom{\alpha}{2}} = \frac{n}{\binom{\alpha+1}{2}}. \qquad \square$$

Proposition 2.9. *Let G be a weakly C_4-free graph on n vertices, and $d = d_{\min}(G)$. Then, for every $t \leq \alpha(G)$,*

$$\omega(G) \geq \frac{d \cdot t - n}{\binom{t}{2}}.$$

Proof. Take an independent set $S = \{x_1, \ldots, x_t\}$ of size t and let A_i be the set of neighbors of x_i in G. Let m be the maximum of $|A_i \cap A_j|$ over all $1 \leq i < j \leq t$. We already know that each $A_i \cap A_j$ must form a clique; hence, $\omega(G) \geq m$. On the other hand, by the Bonferroni inequality (Exercise 1.37) we have that

$$n \geq \left| \bigcup_{i=1}^{t} A_i \right| \geq td - \sum_{i<j} |A_i \cap A_j| \geq td - \binom{t}{2} m \,,$$

from which the desired lower bound on $\omega(G)$ follows. □

Now we are able to prove Theorem 2.6.

Proof of Theorem 2.6. Let a be the average degree of G; hence, $a = 2|E|/n$. By Proposition 2.7, we know that G has an induced subgraph of average degree $\geq a$ and minimum degree $\geq a/2$. So, we may assume w.l.o.g. that the graph G itself has these two properties. We now consider the two possible cases.

If $\alpha(G) \geq 4n/a$, then we apply Proposition 2.9 with* $t = 4n/a$ and obtain

$$\omega(G) \geq \frac{(a/2) \cdot t - n}{\binom{t}{2}} = \frac{n}{\binom{4n/a}{2}} \,.$$

If $\alpha(G) \leq 4n/a$, then we apply Proposition 2.8 and obtain

$$\omega(G) \geq \frac{n}{\binom{\alpha(G)+1}{2}} \geq \frac{n}{\binom{4n/a+1}{2}} \,.$$

In both cases we obtain

$$\omega(G) \geq \frac{n}{\binom{4n/a+1}{2}} = \frac{a^2}{8n + 2a} \geq 0.1 \frac{a^2}{n} \,.$$ □

2.4 Zarankiewicz's problem

At most how many 1s can an $n \times n$ 0-1 matrix contain if it has no $a \times b$ submatrix whose entries are all 1s? Zarankiewicz (1951) raised the problem of the estimation of this number for $a = b = 3$ and $n = 4, 5, 6$ and the general problem became known as *Zarankiewicz's problem.*

It is worth reformulating this problem in terms of bipartite graphs. A bipartite graph with parts of size n is a triple $G = (V_1, V_2, E)$, where V_1 and V_2 are disjoint n-element sets of *vertices* (or *nodes*), and $E \subseteq V_1 \times V_2$ is the set of *edges*. We say that the graph contains an $a \times b$ *clique* if there exist an

* For simplicity, we ignore ceilings and floors.

a-element subset $A \subseteq V_1$ and a b-element subset $B \subseteq V_2$ such that $A \times B \subseteq E$. (Note that an $a \times b$ clique is not the same as a $b \times a$ clique, unless $a = b$.)

Let $k_a(n)$ be the minimal integer k such that *any* bipartite graph with parts of size n and more than k edges contains at least one $a \times a$ clique. Using the probabilistic argument, it can be shown (see Exercise 20.6) that

$$k_a(n) \geq c \cdot n^{2-2/a},$$

where $c > 0$ is a constant, depending only on a. It turns out that this bound is not very far from the best possible, and this can be proved using the double counting argument. The result is essentially due to Kővári, Sós and Turán (1954). For $a = 2$, a lower bound $k_2(n) \leq 3n^{3/2}$ was proved by Erdős (1938). He used this to prove that, if a set $A \subseteq [n]$ is such that the products of any two of its different members are different, then $|A| \leq \pi(n) + O(n^{3/4})$, where $\pi(n)$ is the number of primes not exceeding n.

Theorem 2.10. *For all natural numbers $n \geq a \geq 2$ we have*

$$k_a(n) \leq (a-1)^{1/a} n^{2-1/a} + (a-1)n.$$

Proof. The proof is a direct generalization of a double counting argument we used in the proof of Theorem 2.4. Our goal is to prove the following: let $G = (V_1, V_2, E)$ be a bipartite graph with parts of size n, and suppose that G *does not* contain an $a \times a$ clique; then $|E| \leq (a-1)^{1/a} n^{2-1/a} + (a-1)n$.

By a *star* in the graph G we will mean a set of any a of its edges incident with one vertex $x \in V_1$, i.e., a set of the form

$$S(x, B) := \{(x, y) \in E \ : \ y \in B\},$$

where $B \subseteq V_2$, $|B| = a$. Let Δ be the total number of such stars in G. We may count the stars $S(x, B)$ in two ways, by fixing either the vertex x or the subset B.

For a fixed subset $B \subseteq V_2$, with $|B| = a$, we can have at most $a - 1$ stars of the form $S(x, B)$, because otherwise we would have an $a \times a$ clique in G. Thus,

$$\Delta \leq (a-1) \cdot \binom{n}{a}. \tag{2.5}$$

On the other hand, for a fixed vertex $x \in V_1$, we can form $\binom{d(x)}{a}$ stars $S(x, B)$, where $d(x)$ is the degree of vertex x in G (i.e., the number of vertices adjacent to x). Therefore,

$$\sum_{x \in V_1} \binom{d(x)}{a} \leq (a-1) \cdot \binom{n}{a}. \tag{2.6}$$

We are going to estimate the left-hand side from below using Jensen's inequality. Unfortunately, the function $\binom{x}{a} = x(x-1)\cdots(x-a+1)/a!$ is convex only for $x \geq a-1$. But we can set $f(z) := \binom{x}{a}$ if $x \geq a-1$, and $f(x) := 0$

otherwise. Then Jensen's inequality (1.14) (with $\lambda_x = 1/n$ for all $x \in V_1$) yields

$$\sum_{x \in V_1} \binom{d(x)}{a} \geq \sum_{x \in V_1} f(d(x)) \geq n \cdot f\Big(\sum_{x \in V_1} d(x)/n\Big) = n \cdot f(|E|/n)\,.$$

If $|E|/n < a-1$, there is nothing to prove. So, we can suppose that $|E|/n \geq a-1$. Then we have that

$$n \cdot \binom{|E|/n}{a} = n \cdot f(|E|/n) \leq \sum_{x \in V_1} \binom{d(x)}{a} \leq (a-1)\binom{n}{a}.$$

Expressing the binomial coefficients as quotients of factorials, this inequality implies

$$n\,(|E|/n - (a-1))^a \leq (a-1)n^a,$$

and therefore $|E|/n \leq (a-1)^{1/a} n^{1-1/a} + a - 1$, from which the desired upper bound on $|E|$ follows. □

The theorem above says that any bipartite graph with many edges has large cliques. In order to destroy such cliques we can try to remove some of their vertices. We would like to remove as few vertices as possible. Just how few says the following result.

Theorem 2.11 (Ossowski 1993)**.** *Let $G = (V_1, V_2, E)$ be a bipartite graph with no isolated vertices, $|E| < (k+1)r$ edges and $d(y) \leq r$ for all $y \in V_2$. Then we can delete at most k vertices from V_1 so that the resulting graph has no $(r-a+1) \times a$ clique for $a = 1, 2, \ldots, r$.*

For a vertex x, let $N(x)$ denote the set of its neighbors in G, that is, the set of all vertices adjacent to x; hence, $|N(x)|$ is the degree $d(x)$ of x. We will use the following lemma relating the degree to the total number of vertices.

Lemma 2.12. *Let (X, Y, E) be a bipartite graph with no isolated vertices, and $f : Y \to [0, \infty)$ be a function. If the inequality $d(y) \leq d(x) \cdot f(y)$ holds for each edge $(x, y) \in E$, then $|X| \leq \sum_{y \in Y} f(y)$.*

Proof. By double counting,

$$\begin{aligned} |X| &= \sum_{x \in X} \sum_{y \in N(x)} \frac{1}{d(x)} \leq \sum_{x \in X} \sum_{y \in N(x)} \frac{f(y)}{d(y)} \\ &= \sum_{y \in Y} \sum_{x \in N(y)} \frac{f(y)}{d(y)} = \sum_{y \in Y} \frac{f(y)}{d(y)} \cdot |N(y)| = \sum_{y \in Y} f(y). \quad \square \end{aligned}$$

Proof of Theorem 2.11. (Due to F. Galvin 1997). For a set of vertices $Y \subseteq V_2$, let $N(Y) := \bigcap_{y \in Y} N(y)$ denote the set of all its *common neighbors* in G, that is, the set of all those vertices in V_1 which are joined to each vertex of Y;

hence $|N(Y)| \leq r$ for all $Y \subseteq V_2$. Let $X \subseteq V_1$ be a minimal set with the property that $|N(Y) \setminus X| \leq r - |Y|$ whenever $Y \subseteq V_2$ and $1 \leq |Y| \leq r$. Put otherwise, X is a minimal set of vertices in V_1, the removal of which leads to a graph without $(r-a+1) \times a$ cliques, for all $a = 1, \ldots, r$.

Our goal is to show that $|X| \leq k$.

Note that, for each $x \in X$ we can choose $Y_x \subseteq V_2$ so that $1 \leq |Y_x| \leq r$, $x \in N(Y_x)$ and

$$|N(Y_x) \setminus X| = r - |Y_x|;$$

otherwise X could be replaced by $X \setminus \{x\}$, contradicting the minimality of X. We will apply Lemma 2.12 to the bipartite graph $G' = (X, V_2, F)$, where

$$F = \{(x, y) \ : \ y \in Y_x\}.$$

All we have to do is to show that the hypothesis of the lemma is satisfied by the function (here $N(y)$ is the set of neighbors of y in the original graph G):

$$f(y) := \frac{|N(y)|}{r},$$

because then

$$|X| \leq \sum_{y \in V_2} f(y) = \frac{1}{r} \sum_{y \in V_2} |N(y)| = \frac{|E|}{r} < k + 1.$$

Consider an edge $(x, y) \in F$; we have to show that $d(y) \leq d(x) \cdot f(y)$, where

$$d(x) = |Y_x| \ \text{ and } d(y) = |\{x \in X \ : \ y \in Y_x\}|$$

are the degrees of x and y in the graph $G' = (X, V_2, F)$. Now, $y \in Y_x$ implies $N(Y_x) \subseteq N(y)$, which in its turn implies

$$|N(y) \setminus X| \geq |N(Y_x) \setminus X| = r - |Y_x|;$$

hence

$$\begin{aligned} d(y) &\leq |N(y) \cap X| = |N(y)| - |N(y) \setminus X| \\ &\leq |N(y)| - r + |Y_x| = r \cdot f(y) - r + d(x), \end{aligned}$$

and so

$$\begin{aligned} d(x) \cdot f(y) - d(y) &\geq d(x) \cdot f(y) - r \cdot f(y) + r - d(x) \\ &= (r - d(x)) \cdot (1 - f(y)) \geq 0 \,. \quad \square \end{aligned}$$

2.5 Density of 0-1 matrices

Let H be an $m \times n$ 0-1 matrix. We say that H is α-*dense* if at least an α-fraction of all its mn entries are 1s. Similarly, a row (or column) is α-*dense* if at least an α-fraction of all its entries are 1s.

The next result says that any dense 0-1 matrix must either have one "very dense" row or there must be many rows which are still "dense enough."

Lemma 2.13 (Grigni and Sipser 1995). *If H is 2α-dense then either*
(a) *there exists a row which is $\sqrt{\alpha}$-dense, or*
(b) *at least $\sqrt{\alpha} \cdot m$ of the rows are α-dense.*

Note that $\sqrt{\alpha}$ is larger than α when $\alpha < 1$.

Proof. Suppose that the two cases do not hold. We calculate the density of the entire matrix. Since (b) does not hold, less than $\sqrt{\alpha} \cdot m$ of the rows are α-dense. Since (a) does not hold, each of these rows has less than $\sqrt{\alpha} \cdot n$ 1s; hence, the fraction of 1s in α-dense rows is strictly less than $(\sqrt{\alpha})(\sqrt{\alpha}) = \alpha$. We have at most m rows which are not α-dense, and each of them has less than αn ones. Hence, the fraction of 1s in these rows is also less than α. Thus, the total fraction of 1s in the matrix is less than 2α, contradicting the 2α-density of H. $\square$

Now consider a slightly different question: if H is α-dense, how many of its rows *or* columns are "dense enough"? The answer is given by the following general estimate due to Johan Håstad. This result appeared in the paper of Karchmer and Wigderson (1990) and was used to prove that the graph connectivity problem cannot be solved by monotone circuits of logarithmic depth.

Suppose that our universe is a Cartesian product $A = A_1 \times \cdots \times A_k$ of some finite sets $A_1, \ldots, A_k$. Hence, elements of A are strings $\boldsymbol{a} = (a_1, \ldots, a_k)$ with $a_i \in A_i$. Fix now a subset of strings $H \subseteq A$ and a point $b \in A_i$. The *degree* of b in H is the number $d_H(b) = |\{\boldsymbol{a} \in H : a_i = b\}|$ of strings in H whose i-th coordinate is b.

Say that a point $b \in A_i$ from the i-th set is *popular* in H if its degree $d_H(b)$ is at least a $1/2k$ fraction of the average degree of an element in A_i, that is, if

$$d_H(b) \geq \frac{1}{2k} \frac{|H|}{|A_i|} .$$

Let $P_i \subseteq A_i$ be the set of all popular points in the i-th set A_i, and consider the Cartesian product of these sets:

$$P := P_1 \times P_2 \times \cdots \times P_k .$$

Lemma 2.14 (Håstad). $|P| > \frac{1}{2}|H|$.

Proof. It is enough to show that $|H \setminus P| < \frac{1}{2}|H|$. For every non-popular point $b \in A_i$, we have that

$$|\{\boldsymbol{a} \in H \;:\; a_i = b\}| < \frac{1}{2k}\frac{|H|}{|A_i|}\,.$$

Since the number of non-popular points in each set A_i does not exceed the total number of points $|A_i|$, we obtain

$$\begin{aligned}|H \setminus P| &\leq \sum_{i=1}^{k}\sum_{b \notin P_i} |\{\boldsymbol{a} \in H \;:\; a_i = b\}| < \sum_{i=1}^{k}\sum_{b \notin P_i} \frac{1}{2k}\frac{|H|}{|A_i|} \\ &\leq \sum_{i=1}^{k}\frac{1}{2k}|H| = \frac{1}{2}|H|\,. \quad \square\end{aligned}$$

Corollary 2.15. *In any* 2α*-dense* 0-1 *matrix* H *either a* $\sqrt{\alpha}$*-fraction of its rows or a* $\sqrt{\alpha}$*-fraction of its columns (or both) are* $(\alpha/2)$*-dense.*

Proof. Let H be an $m \times n$ matrix. We can view H as a subset of the Cartesian product $[m] \times [n]$, where $(i,j) \in H$ iff the entry in the i-th row and j-th column is 1. We are going to apply Lemma 2.14 with $k = 2$. We know that $|H| \geq 2\alpha mn$. So, if P_1 is the set of all rows with at least $\frac{1}{4}|H|/|A_1| = \alpha n/2$ ones, and P_2 is the set of all columns with at least $\frac{1}{4}|H|/|A_2| = \alpha m/2$ ones, then Lemma 2.14 implies that

$$\frac{|P_1|}{m}\cdot\frac{|P_2|}{n} \geq \frac{1}{2}\frac{|H|}{mn} \geq \frac{1}{2}\cdot\frac{2\alpha mn}{mn} = \alpha\,.$$

Hence, either $|P_1|/m$ or $|P_2|/n$ must be at least $\sqrt{\alpha}$, as claimed. □

2.6 The Lovász–Stein theorem

This theorem was used by Stein (1974) and Lovász (1975) in studying some combinatorial covering problems. The advantage of this result is that it can be used to get existence results for some combinatorial problems using constructive methods rather than probabilistic methods.

Given a family $\mathcal{F}$ of subsets of some finite set X, its *cover number* of $\mathcal{F}$, $\mathrm{Cov}\,(\mathcal{F})$, is the minimum number of members of $\mathcal{F}$ whose union covers all points (elements) of X.

Theorem 2.16. *If each member of* $\mathcal{F}$ *has at most* a *elements, and each point* $x \in X$ *belongs to at least* v *of the sets in* $\mathcal{F}$*, then*

$$\mathrm{Cov}\,(\mathcal{F}) \leq \frac{|\mathcal{F}|}{v}(1 + \ln a)\,.$$

Proof. Let $N = |X|$, $M = |\mathcal{F}|$ and consider the $N \times M$ 0-1 matrix $A = (a_{x,i})$, where $a_{x,i} = 1$ iff $x \in X$ belongs to the i-th member of $\mathcal{F}$. By our assumption, each row of A has at least v ones and each column at most a ones. By double counting, we have that $Nv \geq Ma$, or equivalently,

$$\frac{M}{v} \leq \frac{N}{a}. \tag{2.7}$$

Our goal is to show that then A must contain an $N \times K$ submatrix C with no all-0 rows and such that

$$K \leq N/a + (M/v)\ln a \leq (M/v)(1 + \ln a).$$

We describe a constructive procedure for producing the desired submatrix C. Let $A_a = A$ and define A'_a to be any maximal set of columns from A_a whose supports[†] are pairwise disjoint and whose columns each have a ones. Let $K_a = |A'_a|$. Discard from A_a the columns of A'_a and any row with a one in A'_a. We are left with a $k_a \times (M - K_a)$ matrix A_{a-1}, where $k_a = N - aK_a$. Clearly, the columns of A_{a-1} have at most $a-1$ ones (indeed, otherwise such a column could be added to the previously discarded set, contradicting its maximality). We continue by doing to A_{a-1} what we did to A_a. That is we define A'_{a-1} to be any maximal set of columns from A_{a-1} whose supports are pairwise disjoint and whose columns each have $a-1$ ones. Let $K_{a-1} = |A'_{a-1}|$. Then discard from A_{a-1} the columns of A'_{a-1} and any row with a one in A'_{a-1} getting a $k_{a-1} \times (M - K_a - K_{a-1})$ matrix A_{a-2}, where $k_{a-1} = N - aK_a - (a-1)K_{a-1}$.

The process will terminate after at most a steps (when we have a matrix containing only zeros). The union of the columns of the discarded sets form the desired submatrix C with $K = \sum_{i=1}^{a} K_i$. The first step of the algorithm gives $k_a = N - aK_a$, which we rewrite, setting $k_{a+1} = N$, as

$$K_a = \frac{k_{a+1} - k_a}{a}.$$

Analogously,

$$K_i = \frac{k_{i+1} - k_i}{i} \qquad \text{for } i = 1, \ldots, a.$$

Now we derive an upper bound for k_i by counting the number of ones in A_{i-1} in two ways: every row of A_{i-1} contains at least v ones, and every column at most $i-1$ ones, thus

$$vk_i \leq (i-1)(M - K_a - \cdots - K_{i+1}) \leq (i-1)M,$$

or equivalently,

$$k_i \leq \frac{(i-1)M}{v}.$$

[†] The *support* of a vector is the set of its nonzero coordinates.

So,

$$
\begin{aligned}
K &= \sum_{i=1}^{a} K_i = \sum_{i=1}^{a} \frac{k_{i+1} - k_i}{i} \\
&= \frac{k_{a+1}}{a} + \frac{k_a}{a(a-1)} + \frac{k_{a-1}}{(a-1)(a-2)} + \cdots + \frac{k_2}{2 \cdot 1} - k_1 \\
&\le \frac{N}{a} + \frac{M}{v}\left(\frac{1}{a} + \frac{1}{a-1} + \cdots + \frac{1}{2}\right) \le \frac{N}{a} + \frac{M}{v} \ln a\,.
\end{aligned}
$$

The last inequality here follows because $1 + 1/2 + 1/3 + \cdots + 1/n$ is the n-th harmonic number which is known to lie between $\ln n$ and $\ln n + 1$. Together with (2.7), this yields $K \le (M/v)(1 + \ln a)$, as desired. □

The advantage of this proof is that it can be turned into a simple greedy algorithm which constructs the desired $N \times K$ submatrix A' with column-set C, $|C| = K$:

1. Set $C := \emptyset$ and $A' := A$.
2. While A' has at least one row do:
   - find a column c in A' having a maximum number of ones;
   - delete all rows of A' that contain a 1 in column c;
   - delete column c from A';
   - set $C := C \cup \{c\}$.

2.6.1 Covering designs

An (n, k, l) *covering design* is a family $\mathcal{F}$ of k-subsets of an n-element set (called *blocks*) such that every l-subset is contained in at least one of these blocks. Let $M(n, k, l)$ denote the minimal cardinality of such a design. A simple counting argument (Exercise 1.26) shows that $M(n, k, l) \ge \binom{n}{l}/\binom{k}{l}$.

In 1985, Rödl proved a long-standing conjecture of Erdős and Hanani that for fixed k and l, coverings of size $\binom{n}{l}/\binom{k}{l}(1 + o(1))$ exist. Rödl used non-constructive probabilistic arguments. We will now use the Lovász–Stein theorem to show how to *construct* an (n, k, l) covering design with only $\ln\binom{k}{l}$ times more blocks. This is not as sharp as Rödl's celebrated result, but it is constructive. A polynomial-time covering algorithm, achieving Rödl's bound, was found by Kuzjurin (2000).

Theorem 2.17. $M(n, k, l) \le \binom{n}{l}/\binom{k}{l}\left[1 + \ln\binom{k}{l}\right]$.

Proof. Let $X = (x_{S,T})$ be an $N \times M$ 0-1 matrix with $N = \binom{n}{l}$ and $M = \binom{n}{k}$. Rows of X are labeled by l-element subsets $S \subseteq [n]$, columns by k-element subsets $T \subseteq [n]$, and $x_{S,T} = 1$ iff $S \subseteq T$. Note that each row contains exactly $v = \binom{n-l}{k-l}$ ones, and each column contains exactly $a = \binom{k}{l}$ ones.

By the Lovász–Stein theorem, there is an $N \times K$ submatrix X' such that X' does not contain an all-0 row and

$$K \le (M/v)(1+\ln a) = \binom{n}{k} \Big/ \binom{n-l}{k-l} \left[1 + \ln \binom{k}{l}\right]$$
$$= \binom{n}{l} \Big/ \binom{k}{l} \left[1 + \ln \binom{k}{l}\right],$$

as $\binom{n}{l}\binom{n-l}{k-l} = \binom{n}{k}\binom{k}{l}$ (see Exercise 1.12). By the definition of X and the property of X' (no all-0 row), the k-subsets that correspond to the columns of X' form an (n,k,l) covering design. □

Exercises

2.1. Let $A_1, \ldots, A_m$ be subsets of an n-element set such that $|A_i \cap A_j| \le t$ for all $i \ne j$. Prove that $\sum_{i=1}^{m} |A_i| \le n + t \cdot \binom{m}{2}$.

2.2. Let $A = (a_{ij})$ be an $n \times n$ matrix ($n \ge 4$ even). The matrix is filled with integers and each integer appears exactly twice. Show that there exists a permutation π of $\{1, \ldots, n\}$ such that all the numbers $a_{i,\pi(i)}$, $i = 1, \ldots, n$ are distinct. (Such a permutation π is also called a *Latin transversal* of A.) *Hint*: Look at how many pairs of entries are "bad," i.e., contain the same number, and show that strictly less than $n!$ of all permutations can go through such pairs.

2.3. Let $\mathcal{F}$ be a family of m subsets of a finite set X. For $x \in X$, let $p(x)$ be the number of pairs (A, B) of sets $A, B \in \mathcal{F}$ such that either $x \in A \cap B$ or $x \notin A \cup B$. Prove that $p(x) \ge m^2/2$ for every $x \in X$. *Hint*: Let $d(x)$ be the degree of x in $\mathcal{F}$, and observe that $p(x) = d(x)^2 + (m - d(x))^2$.

2.4. Let $\mathcal{F}$ be a family of nonempty subsets of a finite set X that is closed under union (i.e., $A, B \in \mathcal{F}$ implies $A \cup B \in \mathcal{F}$). Prove or give a counterexample: there exists $x \in X$ such that $d(x) \ge |\mathcal{F}|/2$. (Open conjecture, due to Peter Frankl.)

2.5. A *projective plane* of order $r-1$ is a family of $n = r^2 - r + 1$ r-element subsets (called *lines*) of an n-element set of points such that each two lines intersect at precisely one point and each point belongs to precisely r lines (cf. Sect. 12.4). Use this family to show that the bound given by Corrádi's lemma (Lemma 2.1) is optimal.

2.6. Theorem 2.10 gives a sufficient condition for a bipartite graph with parts of the same size n to contain an $a \times a$ clique. Extend this result to not necessarily balanced graphs. Let $k_{a,b}(m,n)$ be the minimal integer k such that any bipartite graph with parts of size m and n and more than k edges contains at least one $a \times b$ clique. Prove that for any $0 \le a \le m$ and $0 \le b \le n$,

$$k_{a,b}(m,n) \le (a-1)^{1/b} n m^{1-1/b} + (b-1)m.$$

2.7. (Paturi–Zane 1998). Extend Theorem 2.10 to r-partite graphs as follows. An r-partite m-clique is a Cartesian product $V_1 \times V_2 \times \cdots \times V_r$ of m-element sets $V_1, \ldots, V_r$. An r-partite graph with parts of size m is a subset E of an r-partite m-clique. Let $\mathrm{ex}(m, r, 2)$ denote the maximum size $|E|$ of such a graph E which does not contain an r-partite 2-clique. Erdős (1959, 1964b) proved that

$$cm^{r-r/2^{r-1}} \le \mathrm{ex}(m, r, 2) \le m^{r-1/2^{r-1}},$$

where $c = c(r) > 0$ is a constant depending only on r. A slightly weaker upper bound $\mathrm{ex}(m, r, 2) < 2m^{r-1/2^{r-1}}$ can be derived from Lemma 2.2. Show how to do this. *Hint*: Argue by induction on r. For the induction step take $X = V_1 \times \cdots \times V_{r-1}$ and consider m subsets $A_v = \{x \in X : (x, v) \in E\}$ with $v \in V_r$. Apply Lemma 2.2 with $n = m^{r-1}$, $N = m$ and $w = \frac{1}{2}m^{1/2^{r-1}}$, to obtain a pair of points $u \neq v \in V_k$ for which the graph $E' = A_u \cap A_v$ is large enough, and use the induction hypothesis.

2.8. Let $\mathcal{F} = \{A_1, \ldots, A_N\}$ be a family of subsets of some set X. Use (1.11) to prove that for every $1 \le s \le N$,

$$\sum_{x \in X} d(x)^s = \sum_{(i_1, i_2, \ldots, i_s)} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_s}|,$$

where the last sum is over all s-tuples $(i_1, i_2, \ldots, i_s)$ of (not necessarily distinct) indices.

2.9. Use the previous exercise and the argument of Lemma 2.2 to prove Lemma 2.3.

2.10. Let $A_1, \ldots, A_N$ be subsets of some n-element set X, and suppose that these sets have average size at least αn. Show that for every $s \le (1 - \epsilon)\alpha N$ with $0 < \epsilon < 1$, there are indices $i_1, i_2, \ldots, i_s$ such that

$$|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_s}| \ge (\epsilon\alpha)^s n.$$

Hint: Consider the bipartite graph $G = (X, V, E)$ where $V = \{1, \ldots, N\}$, and $(x, i) \in E$ if and only if $x \in A_i$. Observe that $|E| \ge \alpha n N$ and argue as in the proof of Theorem 2.10.

2.11. Prove the following very useful averaging principle for partitions. Let $X = A_1 \cup A_2 \cup \cdots \cup A_m$ be a partition of a finite set X into m mutually disjoint sets (blocks), and $a = \sum_{i=1}^{m} |A_i|/m$ be the average size of a block in this partition. Show that for every $1 \le b \le a$, at least $(1 - 1/b)|X|$ elements of X belong to blocks of size at least a/b. How many elements of X belong to blocks of size at most ab? *Hint*: $m \cdot (a/b) \le |X|/b$.

2.12. Let $A_1, \ldots, A_r$ be a sequence of (not necessarily distinct) subsets of an n-element set X such that each set has size n/s and each element $x \in X$ belongs to least one and to at most k of them; hence $r \le ks$. Let $K := \sum_{i=0}^{k} \binom{r}{i}$ and assume that $s > 2k$. Prove that there exist two disjoint subsets X_1 and X_2 of X such that $|X_i| \ge n/(2K)$ for both $i = 1, 2$, and none of the sets $A_1, \ldots, A_r$ contains points from both sets X_1 and X_2. *Hint*: Associate with each $x \in X$ its *trace* $T(x) = \{i : x \in A_i\}$ and partition the elements of X according to their traces. Use the previous exercise to show that at least $n/2$ elements belong to blocks of size at least $n/(2K)$. Show that some two of these elements x and y must have *disjoint* traces, $T(x) \cap T(y) = \emptyset$.

2.13. Let $X = A_1 \cup A_2 \cup \cdots \cup A_m$ be a partition of a finite set X into mutually disjoint blocks. Given a subset $Y \subseteq X$, we obtain its partition $Y = B_1 \cup B_2 \cup \cdots \cup B_m$ into blocks $B_i = A_i \cap Y$. Say that a block B_i is λ-*large* if $|B_i|/|A_i| \geq \lambda \cdot |Y|/|X|$. Show that, for every $\lambda > 0$, at least $(1-\lambda)\cdot|Y|$ elements of Y belong to λ-large blocks.

2.14. Given a family $S_1, \ldots, S_n$ of subsets of $V = \{1, \ldots, n\}$, its *intersection graph* $G = (V, E)$ is defined by: $\{i, j\} \in E$ if and only if $S_i \cap S_j \neq \emptyset$. Suppose that: (i) the sets have average size at least r, and (ii) the average size of their pairwise intersections does not exceed k. Show that $|E| \geq \frac{n}{k} \cdot \binom{r}{2}$. *Hint*: Consider the sum $\sum_{i<j} |S_i \cap S_j|$.

2.15. Let H be a 2α-dense 0-1 matrix. Prove that at least an $\alpha/(1-\alpha)$ fraction of its rows must be α-dense.

2.16. (Alon 1986). Let S be a set of strings of length n over some alphabet. Suppose that every two strings of S differ in at least d coordinates. Let k be such that $d > n(1 - 1/\binom{k}{2})$. Show that any k distinct strings $v_1, \ldots, v_k$ of S attain k distinct values in at least one coordinate. *Hint*: Assume the opposite and count the sum of distances between the $\binom{k}{2}$ pairs of v_i's.

http://www.springer.com/978-3-642-17363-9

Extremal Combinatorics
With Applications in Computer Science
Jukna, S.
2011, XXIV, 412 p., Hardcover
ISBN: 978-3-642-17363-9