Semidefinite approximations for bicliques and biindependent pairs

Monique Laurent[1,2], Sven Polak[1], Luis Felipe Vargas[1]

[1] Centrum Wiskunde & Informatica
[2] Tilburg University
{m.laurent, s.c.polak, luis.vargas}@cwi.nl

Abstract. A *(bipartite) biindependent pair* in a bipartite graph $G = (V_1 \cup V_2, E)$ is a pair (A, B), where $A \subseteq V_1$, $B \subseteq V_2$ and the union $A \cup B$ is independent in G. We investigate the following two parameters $g(G)$ and $h(G)$, that are defined, respectively, as the maximum product $|A|\cdot|B|$ and the maximum ratio $\frac{|A|\cdot|B|}{|A|+|B|}$, taken over all such biindependent pairs (A, B) in G. These parameters have many applications, in particular, for bounding maximum product-free subsets in groups and the nonnegative rank of a matrix. We define semidefinite programming upper bounds on $g(G)$ and $h(G)$. We show they can be seen as quadratic variations of the Lovász ϑ-number, a well-known upper bound on $\alpha(G)$, equal to it for G bipartite. We also show links among them as well as with an earlier parameter by Haemers. In addition we formulate closed-form eigenvalue bounds, which coincide with the semidefinite bounds for edge-transitive graphs.

Keywords: bicliques, independent sets, semidefinite programming

Biindependent pairs and bicliques

Given a graph $G = (V, E)$, a *biindependent pair* in G is a pair (A, B) of disjoint subsets of V such that no pair of nodes $\{i, j\} \in A \times B$ is an edge of G. When G is bipartite, with bipartition $V = V_1 \cup V_2$, one is also interested in *bipartite biindependent pairs*, which means satisfying $A \subseteq V_1$ and $B \subseteq V_2$. The maximum cardinality $|A| + |B|$ of such a biindependent pair is $\alpha(G)$, the stability number of G, well-known to be computable in polynomial time using matching algorithms. We consider the following two other parameters, asking for the maximum product $|A| \cdot |B|$ and the maximum ratio $\frac{|A|\cdot|B|}{|A|+|B|}$:

$$g(G) = \max\{|A| \cdot |B| : A \subseteq V_1, B \subseteq V_2, (A, B) \text{ is a biindependent pair in } G\},$$

$$h(G) = \max\left\{\frac{|A|\cdot|B|}{|A|+|B|} : A \subseteq V_1, B \subseteq V_2, , (A, B) \text{ is a biindependent pair in } G\right\}.$$

While computing the parameter $g(G)$ has been shown to be NP-hard by Peeters [9], the exact complexity status of the parameter $h(G)$ is still unknown.

The parameter $h(G)$ was introduced by Vallentin [10], who observed its relevance to maximum product-free subsets in groups in work of Gowers [4]. The parameter $g(G)$ has many applications, in particular, to bounding the rectangle covering number and the nonnegative rank of nonnegative matrices.

Semidefinite and eigenvalue-based bounds

The parameters $g(G)$ and $h(G)$ can be formulated as polynomial optimization problems, which leads to hierarchies of semidefinite programming (SDP) upper bounds, able to find the original parameters in finitely many steps (in fact, in $\alpha(G)$ steps). We investigate in particular the SDP bounds obtained at the lowest level, which take the form

$$h_1(G) = \max_{X \in \mathcal{S}^V} \{\langle C, X\rangle : \mathrm{Tr}(X) = 1,\ X_{ij} = 0\ (\{i,j\} \in E),\ X \succeq 0\}, \tag{1}$$

$$g_1(G) = \max_{X \in \mathcal{S}^V} \left\{\langle C, X\rangle : \begin{pmatrix} 1 & \mathrm{diag}(X)^\mathsf{T} \\ \mathrm{diag}(X) & X \end{pmatrix} \succeq 0,\ X_{ij} = 0 \text{ if } \{i,j\} \in E\right\}, \tag{2}$$

setting $C := \frac{1}{2}\begin{pmatrix} 0 & J_{V_1,V_2} \\ J_{V_1,V_2}^\mathsf{T} & 0 \end{pmatrix}$, where J_{V_1,V_2} is the all-ones matrix in $\mathbb{R}^{V_1 \times V_2}$. These two bounds can be seen as quadratic variations of the parameter $\vartheta(G)$, introduced by Lovász [8] as upper bound on $\alpha(G)$ for any G (and equal to $\alpha(G)$ when G is bipartite). Indeed, if we replace the objective $\langle C, X\rangle$ by $\langle J, X\rangle$ in program (1) and by $\mathrm{Tr}(X)$ in program (2), then we obtain $\vartheta(G)$.

We show the following relations between the parameters $h(G)$, $g(G)$, $h_1(G)$, $g_1(G)$, and $\alpha(G)$ for any bipartite graph G.

Proposition 1. *For any bipartite graph G we have*

$$h(G) \le \tfrac{1}{2}\sqrt{g(G)} \le h_1(G) \le \tfrac{1}{2}\sqrt{g_1(G)} \le \tfrac{1}{4}\alpha(G).$$

When G is r-regular we can give eigenvalue-based closed-form upper bounds.

Proposition 2. *Assume G is bipartite r-regular, set $n := |V_1| = |V_2|$, and let λ_2 be the second largest eigenvalue of the adjacency matrix of G. Then*

$$h_1(G) \le \widehat{h}(G) := \frac{n\lambda_2}{2(\lambda_2 + r)}, \qquad g_1(G) \le \widehat{g}(G) := \begin{cases} \frac{n^2\lambda_2^2}{(\lambda_2+r)^2} & \text{if } r \le 3\lambda_2, \\ \frac{n^2\lambda_2}{8(r-\lambda_2)} & \text{otherwise,} \end{cases}$$

with equality $h_1(G) = \widehat{h}(G)$ and $g_1(G) = \widehat{g}(G)$ when G is edge-transitive.

Observe that the upper bound $\widehat{h}(G)$ sharpens the bound $h(G) \le \frac{n}{r}\lambda_2$ from [10].

Application to biindependent pairs and bicliques in arbitrary graphs

One may also consider bindependent pairs in an arbitrary graph G (not necessarily bipartite). However, they correspond to the bipartite biindependent pairs in an associated bipartite graph $B_0(G)$, whose node set is $V \cup V'$, where $V' = \{i' : i \in V\}$ is a disjoint copy of V, and whose edges are the pairs $\{i, i'\}$ $(i \in V)$, $\{i, j'\}$ and $\{j, i'\}$ for $\{i, j\} \in E$. Indeed, a pair (A, B) is biindependent in G precisely when $(A, B' := \{i' : i \in B\})$ is biindependent in $B_0(G)$ (with $A \subseteq V$ and $B' \subseteq V'$). Hence, the maximum product $|A| \cdot |B|$ and ratio $\frac{|A|\cdot|B|}{|A|+|B|}$, for biindependent pairs in G, are captured by the parameters $g(B_0(G))$ and $h(B_0(G))$

for the bipartite graph $B_0(G)$. So we obtain hierarchies of SDP bounds also for these parameters. Interestingly, the SDP bound $h_1(B_0(G))$ recovers an earlier parameter introduced by Haemers [5]. Finally, one can also model bicliques in any graph G, i.e., the pairs (A,B) of disjoint vertex subsets with $A \times B \subseteq E$, since they correspond to the biindependent pairs in the complementary graph $\overline{G} = (V, \overline{E})$.

Applications

We now briefly describe two applications of the parameters $g(G)$ and $h(G)$.

Let $M \in \mathbb{R}^{V_1 \times V_2}$ be a nonnegative matrix. Its *nonnegative rank* $\mathrm{rank}_+(M)$ is the smallest integer r for which there exist nonnegative vectors $a_\ell \in \mathbb{R}_+^{V_1}$ and $b_\ell \in \mathbb{R}_+^{V_2}$ ($\ell \in [r]$) such that $M = \sum_{\ell=1}^r a_\ell b_\ell^T$. The nonnegative rank is an important parameter, which is hard to compute [11]. Hence one needs good bounds for it. One such bound is provided by the *rectangle covering bound* $\mathrm{rc}(M)$, defined as the smallest number of admissible rectangles $A \times B \subseteq V_1 \times V_2$ needed to cover the support $S_M := \{(i,j) : M_{ij} \neq 0\}$ of M. Here $A \times B \subseteq V_1 \times V_2$ is an *admissible rectangle* if $A \times B \subseteq S_M$. Then we have $\mathrm{rc}(M) \leq \mathrm{rank}_+(M)$. The rectangle covering bound can be very useful; it was used, e.g., in [2] to show an exponential lower bound on the extension complexity of combinatorial polytopes such as the traveling salesman and correlation polytopes.

Also the parameter $\mathrm{rc}(M)$ is not easy to compute. To approximate it, one can consider the bipartite graph B_M, with vertex set $V_1 \cup V_2$ and edge set $E_M := (V_1 \times V_2) \setminus S_M$. Then admissible rectangles for M correspond precisely to biindependent pairs in B_M and one can show that

$$\mathrm{rc}(M) \cdot g(B_M) \geq |S_M|.$$

Hence, an upper bound on $g(B_M)$ gives directly a lower bound on $\mathrm{rc}(M)$ and thus a lower bound on the nonnegative rank $\mathrm{rank}_+(M)$.

The parameter $h(G)$ is useful for bounding the maximum size of a sum-free subset in a group. Let Γ be a finite group. Then a set $A \subseteq \Gamma$ is called *sum-free* if $ab \notin A$ for all $a, b \in A$, and one is interested in finding a largest such set (see [4, 7] for background on this problem).

Given $A \subseteq \Gamma$, let $G_A = (V_1 \cup V_2, E)$ be the associated bipartite Caley graph, with $V_1 = V_2 = \Gamma$ and $\{x, y\} \in E$ if and only if $y = ax$ for some $a \in A$. If A is sum-free, then (A, A) is a biindependent pair in G_A and thus we have $\frac{|A|}{2} = \frac{|A| \cdot |A|}{2|A|} \leq h(G)$. Hence, upper bounds on $h(G_A)$ give rise to upper bounds on sum-free subsets in Γ. In this way, Vallentin [10] could recover a result by Gowers [4]. Note that for this application we are in fact only interested in *balanced* biindependent pairs, i.e., with $|A| = |B|$. This motivates considering the analogues of the parameters $\alpha(G)$, $g(G)$ and $h(G)$, where one restricts the optimization to balanced pairs. The resulting parameters are equal (up to scaling) and hard to compute [3]. The complexity of determining whether a bipartite graph admits a balanced maximum stable set remains unknown. However, hardness of this problem would imply hardness of computing the parameter $h(G)$.

References

1. S. Fiorini, V. Kaibel, K. Pashkovich, D.O. Theis. Combinatorial bounds on nonnegative rank and extended formulations. *Discrete Mathematics*, 313:67–83, 2013.
2. S. Fiorini, S. Massar, S. Pokutta, H.R. Tiwary, R. de Wolf. Linear vs. semidefinite extended formulations: exponential separation and strong lower bounds. In: Proc. STOC 2012, 95–106, 2012.
3. M.R. Garey and D.S. Johnson. *Computers and Intractability. A Guide to the Theory of NP-Completeness,* Freeman, San Francisco, 1979.
4. W.T. Gowers, Quasirandom groups, *Combinatorics, Probability and Computing*, 17:363–387, 2008.
5. W.H. Haemers. Bicliques and eigenvalues. *Journal of Combinatorial Theory, Series B*, 82:56–66, 2001.
6. W.H. Haemers. Disconnected vertex sets and equidistant code pairs. *The Electronic Journal of Combinatorics*, 4(1), Research paper R7, 10 pages, 1997.
7. K.S. Kedlaya, Product-free subsets of groups, then and now, *Contemporary Mathematics, American Mathematical Society*, 479:169–177, 2009. .
8. L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25:1–7, 1979.
9. R. Peeters. The maximum edge biclique problem is NP-hard. *Discrete Applied Mathematics*, 131:651–654, 2003.
10. F. Vallentin, Semidefinite programming bounds for product free subsets in groups. See presentation at http://poema-network.eu/meeting/workshop-2-november-2020.
11. S. Vavasis, On the complexity of nonnegative matrix factorization, *SIAM Journal on Optimization*, 20:1364–1377, 2009.