Teaching dimension, VC dimension, and critical sets in Latin squares

Yingjie Qian (Joint Work with Hamed Hatami)

Dept. of Mathematics and Statistics, McGill University, Canada

yingjie.qian@mail.mcgill.ca

Latin squares and critical sets

- A **Latin square** of order n is an $n \times n$ array filled with elements from the set $\{1, 2, \ldots, n\}$ such that every element occurs exactly once in each row and each column.
- A **critical set** is the minimal set of entries in a Latin square L that uniquely identifies it among all Latin squares of order n.

1			
	2		
	4	2	

Conjecture 1 ([Nel][Mah95][BvR99]). *Every critical set for a Latin square of order n is of size at least* $\lfloor n^2/4 \rfloor$.

VC, teaching dimensions

Consider a finite set Ω, and let $\mathcal{F}(\Omega)$ denote the power set of Ω. In computational learning theory, a subset $\mathcal{C} \subseteq \mathcal{F}(\Omega)$ is refered to as a **concept class**, and the elements $c \in C$ are called **concepts**.

- A subset $S \subseteq \Omega$ is said to be **shattered** by $\mathcal{C}$ if for every $T \subseteq S$ there exists a concept c with $c \cap S = T$.
- The **VC-dimension** of $\mathcal{C}$ is the size of the largest set shattered by $\mathcal{C}$.
- A set $S \subseteq \Omega$ is called a **teaching set** for a concept $c \in \mathcal{C}$ if $(c \cap S) \neq (c' \cap S)$ for every concept $c' \neq c$. Let $\mathrm{TD}(c; \mathcal{C})$ denote the smallest size of a teaching set for a concept $c \in \mathcal{C}$.
- The **teaching dimension** and **minimum teaching dimension** are defined as:

$$\mathrm{TD}(\mathcal{C}) = \max_{c \in \mathcal{C}} \mathrm{TD}(c; \mathcal{C})$$

$$\mathrm{TD}(\mathcal{C}) = \min_{c \in \mathcal{C}} \mathrm{TD}(c; \mathcal{C})$$

- The **recursive teaching dimension** is:

$$\mathrm{RTD}(\mathcal{C}) = \max_{\mathcal{C}' \subseteq \mathcal{C}} \mathrm{TD}_{\min}(\mathcal{C}').$$

Results

Theorem 2 (Main theorem). *For sufficiently large n, every critical set for a Latin square of order n is of size at least* $10^{-4} n^2$.

Theorem 3. *The VC-dimension of the class of Latin squares of order n is at least* $n^2 - (e + o(1))n^{5/3}$.

Theorem 4. *The recursive teaching dimension of the class of Latin squares of order n is at least* $n^2 - (e + o(1))n^{5/3}$.

Previous Work on Critical sets

- Fu, Fu, and Rodger [FFR97] showed a lower-bound of $\lfloor (7n-3)/6 \rfloor$ for $n \geq 20$.
- Horak, Aldred, and Fleischner [FFR97] improved it to $\lfloor (4n-8)/3 \rfloor$ for $n \geq 8$.
- Finally in 2007, Cavenagh [Cav07] gave the first superlinear lower-bound by showing that the size of a critical set in a Latin square is at least $n\lfloor (\log n)^{1/3}/2 \rfloor$.
- Our Theorem 2 is the first to establish that as it was predicted in Conjecture 1, the size of the smallest critical set is of quadratic order.

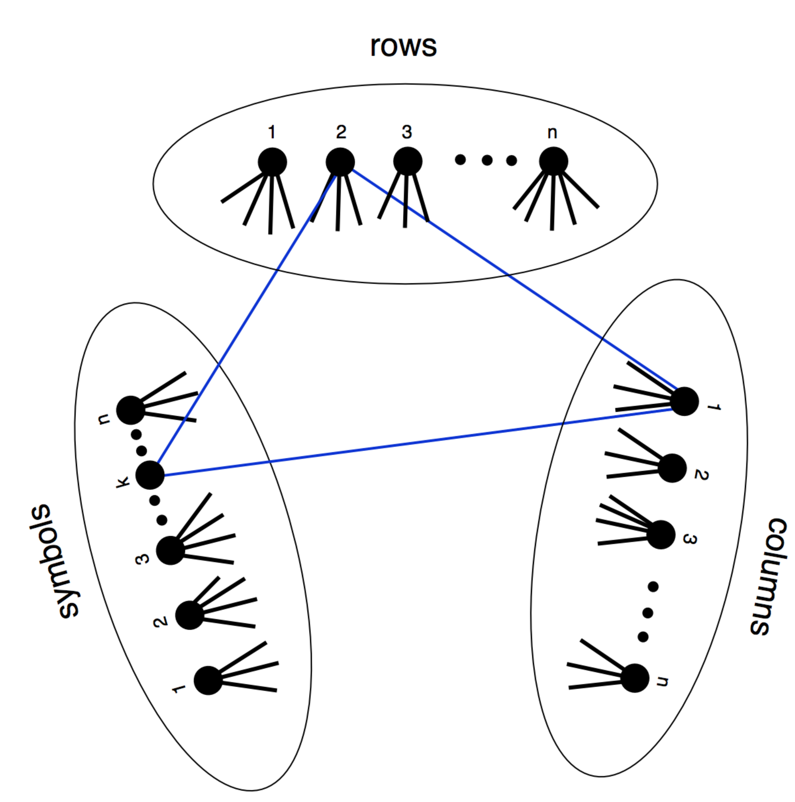
However, determining the exact coefficient and resolving Conjecture 1 remain unsolved.

Proof of Theorem 2

Lemma 5 (See [BKL+16, Corollary 1.6] and [Duk15, Theorem 1.3]). *Let $\gamma > 0$ and $n > n_0(\gamma)$. Every balanced and locally balanced 3-partite graph on $3n$ vertices with minimum degree at least $(101/52 + \gamma)n$, admits a K_3-decomposition.*

Set $\epsilon = 10^{-4}$ and $\lambda = 0.012$.

O		X	X	⋯ XX ⋯		X
k	O	X		X ⋯ ⋯		
	X		X	XX ⋯ ⋯	X	
	X	X		⋯ X ⋯		O
:	:	:	:	⋯		
:	X	O	:	⋯		
:	X	X	:	⋯		
:	:	:	:	⋯		
X				⋯	O	



Proof of Theorem 3 and 4

Lemma 6 ([vLW92, Theorem 17.2]). *Let $\mathcal{L}_n$ be the set of all Latin squares of order n. Then*

$$|\mathcal{L}_n| \geq \frac{(n!)^{2n}}{n^{n^2}}.$$

Lemma 7 ([GHM05, Theorem 3]). *Let $\mathcal{T}_{n,k}$ be the set of all partial Latin squares of order n and of size k. Then*

$$|\mathcal{T}_{n,k}| \leq \binom{n^2}{k} \frac{n!^{2n-\frac{k}{n}} e^{n(3+\frac{\ln(2\pi n)^2}{4})}}{(n-\frac{k}{n})!^{2n} e^k}.$$

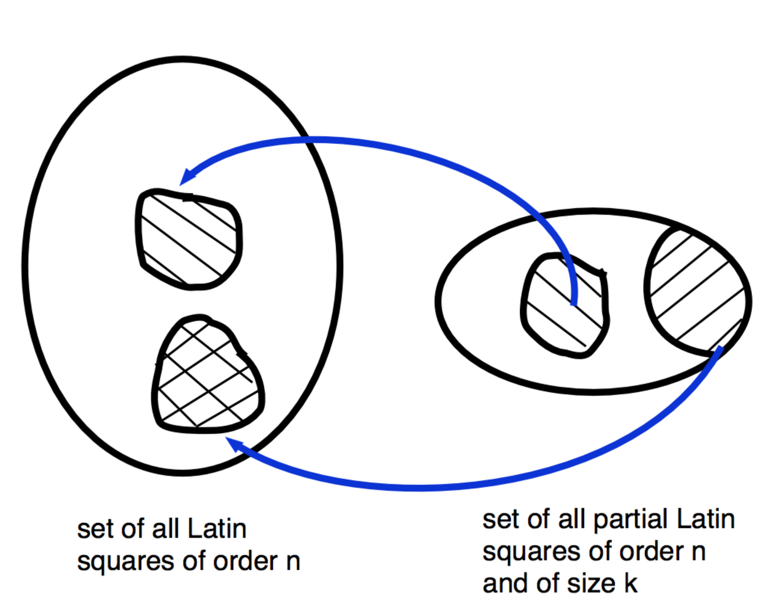
Lemma 8 ([Paj85]). *Every finite set family $\mathcal{F}$ shatters at least $|\mathcal{F}|$ sets.*

Sketch of Proof of Theorem 3:

For $d = VC(\mathcal{L}_n)$, we have $\sum_{k=0}^{d-1} |\mathcal{T}_{n,k}| < |\mathcal{L}_n|$.

So it suffices to show that for every $k \leq n^2 - (e + o(1))n^{5/3}$, we have $|\mathcal{T}_{n,k}| < \frac{|\mathcal{L}_n|}{n^2}$. Then, use similar calculations in Ghandehari, Hatami and Mahmoodian [GHM05].

Sketch of Proof of Theorem 4:



Remarks on Theorem 3 and 4

In Theorems 3 and 4 we established a lower bound of $n^2 - (e + 0(1))n^{5/3}$ for both VC-dimension and the recursive teaching dimension of the class of Latin squares of order n. One can easily obtain an upper bound of the form $n^2 - \Omega(n)$ for VC-dimension, but obtaining a stronger upper-bound and more ambitiously determining the exact asymptotics of the VC-dimension seems difficult. For the teaching dimension and consequently recursive teaching dimension, a stronger upper-bound of $n^2 - \frac{\sqrt{\pi}}{2} n^{3/2}$ follows from the results of [GHM05]. Hence

$$n^2 - (e + 0(1))n^{5/3} \leq \mathrm{RTD}(\mathcal{L}_n) \leq \mathrm{TD}(\mathcal{L}_n) \leq n^2 - \frac{\sqrt{\pi}}{2} n^{3/2}.$$

It would be very interesting to improve either of the constants $5/3$ and $3/2$ appearing in the power of n in the above bounds.

References

[BKL+16] Ben Barber, Daniela Kuhn, Allan Lo, Deryk Osthus, and Amelia Taylor. Clique decomposition of multipartite graphs and completion of latin squares. *arXiv preprint arXiv:1603.01043*, 2016.

[BvR99] John A. Bate and G. H. John van Rees. The size of the smallest strong critical set in a Latin square. *Ars Combin.*, 53:73–83, 1999.

[Cav07] Nicholas J. Cavenagh. A superlinear lower bound for the size of a critical set in a Latin square. *J. Combin. Des.*, 15(4):269–282, 2007.

[Duk15] Peter J. Dukes. Fractional triangle decompositions of dense 3-partite graphs. *arXiv preprint arXiv:1510.08998*, 2015.

[FFR97] Chin-Mei Fu, Hung-Lin Fu, and C. A. Rodger. The minimum size of critical sets in Latin squares. *J. Statist. Plann. Inference*, 62(2):333–337, 1997.

[GHM05] Mahya Ghandehari, Hamed Hatami, and Ebadollah S. Mahmoodian. On the size of the minimum critical set of a Latin square. *Discrete Math.*, 293(1-3):121–127, 2005.

[Mah95] Ebadollah S. Mahmoodian. Some problems in graph colorings. In *Proceedings of the 26th Annual Iranian Mathematics Conference, Vol. 2 (Kerman, 1995)*, pages 215–218. Shahid Bahonar Univ. Kerman, Kerman, 1995.

[Nel] John Nelder. Private communication to Jennifer Seberry, 1979.

[Paj85] Alain Pajor. *Sous-espaces l_1^n des espaces de Banach*, volume 16 of *Travaux en Cours [Works in Progress]*. Hermann, Paris, 1985. With an introduction by Gilles Pisier.

[vLW92] Jacobus H. van Lint and Richard M. Wilson. *A course in combinatorics*. Cambridge University Press, Cambridge, 1992.