# Bounds on the Norms of Locally Random Matrices

Dhruv Medarametla
Mentor: Aaron Potechin

PRIMES Conference
May 16, 2015

# Matrices

## Definition

A $m \times n$ matrix $M$ is defined to be a rectangular array of numbers consisting of $m$ rows and $n$ columns; $M^T$ is an $n \times m$ matrix consisting of $M$ flipped over its main diagonal.

$$M = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

$$M^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

# Matrix Multiplication

## Definition

The matrix product $M = AB$ exists if and only if $A$ is an $m \times n$ matrix and $B$ is an $n \times p$ matrix, in which case

$$M_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} (1 \cdot 1 + 2 \cdot 1) & (1 \cdot -1 + 2 \cdot 0) \\ (0 \cdot 1 + 1 \cdot 1) & (0 \cdot -1 + 1 \cdot 0) \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 1 & 0 \end{bmatrix}$$

# Matrix Norms

- We focus on the **spectral norm**.
- The norm of a vector $V = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$, denoted by $||V||$, is defined as $||V|| = \sqrt{{a_1}^2 + {a_2}^2 + \cdots + {a_n}^2}$.

**Definition**

Given an $m \times n$ matrix $A$, the maximum possible value of $||AV||$ is defined to be $||A||$, where $V$ is an $n$-dimensional vector with $||V|| = 1$.

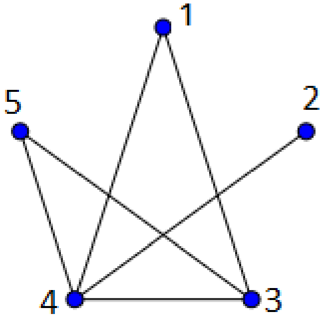
# Properties of the Spectral Norm, Eigenvalues, and Traces

- $||A+B|| \leq ||A|| + ||B||$.
- $||AB|| \leq ||A|| \cdot ||B||$.
- The trace of a $n \times n$ matrix $M$ is the sum of the elements on the main diagonal of $M$.
- If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the $n$ eigenvalues of a matrix $M$, then $\mathrm{tr}(M^k) = \lambda_1^k + \lambda_2^k + \cdots + \lambda_n^k$.
- If $\lambda_{\max}(M)$ is the largest eigenvalue of a matrix $M$, then $||A|| = \sqrt{\lambda_{\max}(AA^T)}$.
- Therefore,

$$\lim_{k \to \infty} \sqrt[2k]{\mathrm{tr}((AA^T)^k)} \geq ||A||.$$

# A Bipartite Graph and a Random Graph

- Consider a bipartite graph $H$ with partite sets $U = \{u_1, u_2, \ldots, u_k\}$ and $V = \{v_1, v_2, \ldots, v_l\}$.
- In addition, let $G$ be a graph with $n$ vertices labeled from 1 to $n$ chosen randomly from the set of all possible graphs on $n$ vertices labeled from 1 to $n$.

# $M$'s Entries

## Definition

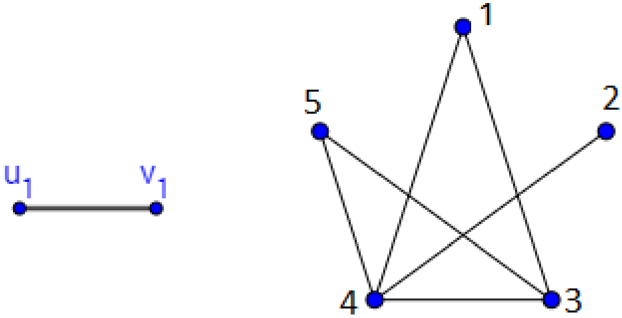
The locally random matrix of $H$, denoted by $M$, is a $\binom{n}{k} \times \binom{n}{l}$ matrix where each row or column corresponds to a sequence of $k$ or $l$ distinct integers from 1 to $n$ respectively.

The entry in row $A$, which corresponds to the sequence $\{a_1, a_2, \ldots a_k\}$, and column $B$, which corresponds to the sequence $\{b_1, b_2, \ldots b_l\}$, is

$$(-1)^{E(A,B)}$$

if $A$ and $B$ are disjoint, where $E(A, B)$ is the number of pairs $(i, j)$ such that edge $(a_i, b_j)$ exists in $G$ and edge $(u_i, v_j)$ exists in $H$. If $A$ and $B$ are not disjoint, the entry is 0.

# An Example of $M$




$$M = \begin{bmatrix} 0 & 1 & -1 & -1 & 1 \\ 1 & 0 & 1 & -1 & 1 \\ -1 & 1 & 0 & -1 & -1 \\ -1 & -1 & -1 & 0 & -1 \\ 1 & 1 & -1 & -1 & 0 \end{bmatrix}$$

# Motivation

- The sum of squares hierarchy is one of the most powerful tools we know of for approximation algorithms. However, its performance is not well understood.
- Analyzing the norms of these matrices is needed for analyzing the performance of the sum of squares hierarchy, especially on the planted clique problem.

# Example 1

- Simplest case: $H$ is simply an edge between two vertices.
- Then, $M$ is simply an $n \times n$ matrix, in which the entry in row $i$ and column $j$ is 1 if edge $(i, j)$ does not exist in $G$ and $-1$ otherwise.
- This case has been extensively studied, and the answer is already known due to Wigner's Semicircle Law.

# Bounding

- Recall $||M|| \leq \sqrt[2k]{\mathrm{tr}((MM^T)^k)} = \sqrt[2k]{\mathrm{tr}(M^{2k})}$.
- We upper bound the expected value of $||M||$ by upper bounding the expected value of $\mathrm{tr}(M^{2k})$.
- To do so, we create an interdependence graph.

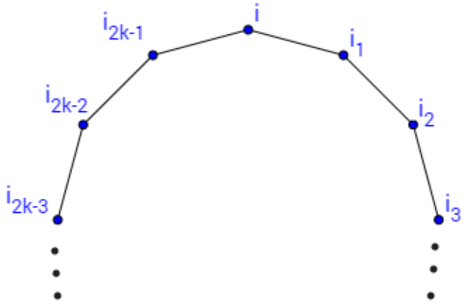
# Interdependence Graphs

- Note that the following is true:

$$(M)^{2k}_{ii} = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_{2k-1}=1}^{n} (M)_{ii_1} (M)_{i_1 i_2} \cdots (M)_{i_{2k-2} i_{2k-1}} (M)_{i_{2k-1} i}$$

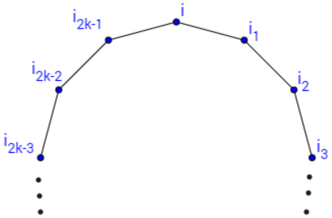
.

- We can represent each product in this sum with the following interdependence graph:



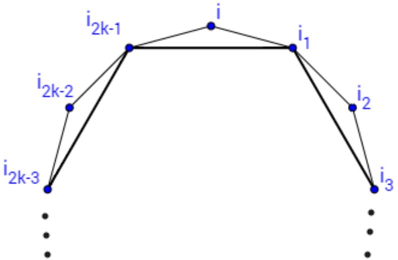
# Interdependence Graphs Cont.

- Therefore, $\mathrm{tr}((MM^T)^k)$ is equal to the sum of all possible products created by the interdependence graphs.
- Note that the expected value of most products is zero; we want to find the number of interdependence graphs that have a nonzero expected value for a product.
- Because the expected value for a given interdependence graph is at most 1, we can bound the expected value of the trace by the number of interdependence graphs with nonzero expected value.



# Interdependence Graphs Cont.

- We count the minimum number of constraints, which we represent by thicker lines, required to create a nonzero expected value.
- Any term $A_{i_a i_{a+1}}$ that does not have a counterpart automatically reduces the expected value of the product to zero.
- Therefore, all edges must have counterparts.



# Proposition

## Theorem

*Any interdependency graph that is a cycle with $2k$ vertices requires at least $k-1$ equalities.*

## Proof.

- Prove by induction and casework.
- Obviously, when $k = 2$, it is true.
- If every vertex is equal to at least one other, there are at least $k$ equalities.
- If one vertex is not equal to any others, its two adjacent vertices are equal, reducing the problem to the Induction Hypothesis.

□

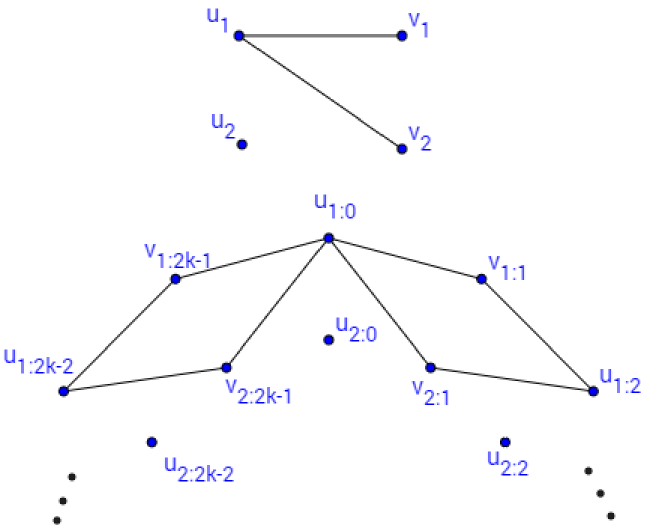
# Example 1's Solution

- The previous theorem implies that at least $k-1$ equalities are required, so there are at most $k+1$ independencies; therefore, $\mathrm{tr}((MM^T)^k)$ is at most some constant factor times $n^{k+1} = n^k \cdot n$.
- Therefore, the expected norm of this matrix is at most $(n^k * n)^{\frac{1}{2k}} = \sqrt{n}\log(n)$.
- In reality, the norm of this particular type of random matrix can be shown to be $\sqrt{n}$ by Wigner's Semicircle Law.

# Generalizing for all H

- Consider an arbitrary $H$; what can be said for the norm?
- Through generalizations of the previous theorems, we can actually bound this norm in all such cases.
- Once again, consider the interdependency graph, though it will look different.

# The Generalized Interdependency Graph Example



# A Generalized Interdependency Graph

- Through bounding arguments, we can show that we need only consider the case in which the $i^{th}$ vertex in a certain column can only be in equalities with other $i^{th}$ vertices in columns
- Applying König's Theorem, and splitting the graph into several subgraphs, we can come up with the following theorem.

# Main Theorem

**Theorem**

*Given a bipartite graph H, and defining G and M as before,*

$$tr((MM^T)^k) \approx n^{(X-Z)k+Z} \longrightarrow ||M|| \leq n^{\frac{X-Z}{2}} * \log^Z(n)$$

*where Z is the size of the maximum matching of H and X is the number of vertices in H.*

# Future Work

- We can consider a matrix $P = MN$, where $M$ and $N$ are two separate Locally Random Matrices created by two independent bipartite graphs $H$ and $Y$.
- We can consider a matrix $Q = (MN) \circ (P)$, where in this case $M$, $N$, and $P$ are all Locally Random Matrices, but $(MN) \circ (P)$ refers to the entrywise product of $MN$ and $P$.
- We can consider classifying the full spectrum of $M$ instead of just the spectral norm.

# Acknowledgements


- PRIMES-USA
- Aaron Potechin
- Tanya Khovanova
- MIT