GRAPH CLUSTERING AND THE NUCLEAR WASSERSTEIN METRIC.

DANIEL DE ROUX AND MAURICIO VELASCO

Abstract. We study the problem of discovering the cluster structure of a random graph $\mathcal{G}$ from an independent sample of size N. We propose a Wasserstein robust formulation of this optimization problem and prove that it can be reformulated as a tractable convex optimization problem. We give theoretical performance guarantees for this problem when the Wasserstein metric is induced by the nuclear norm and $\mathcal{G}$ is distributed according to the stochastic block model. Finally we present our Julia implementation of the proposed algorithm and use it to analyze the voting patterns of the colombian senate in ♣♣♣ Mauricio: [period?].

1. Introduction

Let $\mathcal{G}$ be a random graph with n vertices. By a deterministic summary of $\mathcal{G}$ we mean a (deterministic) graph H^* which, on average, differs from $\mathcal{G}$ by as few edges as possible. In this article we study the problem of finding deterministic summaries *from an independent sample* of $\mathcal{G}$ of size N. More precisely we will address the following problem:

Problem 1.1. *Given adjacency matrices $B_1, \ldots, B_N$ of an independent sample of $\mathcal{G}$ find a symmetric matrix A^* in $\arg\min_A \mathbb{E}_{B\sim\mathcal{G}}[\|A-B\|_1]$.*

Special cases of this problem arise in cluster detection and in data summarization, both heavily studied in the literature (♣♣♣ Mauricio: [refs?]).

A possible approach to problem 1.1 is to use the samples to construct the empirical measure $\hat{\mu} := \sum_{i=1}^{N} \frac{1}{N}\delta_{B_i}$ as an approximation of the distribution of $\mathcal{G}$ and to find a minimizer A of the resulting empirical risk

$$\mathbb{E}_{B\sim\mu}[\|A-B\|_1] = \frac{1}{N}\sum_{i=1}^{N}\|A-B_i\|_1.$$

When the sample size N is not sufficiently large for $\hat{\mu}$ to be a good approximation for the distribution of $\mathcal{G}$ this approach leads to overfitting. To mitigate this problem we propose a robust version of Problem 1.1. In the robust version one aims to minimize the worst-case risk when the distribution of B is allowed to vary in a ball $\mathcal{N}_\delta(\hat{\mu})$ of radius $\delta > 0$ centered at $\hat{\mu}$ *in a suitable metric.* More precisely,

2000 *Mathematics Subject Classification.* Primary 15A29 Secondary 15B52,52A22.

Key words and phrases. Compressed sensing, truncated moment problems, Kostlan-Shub-Smale polynomials.

Problem 1.2. *Given adjacency matrices $B_1, \dots, B_N$ of an independent sample of $\mathcal{G}$ find a symmetric matrix $\overline{A}$ which minimizes the robust worst-case risk*

$$R_\delta(A) := \left(\sup_{\nu \in \mathcal{N}_\delta(\hat{\mu})} \mathbb{E}_{B\sim G}[\|A-B\|_1]\right).$$

The robust worst-case risk is off course dependent on the chosen metric among probability distributions. In this article we will use the Wasserstein metric $W_{\|\bullet\|}$ associated to a norm $\| \bullet \|$ on the space K of symmetric matrices with entries in $[0,1]$. If ν_1, ν_2 are probability measures on K and $\Pi(\nu_1,\nu_2)$ is the set of random variables (Z_1, Z_2) taking values in $K \times K$ with $Z_i \sim \nu_i$ then this Wasserstein metric is given by

$$W_{\|\bullet\|}(\nu_1,\nu_2) := \inf_{(Z_1,Z_2)\in\Pi(\nu_1,\nu_2)} \mathbb{E}[\|Z_1 - Z_2\|].$$

The seminal work of Esfahani and Kuhn [**?**] shows that robust formulations using the Wasserstein metric lead to tractable convex optimization problems. Our first result is that this is also true for Problem 2 when the Wasserstein metric is induced by a semidefinitely representable norm.

Theorem 1.3. *Let K be the set of $n \times n$ symmetric matrices with entries in $[0,1]$, let $\| \bullet \|$ be a norm on K and let $\delta > 0$. Problem 2 is equivalent to:*

$$\min_{(A,s_1,\dots,s_N,\lambda)\in\mathcal{T}} \left(\lambda\delta + \frac{1}{N}\sum_{i=1}^{N} s_i + \frac{1}{N}\sum \|A - B_i\|_1\right)$$

where $\mathcal{T}$ is the set of $(A, \vec{s}, \lambda) \in K \times \mathbb{R}^n \times \mathbb{R}$ satisfying the inequalities $\lambda \geq 0$ and $\eta_{E,i}(\lambda) \leq s_i$ as $i = 1, \dots, N$ and E ranges over the set of all $\{0,1\}$-symmetric matrices, where

$$\eta_{E,i}(\lambda) := \sup_{Y\in K} \left(\langle E, A - Y\rangle - \lambda\|B_i - Y\|\right)$$

The reformulation in Theorem 1.3 is a finite-dimensional convex optimization problem which unfortunately contains an exponential number of constraints. To address this proble we introduce:

(1) A more tractable simplification which agrees with the original problem whenever the optimal $\overline{A}$ occurs at a matrix with entries in $\{0,1\}$.
(2) A relaxation of (1) which takes the form of a regularized empirical risk minimization problem.

More specifically, we prove the following

Theorem 1.4. *If A is a symmetric $\{0,1\}$-matrix and $\delta > 0$ then the following statements hold:*

(1) The following equality holds:

$$R_\delta(A) = \min_{\mathcal{H}} \left(\lambda\delta + \frac{1}{N}\sum_{i=1}^{N} s_i + \frac{1}{N}\sum_{i=1}^{N} \|A - B_i\|_1\right)$$

where $\mathcal{H}$ is the set of $(\lambda, s_1, \dots, s_N, \Lambda, W) \in \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ ♣♣♣ Mauricio: [Λ y W son simétricas verdad? asi que deberia ser K] *which satisfy the inequalities:*

$$\begin{aligned}
&\|W\|_* \leq \lambda \\
&\Lambda \geq 0 \\
&2A - 11^t - W + \Lambda \geq 0 \\
&\|\Lambda\|_1 \leq s_i - \langle 2A - 11^t - W, B_i\rangle \text{ for } i = 1, \dots, N
\end{aligned}$$

(2) *The following inequality holds:*

$$R_\delta(A) \leq \frac{1}{N}\sum_{i=1}^{N} \|A - B_i\|_1 + \delta\|2A - 11^t\|_*.$$

Moreover, if $\|\cdot\|$ *is a semidefinitely representable function then* (1) *and the upper bound in* (2) *are semidefinite programming problems.*

Solving the optimization problems in Theorem 1.4 leads to a new algorithm for estimating deterministic graph summaries which we call *Wasserstein robust graph summarization.*

We applied this algorithm to a variety of graphs G distributed according to the stochastic block model and observed that the regularized problem with the Wasserstein metric induced by the spectral norm was able to recover the correct cluster structure using only very few samples outperforming all others. More precisely the resulting *Wasserstein robust nuclear norm summarization* problem is given by

$$\min_{A\in K}\left(\frac{1}{N}\sum_{i=1}^{N} \|A - B_i\|_1 + \delta\|2A - 11^t\|_*\right) \tag{1}$$

where the nuclear norm $\|B\|_*$ of a symmetric matrix B equals the sum of the absolute values of its eigenvalues.

Our next result is a performance guarantee for this algorithm which shows that exact recovery occurs for suitable $\delta > 0$ with overwhelming probability on samples constructed via the stochastic block model, explaining the good practical performance of Wasserstein nuclear norm summarization in cluster detection.

In order to describe our performance guarantee we need to establish notation for the parameters of the stochastic block model. Recall that a random graph $\mathcal{G}$ has distribution given by the stochastic block model in n vertices if there is a partition of $[n]$ into disjoint subsets $C_1, \dots, C_k$, real numbers $0 \leq q < \frac{1}{2} < p_1, \dots, p_k \leq 1$ and edges are added independently with probability p_{ij} of joining vertices i, j given by

$$p_{ij} := \begin{cases} p_t, \text{ if } \{i,j\} \subseteq C_t \\ q, \text{ else.} \end{cases}$$

Such a random graph $\mathcal{G}$ has a unique deterministic summary A^* obtained by putting edges only between vertices belonging to the same cluster.

Theorem 1.5. *Suppose $B_1, \dots, B_N$ are independent and are distributed according to $\mathcal{G}$. If $\alpha = \min(|p_t - \frac{1}{2}|, |q - \frac{1}{2}|)$ and δ^* is the maximum of $a(\delta) := \frac{\delta\left(\alpha - \frac{\delta}{n}\right)^2}{\left(1+\frac{2\delta}{n}\right)}$ in $[0, \alpha n]$ then the probability that A^* is not the unique minimizer of (1) is bounded above by*

$$\exp\left(-\frac{2N(n-1)^3(\alpha n - \delta^*)^2}{n}\right) + e^{-N(\delta^* a(\delta^*))} \prod_{i \neq j} \left(1 + (e^{-4a}\widetilde{p_{ij}} + \widetilde{q_{ij}})^{\frac{N}{2}}\right).$$

which decreases exponentially with the sample size N.

The key point of the proof of the Theorem, discussed at length in Section **??**, is that the nuclear norm is a good regularizer for graph summarization problems because the subdifferential of the regularization term at A^* is sufficiently rich so as to contain enough transportation matrices.

Finally we focus on the practical performance of the proposed algorithm. Solving the optimization problems appearing in Theorem 1.4 require solving large semidefinite programs which are beyond the capacity of standard off-the-shelf software even for relatively small graphs (of say 40 vertices with $N = 4$). One possible reason is that off-the-shelf solvers often use interior point methods, which are highly accurate but often do not scale well. A better alternative, especially well suited for solving **??** is to use first order numerical optimization methods such as ADMM. In Section **??** we adapt the ADMM algorithm for our regularized problem and present our open source Julia implementation. This implementation can run the Wasserstein nuclear norm summarization algorithm in graphs of up to 10000 vertices on a common laptop. In Section **??** we use it to find a deterministic summary of voting patterns in the colombian senate during the period ♣♣♣ Mauricio: [Period].

1.1. **Relation to previous work.** ♣♣♣ Mauricio: [Daniel porfa escribir acá los párrafos que me mencionó y enriquecer enormemente las referencias (estan en el archivo lit.bib en formato bibtex, para poner nuevas referencias solo pongalas ahi, es facil cargarlas de MathScinet en formato bibtex).]

1.2. **Acknowledgments.** We wish to thank Fabrice Gamboa, Mauricio Junca and Thierry Klein for useful conversations during the completion of this project. ♣♣♣ Mauricio: [CienciaPolitica]. D. De Roux was partially supported by a grant from Facultad de Ciencias, Universidad de los Andes. M. Velasco was partially supported by research funds from Universidad de los Andes.

2. Robust learning of deterministic summaries

Let $[n] := \{1, 2, \dots, n\}$. By a graph G on $[n]$ we mean a finite loopless undirected graph with vertex set $[n]$. The adjacency matrix of such a graph is a symmetric $n \times n$ matrix with entries in $\{0, 1\}$ and ones in positions (i, j) whenever vertices i and j are adjacent. Let $\mathcal{A}_n$ be the set of adjacency matrices of graphs on $[n]$ (i.e. $\{0, 1\}$-matrices of size $n \times n$ which are symmetric and have zeroes in the diagonal). Throughout the article we will use graphs and their adjacency matrices interchangeably. By a random graph $\mathcal{G}$ on $[n]$ we mean a random variable B taking values in $\mathcal{A}_n$.

Definition 2.1. *A deterministic summary of a random graph $\mathcal{G}$ on $[n]$ is a graph A^* with $A^* \in \arg\min_{A\in\mathcal{A}_n} \mathbb{E}[\|A-\mathcal{G}\|_1]$.*

Here the norm $\|\bullet\|_1$ means the sum of absolute values of all entries of the matrix. Therefore the deterministic summary A^* is a graph which, on average, differs from $\mathcal{G}$ by the smallest possible number of edges. If the distribution of $\mathcal{G}$ is known it is easy to find a deterministic summary (which is often unique). More precisely,

Lemma 2.2. *If $\mathbb{P}\{(i,j)\in\mathcal{G}\} \neq \frac{1}{2}$ for all $i,j\in[n]$ then the unique deterministic summary of $\mathcal{G}$ is given by A^* with*

$$A^*_{ij} = \begin{cases} 1, \text{ if } \mathbb{P}\{(i,j)\in\mathcal{G}\} > \mathbb{P}\{(i,j)\notin\mathcal{G}\} \\ 0, \text{ if } \mathbb{P}\{(i,j)\in\mathcal{G}\} < \mathbb{P}\{(i,j)\notin\mathcal{G}\} \end{cases}$$

Proof. If $A\in\mathcal{A}_n$ then the term coming from entry (i,j) in $\mathbb{E}[\|A-\mathcal{G}\|_1]$ is given by $|A_{ij}-1|p+|A_{ij}|q$ where p (resp q) is the probability that (i,j) is (resp. is not) an edge of $\mathcal{G}$. This quantity is greater than $\min(p,q)$ and equality is achieved, for all i,j with $A=A^*$. $\square$

Motivated by the previous Lemma we define a cluster structure on a random graph.

Definition 2.3. *A random graph $\mathcal{G}$ has a cluster structure if it has a unique deterministic summary A^* and the corresponding graph is a disjoint union of cliques. We call these cliques the clusters of $\mathcal{G}$.*

The main problem that we address in this article is that of *learning* deterministic summaries (an in particular the problem of learning cluster structures on random graphs who have them). By this we mean that our only knowledge about the distribution of the random graph $\mathcal{G}$ is encoded in an independent sample $B_1,\dots,B_N$ of adjacency matrices with the same distribution as $\mathcal{G}$, leading to Problem 1.1 in the Introduction.

Given the sample, define the empirical measure $\hat{\mu} := \frac{1}{N}\sum_{j=1}^N \delta_{B_j}$ as a sum of Dirac delta measures at the sample points. As the number of sample points increases the measure $\hat{\mu}$ converges to the distribution of $\mathcal{G}$ and it is therefore reasonable to try to minimize the objective function in Problem 1.1 with respect to the measure $\hat{\mu}$ instead of $\mathcal{G}$, that is by finding a minimizer of the empirical risk

$$\overline{A} \in \arg\min_{A\in K} \mathbb{E}_{Z\sim\hat{\mu}}[\|A-Z\|]$$

Arguing as in Lemma 2.2 it is immediate that a (generally unique) minimizer $\overline{A}$ is given by counting edge frequencies, that is

$$\overline{A}_{ij} := \begin{cases} 1, \text{ if } |\{t : B^{(t)}_{ij}=1\}| > |\{t : B^{(t)}_{ij}=0\}| \\ 0, \text{ if } |\{t : B^{(t)}_{ij}=1\}| < |\{t : B^{(t)}_{ij}=0\}| \end{cases}$$

The empirical risk minimization approach has lots of advantages, it is easy to implement, scales very well and is guaranteed to be consistent (in the sense that $\overline{A}_{ij}\to\overline{A}$ as the number of samples $N\to\infty$). However it also suffers from some potential drawbacks:

(1) If the sample size N is small then the empirical measure $\hat{\mu}$ could be very far from the distribution of $\mathcal{G}$.
(2) The estimation of $\overline{A}$ is done independently edge by edge and in particular it does not use any global information, for instance the existence of a cluster structure as part of the estimation process.

To mitigate these problems we will use a robust formulation. To this end let $\|\bullet\|$ be a norm in the space of symmetric $n \times n$ matrices and let $W_{\|\bullet\|}$ be the Wasserstein distance induced by $\|\bullet\|$ on probability distributions taking values on K. For a real number $\delta \geq 0$ let $\mathcal{N}_\delta(\hat{\mu})$ be the (closed) ball of radius δ centered at $\hat{\mu}$. We would like to solve the following

Problem. *Given adjacency matrices $B_1, \ldots, B_N$ of an independent sample of $\mathcal{G}$ find a symmetric matrix $\overline{A}$ which minimizes the robust worst-case risk*

$$R_\delta(A) := \left(\sup_{\nu \in \mathcal{N}_\delta(\hat{\mu})} \mathbb{E}_{B \sim G}[\|A - B\|_1] \right).$$

Our first result reformulates the robust optimization above as a finite convex optimization problem.

Proof of Theorem 1.3. □

3. Clustering in the stochastic block model and the Wasserstein nuclear norm.

4. An algorithm for Wasserstein nuclear norm summarization

5. A numerical example: the colombian senate voting record.

6. Preliminaries

6.1. **Preliminaries on graphs and norms.** By a graph G we mean a finite loopless undirected graph. We say that G is weighted if it is endowed with a function $w : E(G) \to \mathbb{R}$ which assigns to every edge a real number in $[0, 1]$. If G has n vertices then it is completely specified by its adjacency matrix $A \in \{0, 1\}^{n \times n}$ defined by $A_{ij} = 1$ if and only if vertices i, j are connected. If G is weighted then we use the term adjacency matrix of G to denote the matrix with entries $A_{i,j} = w(i, j)$.

If A is a matrix then we use $\|\bullet\|$, $\|\bullet\|_1$ to denote its operator norm and ℓ^1-norm respectively.

7. A description of the problem

By a random graph B we mean a random variable B taking values on the set of adjacency matrices of graphs (i.e. symmeric matrices in $\{0, 1\}^{n \times n}$ with zero diagonal). By a random weighted graph we mean a random variable taking values in the adjacency matrices of weighted graphs (i.e. symmetric matrices with 0 diagonal all of whose off-diagonal entries lie in $[0, 1]$).

Definition 7.1. *Let B be a random weighted graph and let A be a (weighted) adjacency matrix. Define the risk of choosing $A \in \{0,1\}^n$ as a deterministic summary of B as*

$$R(A) := \mathbb{E}[\|A - B\|_1].$$

We say that A^* is an optimal summary of a random graph B in the set S if it is a minimizer of the optimization problem $\min_{A \in S} R(A)$.

8. Recovering cluster structures in the stochastic block model

Suppose B is a random (undirected loopless) graph on n vertices generated by the stochastic block model. This means that we fix a set partition $C_1, \dots, C_l$ of $[n]$ into sets we call clusters and real numbers $0 \le p_i, \bar{p} \le 1$ for $i = 1, \dots, l$. The edges of B are independent random variables and an edge joins vertices i, j with probability p_t if $\{i,j\} \subseteq C_t$ for some cluster C_t and with probability $\bar{p}$ if $\{i,j\}$ is not contained in any C_t. We let $O \subseteq [n] \times [n]$ be the set of pairs of vertices which are not simultaneously contained in any cluster.

For an integer N let $B_1, \dots, B_N$ be an independent sample of N graphs with the distribution of B. Let A^* be the $n \times n$ matrix with entries in $\{0,1\}$ which captures the underlying cluster structure, namely $A^*_{ij} = 1$ iff there is a cluster C_t which contains both ij. ♣♣♣ Daniel: [or if i=j. This is important as proofs are clearer if we consider that the entries in the diagonal are not in I or in O.] In this section we study the probability, as a function of δ that the optimization problem $\min_A \Delta(A)$

$$\Delta(A) = \delta \|2A - 11^t\|_* + \frac{1}{N} \sum_{k=1}^{N} \|A - B_k\|_1$$

has the correct cluster structure A^* as a minimizer. For vertices $i, j \in [n]$ define $n_1(ij)$ (resp. $n_0(ij)$) the random variables which count the number of times that a given pair is (resp is not) an edge of some B_j, $j = 1, \dots N$. Note that the $n_q(ij)$ for $q = 0, 1$ are binomial random variables.

Lemma 8.1. *The following statements hold:*

1. *The subdifferential of $\frac{1}{N}\sum_{k=1}^{N} \|A - B_k\|_1$ at A^* is the set of symmetric matrices C satisfying the inequalities*
$$\frac{n_0(ij) - n_1(ij)}{N} \le C_{ij} \le 1, \text{ if } \{i,j\} \subseteq C_t \text{ for some } t,$$
$$-1 \le C_{ij} \le \frac{n_0(ij) - n_1(ij)}{N} \text{ if } \{i,j\} \text{ does not belong to any cluster and}$$
$$-N \le C_{ii} \le N \text{ for all } i.$$
2. *The subdifferential of $\delta\|2A - 11^t\|_*$ at A^* is given by the set of symmetric matrices of the form $2\delta C$ where C has spectral norm $\|C\| \le 1$ and satisfies $\langle C, 2A - 11^t \rangle = n$.*

Proof. (1) Since the subdifferential is additive it suffices to understand the subdifferential of the absolute value. If $i, j \in C_t$ then $A^*_{ij} = 1$ and the entry ij of the subdifferential of the sum at A^* is $[-1, 1]$ for each B_i containing the edge and it is 1 for each B_i for which (ij) is not an edge. ♣♣♣ Daniel: [vale la pena mencionar el caso i=j o es obvio la propiedad en la diagonal?] If i, j is not contained in any cluster then $A^*_{ij} = 0$ and the entry ij of the subdifferential of the sum at A^* is -1 for each B_i which contains the edge ij and $[-1, 1]$ for each B_i which does not, proving the claim. (2) It is easy to prove that the subdifferential of any norm $\| \bullet \|$ at a point X is given by those C for which the dual norm $\|C\|_* \leq 1$ and $\langle C, X\rangle = \|X\|$. Claim (2) follows because $\|2A - 11^t\| = Tr(2A - 11^t) = n$ where the first equality holds since $2A - 11^t$ is positive semidefinite.♣♣♣ Mauricio: [Esto es obvio con solo dos clusters (la matriz es uu^t donde u es el vector con 1's en un cluster y -1's en el complemento pero hay que demostrarlo para tres o mas)]. □

In the following section we discuss how tradeoffs between components explain the improved recovery probability induced by the spectral norm.

9. Estimating recovery probabilities

Let Γ be the symmetric matrix with zero diagonal and off-diagonal entries given by $\Gamma_{ij} = \frac{n_0(ij) - n_1(ij)}{N}$. By Lemma 8.1 a symmetric matrix C lies in the subdifferential if and only if it satisfies the inequalities

$$\Gamma_{ij} \leq C_{ij} \leq 1, \text{ if } \{i, j\} \subseteq C_t \text{ for some } t,$$

$$-1 \leq C_{ij} \leq \Gamma_{ij} \text{ if } \{i, j\} \text{ does not belong to any cluster and}$$

$$-N \leq C_{ii} \leq N \text{ for all } i.$$

To simplify these inequalities we define a linear operator $\widetilde{\bullet}$ on symmetric matrices by the formula

$$\widetilde{A} = \begin{cases} A_{ij} \text{ if } i = j \text{ or } ij \in I \text{ and} \\ -A_{ij} \text{ if } ij \in O. \end{cases}$$

in this language C_{ij} belongs to the subdifferential if and only if $\widetilde{\Gamma}_{ij} \leq \widetilde{C}_{ij}$ for $i \neq j$. The following key result gives sufficient conditions for the true cluster structure A^* to be a minimizer of the proposed optimization problem. In order to describe it we introduce the following notation.

Definition 9.1. *Let δ be a positive real number. For a symmetric matrix Γ define the quantities*

$$b(\Gamma, \delta) := \sum_{i \neq j} \max\left(\widetilde{\Gamma_{ij}} + \frac{2\delta}{n}, 0\right) \text{ and } a(\Gamma, \delta) := \sum_{i \neq j} \max\left(-\widetilde{\Gamma_{ij}} - \frac{2\delta}{n}, 0\right)$$

The quantity $b(\Gamma, \delta)$ (resp. $a(\Gamma, \delta)$) measures the total amount by which the matrix $\widetilde{-\frac{2\delta}{n}11^t}$ fails (resp. succeeds) to be in the subdifferential of Lemma 8.1 in the sense that it sums over all ij the amount by which the inequalities $\widetilde{\Gamma_{ij}} \leq \frac{2\delta}{n}11^t$ fail (resp. succeed). The key point of the following Theorem is that if the inequality fails by less than it succeeds then the subdifferential of the spectral norm is sufficiently rich so as to allow us to redistribute these quantities. In this sense the following Theorem explains the success of the spectral norm in cluster recovery algorithms.

Recall that

$$\Delta(A) = \delta \|2A - 11^t\|_* + \frac{1}{N}\sum_{k=1}^{N} \|A - B_k\|_1$$

And define functions f, g:

$$f := \frac{1}{N}\sum_{k=1}^{N} \|A - B_k\|_1, \quad g := \delta\|2A - 11^t\|_*$$

In this vocabulary, we have that

$$\partial(\Delta) = \partial(f)(A^*) + \partial(g)(A^*) \tag{2}$$

Theorem 9.2. *[Alternative proof of theorem 6.2] Assume there are only two clusters. Let $\delta > 0$ with $\left(\frac{\delta}{n} + b(\Gamma, \delta)\right) < \frac{N}{2}$. If $b(\Gamma, \delta) < \min(\delta, a(\Gamma, \delta))$ then A^* is a minimizer of the optimization problem $\min_A \Delta(A)$.*

Proof. We will show that there exists a matrix C such that $-C \in \partial(g)(A^*)$ for which $\widetilde{\Gamma}_{ij} \leq \widetilde{C_{ij}}$ for $i \neq j$. This implies that $C_{ij} \in \partial(f)(A^*)$ and therefore $0 = C - C$ belongs to the subdifferential $\partial(\Delta)(A^*)$ and thus A^* is a minimizer of $\Delta(A)$.

Recall that $-C \in \partial(g)(A^*)$ if and only if

$$\langle H^*, C\rangle = -2\delta n \text{ and } \|C\| \leq 2\delta$$

where $H^* = 2A - 11^t$ and $\| \bullet \|$ is the spectral norm.

Notice that both these conditions are satisfied by setting $C^0 = -\frac{2\delta}{n}H^*$ as

$$\frac{-2\delta}{n}\langle H^*, H^*\rangle = \frac{-2\delta}{n}n^2 = -2\delta n \text{ and } \|C^0\| = \frac{2\delta}{n}\|H^*\| = 2\delta.$$

However this choice of C^0 will not, in general, satisfy the inequalities $\widetilde{\Gamma}_{ij} \leq \widetilde{-\frac{2\delta}{n}H^*}_{ij}$ for $i \neq j$. Therefore, we will correct our candidate matrix C^0 so that it satisfies these inequalities and still belongs to $\partial(g)(A^*)$.

To do this, we will construct a matrix K and add it to C^0. Crucially, K will satisfy that $KH^* = H^*K = 0$ so that we can control the spectral norm of $C^0 + K$.

Let $i < j$ and define the symmetric matrix e^{ij} as follows:

$$e^{ij} = \begin{cases} 1 \text{ if } (i,j) \in I. \\ -1 \text{ if } (i,j) \in O. \\ -1 \textit{ in the entry } ii \text{ and in } jj. \\ 0 \text{ otherwise.} \end{cases} \tag{3}$$

Observe that for any (i, j) the matrix e^{ij} satisfies the following properties:

(1) $H^*e^{ij} = 0 = e^{ij}H^*$.
(2) The inequality $\| - e^{ij} + e^{st}\| \leq 2$ holds for all ij and st. This is immediate noting that the spectral norm is bounded by the product of the induced $1-$*norm* and the induced $\infty-$norm. (The inequality is strict only if $|\{i,j\} \cap \{s,t\}| \geq 1$ and in this case it can take values of $\sqrt{3}$ and 0).

The idea is to use that $b(\Gamma,\delta) < a(\Gamma,\delta)$ so there exists a way to redistribute the quantity $b(\Gamma,\delta)$ by subtracting it from the ij for which $\frac{2\delta}{n} < \widetilde{\Gamma}_{ij}$ and adding it into those st for which $\widetilde{\Gamma}_{st} \leq \frac{2\delta}{n}$.

Let U be the set of entries $\{i,j\}$ where $\widetilde{\Gamma}_{ij} + \frac{2\delta}{n} \leq 0$ and V be the set of entries $\{i,j\}$ where $-(\widetilde{\Gamma}_{ij} + \frac{2\delta}{n}) > 0$. Observe that

$$b(\Gamma,\delta) = \sum_{ij\in U} (\widetilde{\Gamma}_{ij} + \frac{2\delta}{n}) \text{ and } a(\Gamma,\delta) = \sum_{ij\in V} -(\widetilde{\Gamma}_{ij} + \frac{2\delta}{n}).$$

By hypothesis $b(\Gamma,\delta) < a(\Gamma,\delta)$. Let $l_1,..,l_k$ be an enumeration of V were k is it's cardinality. For each entry $ij \in U$, there exists nonnegative coefficients $\gamma^{ij}_{l_1},..,\gamma^{ij}_{l_k}$ such that:

$$\widetilde{\Gamma}_{ij} + \frac{2\delta}{n} - \gamma^{ij}_{l_1} - ... - \gamma^{ij}_{l_k} < 0.$$

and that such that for each γ_{l_p} with $p \in \{1,..,k\}$

$$-(\widetilde{\Gamma}_{l_p} + \frac{2\delta}{n}) - \sum_{ij\in U} \gamma^{ij}_{l_p} > 0. \tag{4}$$

For $ij \in U$ define the matrix

$$W^{ij} := \gamma^{ij}_{l_1}(\widetilde{e^{ij} - e^{l_1}}) + ... + \gamma^{ij}_{l_k}(\widetilde{e^{ij} - e^{l_k}})$$

Given that $ij \in U$, $l_1,..,l_k \in V$ and the sets U and V are disjoint, the support of e^{ij} is disjoint from the support of any of the matrices $e^{l_1},..,e^{l_k}$ (except probably at the diagonal). In particular, if $ij \in I$ the entry ij of the matrix W^{ij}, namely W^{ij}_{ij} is equal to:

$$\gamma^{ij}_{l_1}(1-0) + ... + \gamma^{ij}_{l_k}(1-0) = \gamma^{ij}_{l_1} + ... + \gamma^{ij}_{l_k}.$$

and if $ij \in O$,

$$\gamma^{ij}_{l_1}(-1-0) + ... + \gamma^{ij}_{l_k}(-1-0) = -\gamma^{ij}_{l_1} - ... - \gamma^{ij}_{l_k}.$$

Define the matrix C^1 as:

$$C^1 := C^0 + \sum_{ij\in U} W^{ij}$$

We will now verify that $C^1 \in \partial(f)(A^*)$. For $ij \in V$, we have by definition of V and by equation 4 that $\widetilde{\Gamma}_{ij} < \widetilde{C}^1_{ij}$.

Let $ij \in I \cap U$. We have that $\Gamma_{ij} + \frac{2\delta}{n} > 0$. Now, the entry of C^1 in ij is given by:

$$C^1_{ij} = -\frac{2\delta}{n} H_{ij} + W^{ij}_{ij} = -\frac{2\delta}{n} + \gamma^{ij}_{l_1} + ... + \gamma^{ij}_{l_k}$$

By the construction of the γ's, we have that

$$\Gamma_{ij} + \frac{2\delta}{n} - \gamma^{ij}_{l_1} - ... - \gamma^{ij}_{l_k} < 0$$

it follows that

$$\Gamma_{ij} - C^1_{ij} < 0$$

so that

$$C^1_{ij} > \Gamma_{ij}.$$

For $ij \in O \cap U$, we have that $\frac{2\delta}{n} - \Gamma_{ij} > 0$. The entry of C^1 in ij is given by:

$$C^1_{ij} = -\frac{2\delta}{n} H_{ij} + W^{ij}_{ij} = \frac{2\delta}{n} - \gamma^{ij}_{l_1} - ... - \gamma^{ij}_{l_k}$$

Since

$$\frac{2\delta}{n} - \Gamma_{ij} - \gamma^{ij}_{l_1} - ... - \gamma^{ij}_{l_k} < 0$$

it follows that

$$-C_{ij} > -\Gamma_{ij}.$$

and that $\widetilde{\Gamma}_{ij} < \widetilde{C}^1_{ij}$ for $i \neq j$.

It remains to show that the entries of the diagonal of C^1 are bounded by n. The diagonal entries $\{ss\}$ of C^1 are given by

$$-\frac{2\delta}{n} + \sum_{ij \in U} W^{ij}_{ss}.$$

Notice that by the definition of W^{ij}, each of the matrices $(\widetilde{e^{ij} - e^{l_p}})$ has, in the worst case, a -2 in the entry ss so the entry ss of C^1 is bounded below by

$$-\frac{2\delta}{n} + -2 \left(\sum_{ij \in U} \sum_{l_p \in V} \gamma^{ij}_{l_p} \right)$$

the quantity in the parentheses is all the weight that we have to distribute, i.e $b(\Gamma, \delta)$. Therefore,

$$C^1 \geq -\frac{2\delta}{n} - 2(b(\Gamma, \delta)) = -2\left(\frac{\delta}{n} + b(\Gamma, \delta)\right).$$

By hypothesis, $\frac{N}{2} > \left(\frac{\delta}{n} + b(\Gamma, \delta)\right)$ so we obtain that

$$C^1_{ss} > -N.$$

It is obvious that $N > C^1_{ss}$. We conclude that $C^1 \in \partial(f)(A^*)$.

For each ij, $H^* e^{ij} = 0 = e^{ij} H^*$ so the equality $\langle H^*, C^1 \rangle = \langle H^*, -\frac{2\delta}{n} H^* \rangle = -2\delta n$ holds and moreover

$$\|C^1\| = \max \left(\left\| -\frac{2\delta}{n} H^* \right\|, \left\| \sum_{ij \in U} W^{ij} \right\| \right).$$

The operator norm of the first term in the maximum equals 2δ and that of the second term is bounded by $2b(\Gamma, \delta)$ by the triangle inequality and the definition of $b(\Gamma, \delta)$. We conclude that $\|C\|$ is bounded by 2δ because $b(\Gamma, \delta) \leq \delta$. As a result $-C \in \partial(g)(A^*)$ proving the Theorem. Note that C^1 satisfies all the inequalities that define the membership to $\partial(f)(A^*)$ in 8.1 strictly, so C^1 belongs is interior point of $\partial(f)(A^*)$.

$\square$

We will now prove the general case when there are more than two clusters. The proof will be similar the proof of the previous theorem. We will start with the candidate matrix $\frac{-2\delta}{n}H^*$ and correct it by adding transport matrices that assure that the corrected matrix belong to the subdifferential of $\Delta(A)$ at A^*. The difficulty in applying the tools of the previous theorem to solve the general case is that the matrices K that transport weight from one cluster to another do not, in general, satisfy the relation $KH^* = H^*K = 0$ when there are more than two clusters. This problem can be solved by splitting the matrix H^* into matrices than only take into account 2 clusters, and using the previous theorem. We begin recalling the following simple result:

Claim 9.3. *Let $a_i, b_i \geq 0,\ i = 1, \dots, p$ be two non-negative, finite sequences of real numbers such that*

$$\sum_{i=1}^{p} a_i \geq \sum_{i=1}^{p} b_i.$$

Then there exist a finite sequence of reals c_i such that

- $\sum_{i=1}^{p} c_i = 0.$
- $a_i \geq b_i + c_i\ \forall i.$

Now we proceed to do the proof. For clusters $C_s \neq C_t$ define the matrix $H^{C_sC_t}$ whose entries are given by:

$$H^{C_iC_j}_{uv} = \begin{cases} 1 \text{ if } u, v \in C_s \text{ or } u, v \in C_t. \\ -1 \text{ if } u \in C_s, v \in C_t \text{ or } u \in C_s, v \in C_t. \\ 0 \text{ in any other case.} \end{cases}$$

Notice that this matrix has 4 blocs. Two with only 1 and two with only -1. Moreover, its spectral norm is equal to $|C_s + |C_t|$.

Lemma 9.4. *Assume there are l clusters. Suppose that $b(\Gamma, \delta) < min(\delta, a(\Gamma, \delta))$. Then, A^* is a minimizer of the optimization problem $\min_A \Delta(A)$.*

Proof. First of all, observe that

$$\frac{-2\delta}{n}H^* = \frac{-2\delta}{n}\frac{1}{l-1}\sum_{1\leq s<t\leq l} H^{C_sC_t}.$$

For each, $C_s \neq C_t$ construct a transport matrix Δ_{st} as in the previous theorem, as to assure that $\Delta_{st}H^{C_sC_t} = H^{C_sC_t}\Delta_{st} = 0$. This can be done since $H^{C_sC_t}$ takes into account only two clusters. Recall that the total amount of weight to be corrected is $b(\Gamma, \delta)$. Let $w_{s,t}$ the weight to be distributed from cluster s to cluster t. In the notation of the previous theorem, w_{st} is just the sum of the $\gamma^{ij}_{l_p}$ where $l_p \in C_s$ and $ij \in C_t$.

Let

$$C := -\frac{2\delta}{n}H^* + \sum_{i<j}\Delta_{ij} = \frac{-2\delta}{n}\frac{1}{l-1}\sum_{1\leq i<j\leq l} H^{C_iC_j} + \sum_{i<j}\Delta_{ij}$$

Finally, assume that for each $i < j$, $\frac{1}{l-1}||H^{C_iC_j}|| \geq ||\Delta_{ij}||$.

Then,

$$\begin{aligned}\|C\| &= \left\| \frac{-2\delta}{n}\frac{1}{l-1}\sum_{1\le i<j\le l} H^{C_iC_j} + \sum_{i<j}\Delta_{ij}\right\| \\ &\le \frac{2\delta}{n(l-1)}\sum_{1\le i<j\le l}\left\| H^{C_iC_j} + \frac{n(l-1)}{2\delta}\Delta_{ij}\right\| \\ &= \frac{2\delta}{n(l-1)}\sum_{1\le i<j\le l}\max(\left\|H^{C_iC_j}\right\|, \left\|\frac{n(l-1)}{2\delta}\Delta_{ij}\right\|)\end{aligned}$$

Now notice that

$$\delta \ge b(\Gamma,\delta) \text{ therefore } (l-1)n \ge \frac{(l-1)n}{\delta}b(\Gamma,\delta)$$

$$\begin{aligned}\text{which implies that } \sum_{i<j}\left\|H^{C_iC_j}\right\| &\ge \frac{(l-1)n}{\delta}\sum_{i<j} w_{i,j} \\ &\ge \frac{(l-1)n}{2\delta}\sum_{i<j}\|\Delta_{i,j}\|.\end{aligned}$$

By the claim, we can assume without loss of generality that for each $i<j$,

$$\left\|H^{C_i,C_j}\right\| \ge \frac{(l-1)n}{\delta}\|\Delta_{i,j}\|.$$

This implies that the last sum reduces to

$$\frac{2\delta}{n(l-1)}\sum_{1\le i<j\le l}\left\|H^{C_iC_j}\right\| = \sum_{1\le i<j\le l}\frac{2\delta(|C_i|+|C_j|)}{n(l-1)} = \frac{2\delta(l-1)}{n(l-1)}\sum_{1\le i<j\le l}|C_i|+|C_j| = 2\delta.$$

And so $\|C\| \le 2\delta$.

□

♣♣♣ Daniel: [toca revisar que $< h^*, C >$ es igual a $-2\delta n$ o eso es obvio?]

Uniqueness of the minimizer. In this brief section we discuss an important corollary: A^* is the unique minimizer of the optimization problem $min_A\Delta(A)$. We begin proving a well known lemma.

Lemma 9.5. *Let f be a convex function defined over a region D. Let $\hat{x}$ be a point in it's domain such that the subdifferential of f at $\hat{x}$ is full dimensional and 0 belongs to it's interior. Then, $\hat{x}$ is the unique minimizer of f.*

Proof. Let $B_\epsilon(0)$ be a ball of radius ϵ centered in 0 and contained in $\partial(f)(\hat{x})$. Let $x\in D$ with $x\neq\hat{x}$. Let $Q\in\partial(f)(\hat{x})$. By the property of the elements of the subdifferential at a point,

$$f(x) \ge f(\hat{x}) + \langle Q, x-\hat{x}\rangle.$$

This property holds for every $Q\in\partial(f)(\hat{x})$, so taking supremum we obtain that

$$f(x) \ge \sup_{Q\in\partial(f)(\hat{x})} f(\hat{x}) + \langle Q, x-\hat{x}\rangle..$$

Since $B_\epsilon(0)\subseteq\partial(f)(\hat{x})$ we obtain that

$$f(x) \geq f(\hat{x}) + \sup_{Q \in B_\epsilon(0)} \langle Q, x - \hat{x} \rangle = f(\hat{x}) + \epsilon \|x - \hat{x}\|.$$

As $x \neq \hat{x}$, $\epsilon\|x - \hat{x}\| > 0$. Therefore, $f(x) > f(\hat{x})$ for all $x \in D$ different of $\hat{x}$.

□

Remark 9.6. Under the conditions of theorem 9.2, The matrix C^1 we constructed satisfies the inequalities given by 8.1 strictly, so c^1 is an interior point of $\partial(f)(A^*)$. It follows that the matrix C constructed in 9.4 also satisfies these inequalities strictly and thefore it is also an interior point of $\partial(f)(A^*)$.

Corollary 9.7. *Under the conditions of theorem 9.2, A^* is the unique minimizer of the optimization problem $min_A \Delta(A)$.*

Proof. By 9.6, the matrix C constructed in 9.4 is an interior point of $\partial(f)(A^*)$. Therefore, as the subdifferential of Δ at A^* is equal to the Minkowski sum of the subdifferentials of g and f at A^*,$0 = C - C$ is an interior point of $\partial(\Delta)(A^*)$. It follows by lemma 9.5 that A^* is the unique minimizer of the optimization problem $min_A \Delta(A)$.

□

Using Theorem 9.2 we now estimate the probabilities of perfect recovery of the correct cluster structure.

9.1. **A bound for recovery probabilities.** In this section we will bound the probability that the correct A^* is not an optimal solution of our proposed optimization problem. A key tool will be the following version of Hoeffding's inequality: If $X_1, \ldots, X_T$ are independent random variables with values in $[c_i, d_i]$ and $\Lambda_T := \sum_{i=1}^T X_i$ then the following inequality holds for all $t \geq 0$

$$\mathbb{P}\{\Lambda_T - \mathbb{E}[\Lambda_T] \geq t\} \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^T (d_i - c_i)^2}\right).$$

Theorem 9.8. *Suppose $B_1, \ldots, B_N$ are independent and have the same distribution as $\mathcal{G}$. If $\alpha = \min(|p_t - \frac{1}{2}|, |q - \frac{1}{2}|)$ and δ^* is the maximum of $a(\delta) := \frac{\delta\left(\alpha - \frac{\delta}{n}\right)^2}{\left(1 + \frac{2\delta}{n}\right)}$ in $[0, \alpha n]$ then the probability that A^* is not a minimizer of (??) is bounded above by*

$$\exp\left(-\frac{2N(n-1)^3(\alpha n - \delta^*)^2}{n}\right) + e^{-N(\delta^* a(\delta^*))} \prod_{i \neq j} \left(1 + (e^{-4a} p_{ij} + q_{ij})^{\frac{N}{2}}\right).$$

Moreover this quantity decreases exponentially with the sample size N.

Proof. By Lemma 9.4 and the union bound the probability that A^* is not an optimal solution of problem (2) is bounded above by

$$\mathbb{P}\{b(\Gamma, \delta) \geq a(\Gamma, \delta)\} + \mathbb{P}\{b(\Gamma, \delta) \geq \delta\}. \tag{5}$$

and we will find upper bounds for the individual terms in (5). For $t = 1, \dots, N$ and $i, j \in [n]$ define

$$Z_{ij}^{(t)} := \begin{cases} -1, \text{ if } (B_t)_{ij} = 1 \\ +1, \text{ if } (B_t)_{ij} = 0 \end{cases}$$

and note that for every $i \neq j$ the equality $\sum_{t=1}^{N} \frac{Z_{ij}^{(t)}}{N} = \Gamma_{ij}$ holds. As a result

$$b(\Gamma, \delta) - a(\Gamma, \delta) = \sum_{i \neq j} \left(\widetilde{\Gamma_{ij}} + \frac{2\delta}{n} \right) = \frac{2\delta n(n-1)}{n} + \sum_{t=1}^{N} \sum_{i \neq j} \frac{\widetilde{Z_{ij}^{(t)}}}{N}$$

and therefore if $M := \mathbb{E}\left[\sum_{t=1}^{N} \sum_{i \neq j} \frac{\widetilde{Z_{ij}^{(t)}}}{N}\right]$ then the number M is negative and is given by the formula

$$M = (2q-1)2|O| + \sum_{i=1}^{l} 2\binom{c_i}{2}(1-2p_i) \leq -2\alpha n(n-1)$$

We can therefore bound the probability in the first term with

$$\mathbb{P}\left\{ \frac{2\delta n(n-1)}{n} + \sum_{t=1}^{N} \sum_{i \neq j} \frac{\widetilde{Z_{ij}^{(t)}}}{N} \geq 0 \right\} = \mathbb{P}\left\{ \sum_{t=1}^{N} \sum_{i \neq j} \frac{\widetilde{Z_{ij}^{(t)}}}{N} - M \geq -2\delta(n-1) - M \right\} \leq$$

$$\leq \exp\left(-2\frac{(-M - 2\delta(n-1))^2}{Nn(n-1)(\frac{2}{N})^2} \right) = \exp\left(-\frac{N}{2}\left(1 - \frac{1}{n}\right)\left(\frac{-M}{n-1} - 2\delta\right)^2 \right)$$

where the inequality follows from Hoeffding's inequality applied to the $Nn(n-1)$ independent random variables $\frac{\widetilde{Z_{ij}^{(t)}}}{N}$ which have values in $\left[-\frac{1}{N}, \frac{1}{N}\right]$. The inequality applies whenever $-\frac{M}{2(n-1)} > \delta > 0$. In particular whenever $0 < \delta < \alpha n$ we have

$$\mathbb{P}\{b(\Gamma, \delta) - a(\Gamma, \delta)\} \leq \exp\left(-\frac{2N(n-1)^3(\alpha n - \delta)^2}{n} \right).$$

Bounding the second term is more involved. Recall that

$$b(\Gamma, \delta) = \sum_{i \neq j} \max\left(\widetilde{\Gamma_{ij}} + \frac{2\delta}{n}, 0 \right).$$

Let $Y_{ij} := \widetilde{\Gamma_{ij}} + \frac{2\delta}{n}$ and let $X_{ij} := \max(Y_{ij}, 0)$. In order to prove a concentration inequality for the variables X_{ij} we begin by studying their moment generating functions $m_{X_{ij}}(t)$. Note that for every real number t the equality

$$\exp(tX_{ij}) = 1_{\{Y_{ij} \leq 0\}} + 1_{\{Y_{ij} \geq 0\}} \exp tY_{ij}$$

holds. Now $Y_{ij} = 1 + \frac{2\delta}{n} - 2\frac{n_{ij}}{N}$ where n_{ij} is a binomial random variable with parameters N and p_{ij} given by

$$p_{ij} := \begin{cases} p_t, \text{ if } \{i, j\} \subseteq C_t \\ 1 - q, \text{ else} \end{cases}$$

As a result taking expected values on both sides of the expression above we conclude that

$$m_{X_{ij}}(t) \leq \mathbb{P}\{Y_{ij} \leq 0\} + e^{t\left(1+\frac{2\delta}{n}\right)}\mathbb{E}\left(e^{-\frac{2t}{N}n_{ij}}1_{\{Y_{ij}\geq 0\}}\right).$$

Using the Cauchy-Schwartz inequality and the known formula for the moment generating function of a binomial random variable it follows that

$$m_{X_{ij}}(t) \leq \mathbb{P}\{Y_{ij} \leq 0\} + e^{t\left(1+\frac{2\delta}{n}\right)}\mathbb{E}\left(e^{-\frac{4t}{N}n_{ij}}\right)^{\frac{1}{2}}\mathbb{P}\{Y_{ij} \geq 0\}^{\frac{1}{2}} =$$

$$= \mathbb{P}\{Y_{ij} \leq 0\} + e^{t\left(1+\frac{2\delta}{n}\right)}\left(e^{-\frac{4t}{N}}p_{ij} + q_{ij}\right)^{\frac{N}{2}}\mathbb{P}\{Y_{ij} \geq 0\}^{\frac{1}{2}}$$

where $q_{ij} := 1 - p_{ij}$. By Hoeffding's inequality on Bernoulli random variables we know that

$$\mathbb{P}\{Y_{ij} \geq 0\} \leq \exp\left(-\frac{N}{2}\left(-\frac{2\delta}{n} - (1-2p_{ij})\right)^2\right) \leq \exp\left(-2N(\alpha - \delta/n)^2\right)$$

so if $t = aN$ the inequality

$$e^{t\left(1+\frac{2\delta}{n}\right)}\exp\left(-N(\alpha-\delta/n)^2\right) \leq 1$$

holds whenever $a \leq \frac{(\alpha-\delta/n)^2}{\left(1+\frac{2\delta}{n}\right)}$ and for all such a we have

$$m_{X_{ij}}(aN) \leq \left(1 + (e^{-4a}p_{ij} + q_{ij})^{\frac{N}{2}}\right)$$

We define $a(\delta) := \frac{(\alpha-\delta/n)^2}{\left(1+\frac{2\delta}{n}\right)}$ and will use it to prove a moment concentration inequality for $b(\Gamma, \delta)$ which will give us a bound on the second term in (5). For every $t > 0$ we have

$$\mathbb{P}\left\{b(\Gamma,\delta) \geq \delta\right\} = \mathbb{P}\left\{\exp\left(t\sum_{i\neq j} X_{ij}\right) \geq e^{t\delta}\right\} \leq e^{-t\delta}\prod_{i\neq j}\mathbb{E}[e^{tX_{ij}}] = e^{-t\delta}m_{X_{ij}}(t)$$

Choosing $t = a(\delta)N$ and using the previous inequality we see that

$$\mathbb{P}\left\{b(\Gamma,\delta) \geq \delta\right\} \leq e^{-N\delta a(\delta)}\prod_{i\neq j}\left(1 + (e^{-4a}p_{ij} + q_{ij})^{\frac{N}{2}}\right)$$

Which decreases exponentially in N for any $0 \leq \delta \leq n\alpha$. The rate of decrease of the first term is controlled by the positive factor

$$\delta a(\delta) = \frac{\delta\left(\alpha - \frac{\delta}{n}\right)^2}{\left(1+\frac{2\delta}{n}\right)}.$$

and we let δ^* be a maximizer of this function. $\square$

10. Uniqueness of the minimizer A^*

♣♣♣ Daniel: [The following section is deprecated in favor of the alternative proof of uniqueness.]

In this section, we address the question of whether or not the matrix A^* is the unique minimizer of the optimization problem $min_A \Delta(A)$. We begin by stating two lemmas.

Lemma 10.1. *Let C_1 and C_2 be two convex sets in R^m for some m. Denote the relative interior of a set C by Relint(C) Then,*

$$Relint(C_1) + Relint(C_2) \subseteq Relint(C_1 + C_2).$$

Lemma 10.2. *If $\langle N, Q\rangle \leq 0$ for every Q in a convex set C which contains 0 in its relative interior then the equality $\langle N, Q\rangle = 0$ holds for every $Q \in C$.*

Proof. Because of the inequality in our assumptions the set

$$F := \{Q \in C : \langle N, Q\rangle = 0\}$$

is a face of the convex set C. Moreover this face contains $Q = 0$ which is a point in the relative interior of C. Since the only face of a convex set which intersects its relative interior is C itself we conclude that $F = C$ as claimed. □

We will now prove that with very high probability, A^* is the unique minimizer of $min_A \Delta(A)$. this fact will follow from the next theorem. We begin with some notation. Let C denote the subdifferential of Δ at A^*. $\partial(\Delta(A))(A^*)$. since $\Delta(A)$ is a convex function, C a convex region. Recall that

$$\Delta(A) = \delta\|2A - 11^t\|_* + \frac{1}{N}\sum_{k=1}^{N}\|A - B_k\|_1$$

And define functions f, g:

$$f := \frac{1}{N}\sum_{k=1}^{N}\|A - B_k\|_1, \quad g := \delta\|2A - 11^t\|_*$$

In this vocabulary, we have that

$$C = \partial(f)(A^*) + \partial(g)(A^*) \tag{6}$$

Theorem 10.3. *Assume that $A^* \in argmin\Delta(A)$ and that $0 \in Relint(C)$. Then, A^* is the unique minimizer of the optimization problem $min_A \Delta(A)$.*

Proof. Suppose there exist a minimizer $\bar{A} \in argmin\ \Delta(A)$ with $\bar{A} \neq A^*$. Let $N := \bar{A} - A^*$. We will prove that $N = 0$. Let $Q \in C$. By definition of a subdifferential in A^* we have that

$$\Delta(\bar{A}) \geq \Delta(A^*) + \langle Q, \bar{A} - A^*\rangle$$

Since both $\bar{A}$ and A^* are minimizers of $\Delta(A)$ we have that $\Delta(\bar{A}) = \Delta(A^*)$ and it follows that

$$\langle Q, N\rangle \leq 0.$$

For all $Q \in C$. since $0 \in Relint(C)$ by hypothesis, lemma 10.2 implies that

$$\langle Q, N\rangle = 0.$$

For all $Q \in C$.

Now, observe the key following fact. By the characterizations of the subdifferentials $\partial(f)(A^*)$ and $\partial(g)(A^*)$ discussed in the previous section, we have that $\frac{-2\delta}{n}H^* \in \partial(g)(A^*)$ and that $H^* \in \partial(f)(A^*)$. By (6) we have that $H^* + \frac{-2\delta}{n}H^* \in C$ which implies that

$$\left\langle H^* + \frac{-2\delta}{n}H^*, N \right\rangle = 0$$

and

$$\langle H^*, N\rangle = \left\langle \frac{-2\delta}{n}H^*, N \right\rangle = 0.$$

For $T \in \partial(f)(A^*)$ we have that $T + \frac{-2\delta}{n}H^* \in C$ and

$$0 = \left\langle T + \frac{-2\delta}{n}H^*, N \right\rangle = \langle T, N\rangle + \left\langle \frac{-2\delta}{n}H^*, N \right\rangle = \langle T, N\rangle .$$

Using this same argument with H^* instead of $\frac{-2\delta}{n}H^*$ we have that for all $S \in \partial(g)(A^*)$, $\langle S, N\rangle = 0$.

The two previous facts are the key to show that $N = 0$. To prove this, we will carefully select some members of $\partial(f)(A^*)$ and of $\partial(g)(A^*)$ and use them to show that the entries of N are all 0.

First of all, let D_i be a $n \times n$ diagonal matrix with entries equal to 0 except in the entry ii where we set it to $1 > \epsilon > 0$.

$$D = \begin{bmatrix} 0 & & & \\ & \epsilon & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}$$

Recall that the only restriction for a entry diagonal of a matrix in $\partial(f)(A^*)$ is that its absolute value is bounded by n and that $H^* \in \partial(f)(A^*)$ The matrix $H^* + D_i$ has the same entries of H^* except in the entry ii where its value is $H^*_{ii} + \epsilon = 1 + \epsilon < n$ so that $H^* + D_i \in \partial(f)(A^*)$.

Then,

$$0 = \langle H^* + D_i, N\rangle = \langle H^*, N\rangle + \langle D, N\rangle = \langle D, N\rangle = \epsilon N_{ii}.$$

which implies that $N_{ii} = 0$ for all $1 \le i \le n$. Note that in particular $Tr(N) = 0$.

Let G_I be the set of pairs $\{i, j\}$ such that $\{i, j\} \in I$, $i \neq j$ and $\frac{n_0(ij)-n_1(ij)}{N} < 1$. Let $\{i, j\} \in G_I$ and chose $\epsilon > 0$ small enough such that $\frac{n_0(ij)-n_1(ij)}{N} + \epsilon < 1$. Now, define the matrix $T^{i,j}$ as a matrix of all zeros except at its entries ij, ji where its value is set to ϵ. I.e, $T^{i,j}$ is of the form

$$T^{ij} = \begin{bmatrix} 0 & & & & \\ & 0 & & \epsilon & \\ & & \ddots & & \\ & \epsilon & & 0 & \\ & & & & 0 \end{bmatrix}$$

Notice that the matrix $H^* + T^{ij}$ belongs to $\partial(f)(A^*)$ since all its entries all equal to those of H^*, except at ij and ji, where its entries are $1-\epsilon$ and therefore

$$\frac{n_0(ij) - n_1(ij)}{N} \leq H^*_{ij} + T^{ij}_{ij} < 1$$

We obtain that

$$0 = \left\langle H^* + T^{ij}, N \right\rangle = \langle H^*, N \rangle + \left\langle T^{ij}, N \right\rangle = \left\langle T^{ij}, N \right\rangle = \epsilon N_{ij} + \epsilon N_{ji}.$$

By the symmetry of N it follows that $n_{ij} = n_{ji} = 0 \; \forall \; \{ij\} \in G_I$. Define the set G_O as the set of pairs $\{i,j\}$ such that $\{i,j\} \in O$, $i \neq j$ and $-1 > \frac{n_0(ij)-n_1(ij)}{N}$. Using the same argument as before with $-\epsilon$ instead of ϵ, we can show that $n_{ij} = n_{ji} = 0 \; \forall \; \{ij\} \in G_O$.

It only remains to show that that entries of N which are not in G_I or in G_O are 0. Define the set B_I as $I \setminus G_I$. let $\{i,j\} \in B_I$. This means that $\{i,j\} \in I$ but $\frac{n_0(ij)-n_1(ij)}{N} = 1$. Define the symmetric matrix S^{ij} as follows:

$$S^{ij}_{kl} = \begin{cases} 0 \text{ if } \{k,l\} \in G_I \text{ or if } \{k,l\} \in G_O. \\ 2\delta \text{ if } k = l \text{ but } k \neq i \text{ or } k \neq j. \\ 2\delta \text{ if } \{k,l\} = \{i,j\}. \\ 0 \text{ if } i = k = l \text{ or if } j = k = l. \end{cases} \tag{7}$$

In other words, S^{ij} is of the form

$$S^{ij} = \begin{bmatrix} 2\delta & & & & & & \\ & 2\delta & & & & & \\ & & 0 & & & 2\delta & \\ & & & 2\delta & & & \\ & & & & \ddots & & \\ & & 2\delta & & & 0 & \\ & & & & & & 2\delta \end{bmatrix}$$

First of all, observe that $\{ij\} \in I$ so $H^*_{ij} = H^*_{ji} = 1$, and as the entries in the diagonal of H^* are also 1, we have that

$$\left\langle H^*, S^{ij} \right\rangle = 2\delta(Tr(H^*) - H^*_{ii} - H^*_{jj}) + 2\delta H_{ij} + 2\delta H_{ji} = (n-2)2\delta + 2(2\delta) = 2\delta.$$

Moreover, the spectral norm of S_{ij} is bounded by $\sqrt{\|S^{ij}\|_1 \|S^{ij}\|_\infty}$ by holder inequality. Here these norms are the induced 1 norm and the induced ∞ norm. As the sum of every row and every column of S^{ij} is 2δ, we obtain that

$$\|S^{ij}\| \leq \sqrt{(2\delta)(2\delta)} = 2\delta.$$

We conclude that $S^{ij} \in \partial(g)(A^*)$ and so,

$$0 = \left\langle S^{ij}, N \right\rangle = 2\delta(Tr(N) - N_{ii} - N_{jj}) + 2\delta N_{ij} + 2\delta N_{ji} = 4\delta N_{ij}.$$

The last equality is given by the fact that N is Symmetric. this means that $N_{ij} = Nji = 0$ for all $i,j \in B_I$. Finally, define B_O as $O \setminus G_O$. let $\{i,j\} \in B_O$. This means that $\{i,j\} \in O$ but $\frac{n_0(ij)-n_1(ij)}{N} = -1$. A similar argument using -2δ in the off diagonal entries shows that $N_{ij} = Nji = 0$ for all $i,j \in B_O$.

To summarize, we have shown that all entries of N must equal 0 since any entry belongs to its diagonal,G_I, G_O, B_I or B_O are 0 and these sets are a partition of the entries of N. We conclude that $\bar{A} = A^*$. $\square$

Departamento de matemáticas, Universidad de los Andes, Carrera $1^{\text{ra}}\#18A-12$, Bogotá, Colombia

E-mail address: d.de1033@uniandes.edu.co

Departamento de matemáticas, Universidad de los Andes, Carrera $1^{\text{ra}}\#18A-12$, Bogotá, Colombia

E-mail address: mvelasco@uniandes.edu.co