Edge-Negative Association in Random Forests and Connected Subgraphs

Xian-Yuan WU

Capital Normal University

wuxy@mail.cnu.edu.cn

August 10–14, 2004, RCS, BNU

Contents

- Definitions

Contents

- Definitions
- Backgrounds

Contents

- Definitions
- Backgrounds
- The Problem

Contents

Contents

Contents

Def: Graph and its Subgraph

- Subgraph: Let $G = (V, E)$ be a finite graph, graph $G' = (V', E')$ is called a subgraph of G, if $V' \subset V$, $E' \subset E$.

Def: Graph and its Subgraph

- Subgraph: Let $G = (V, E)$ be a finite graph, graph $G' = (V', E')$ is called a subgraph of G, if $V' \subset V$, $E' \subset E$.

- $G' = (V', E')$ is called a spanning subgraph of G, if $V' = V$. Denote by $\mathcal{G}$ the set of all spanning subgraphs of G.

Def: Graph and its Subgraph

- Subgraph: Let $G = (V, E)$ be a finite graph, graph $G' = (V', E')$ is called a subgraph of G, if $V' \subset V$, $E' \subset E$.

- $G' = (V', E')$ is called a spanning subgraph of G, if $V' = V$. Denote by $\mathcal{G}$ the set of all spanning subgraphs of G.

- Let $\mathcal{E}$ be the set of all subsets of E, $\Omega := \{0, 1\}^E$, then $\mathcal{E}$, Ω and $\mathcal{G}$ can be seen as the same.

Def: Spanning Subgraph

- Spanning Tree: Let $T \in \mathcal{G}$, if T is connected and there is no circuit in T, we call T a spanning tree. Denote by $\mathcal{T}$ the set of all spanning trees of G.

Def: Spanning Subgraph

- Spanning Tree: Let $T \in \mathcal{G}$, if T is connected and there is no circuit in T, we call T a spanning tree. Denote by $\mathcal{T}$ the set of all spanning trees of G.

- Spanning Forest: $F \in \mathcal{G}$, F has no circuit, then F is called spanning forest. Denote by $\mathcal{F}$ the set of all spanning forests of G.

Def: Spanning Subgraph

- Spanning Tree: Let $T \in \mathcal{G}$, if T is connected and there is no circuit in T, we call T a spanning tree. Denote by $\mathcal{T}$ the set of all spanning trees of G.

- Spanning Forest: $F \in \mathcal{G}$, F has no circuit, then F is called spanning forest. Denote by $\mathcal{F}$ the set of all spanning forests of G.

- Connected Subgraph: An element C of $\mathcal{G}$ is called a connected subgraph of G, if G is connected. The set of all connected subgraphs of G is denoted by $\mathcal{C}$.

Def: Spanning Subgraph

- Spanning Tree: Let $T \in \mathcal{G}$, if T is connected and there is no circuit in T, we call T a spanning tree. Denote by $\mathcal{T}$ the set of all spanning trees of G.

- Spanning Forest: $F \in \mathcal{G}$, F has no circuit, then F is called spanning forest. Denote by $\mathcal{F}$ the set of all spanning forests of G.

- Connected Subgraph: An element C of $\mathcal{G}$ is called a connected subgraph of G, if G is connected. The set of all connected subgraphs of G is denoted by $\mathcal{C}$.

- Obviously, $\mathcal{T} = \mathcal{F} \cap \mathcal{C}$.

Def: Random Subgraph

- Random Subgraph: If $\mathbf{G}$ is randomly chosen from $\mathcal{G}$ in law μ, a probability measure on $\mathcal{G}$, we call $\mathbf{G}$ a random subgraph, and μ is its distribution.

Def: Random Subgraph

- Random Subgraph: If $\mathbf{G}$ is randomly chosen from $\mathcal{G}$ in law μ, a probability measure on $\mathcal{G}$, we call $\mathbf{G}$ a random subgraph, and μ is its distribution.
- If $\mathbf{T}$, $\mathbf{F}$ and $\mathbf{C}$ are randomly chosen from $\mathcal{T}$, $\mathcal{F}$ and $\mathcal{C}$ respectively, we call them random spanning tree, spanning forest and connected subgraph.

Def: Random Subgraph

- Random Subgraph: If $\mathbf{G}$ is randomly chosen from $\mathcal{G}$ in law μ, a probability measure on $\mathcal{G}$, we call $\mathbf{G}$ a random subgraph, and μ is its distribution.

- If $\mathbf{T}$, $\mathbf{F}$ and $\mathbf{C}$ are randomly chosen from $\mathcal{T}$, $\mathcal{F}$ and $\mathcal{C}$ respectively, we call them random spanning tree, spanning forest and connected subgraph.

- $\mathbf{T}$, $\mathbf{F}$ and $\mathbf{C}$ are called uniform spanning tree, spanning forest and connected subgraph, if their distributions are uniform distribution on $\mathcal{T}$, $\mathcal{F}$ and $\mathcal{C}$ respectively.

Def: (edge-)negative association

- A random subgraph $\mathbb{S}$ in $\mathcal{E}$ is called edge-negatively associated, if $\forall\ e, f \in E$, $e \neq f$,

$$\mathbb{P}(e, f \in \mathbf{S}) \leq \mathbb{P}(\mathbf{e} \in \mathbf{S})\mathbb{P}(\mathbf{f} \in \mathbf{S}).$$

Def: (edge-)negative association

- A random subgraph $\mathbf{S}$ in $\mathcal{E}$ is called edge-negatively associated, if $\forall\ e, f \in E$, $e \neq f$,

$$\mathbb{P}(e, f \in \mathbf{S}) \leq \mathbb{P}(\mathbf{e} \in \mathbf{S})\mathbb{P}(\mathbf{f} \in \mathbf{S}).$$

- We call a random subgraph $\mathbf{S}$ negatively associated if

$$\mathbb{P}(\mathbf{S} \in A \cap B) \leq \mathbb{P}(\mathbf{S} \in A)\mathbb{P}(\mathbf{S} \in B)$$

for all pairs A, B of increasing events with the **property that there exists $E' \subset E$ such that A is defined on E' and B is defined on its complement $\overline{E'} = E \setminus E'$.**

Def: (edge-)negative association

- A random subgraph $\mathbb{S}$ in $\mathcal{E}$ is called edge-negatively associated, if $\forall\ e, f \in E$, $e \neq f$,

$$\mathbb{P}(e, f \in \mathbf{S}) \leq \mathbb{P}(\mathbf{e} \in \mathbf{S})\mathbb{P}(\mathbf{f} \in \mathbf{S}).$$

- We call a random subgraph $\mathbb{S}$ negatively associated if

$$\mathbb{P}(\mathbf{S} \in A \cap B) \leq \mathbb{P}(\mathbf{S} \in A)\mathbb{P}(\mathbf{S} \in B)$$

  for all pairs A, B of increasing events with the **property that there exists $E' \subset E$ such that A is defined on E' and B is defined on its complement $\overline{E'} = E \setminus E'$.**

- Here increasing is defined as usual. We call event A defined on E', if $\omega'(e) = \omega(e)$ for all $e \in E'$, then either $\omega, \omega' \in A$ or $\omega, \omega' \notin A$.

Bg: RC Measure

- Random-Cluster Measure:

  For $p \in [0,1]$, $q > 0$, the RC Measure $\phi_{p,q}$ is defined in $\Omega := \{0,1\}^E$ as

$$\phi_{p,q}(\omega) = \frac{1}{Z(p,q)} \left\{ \prod_{e \in E} p^{\omega(e)} (1-p)^{1-\omega(e)} \right\} q^{k(\omega)}, \quad \omega \in \Omega.$$

Bg: RC Measure

- Random-Cluster Measure:

  For $p \in [0,1]$, $q > 0$, the RC Measure $\phi_{p,q}$ is defined in $\Omega := \{0,1\}^E$ as

$$\phi_{p,q}(\omega) = \frac{1}{Z(p,q)} \left\{ \prod_{e \in E} p^{\omega(e)} (1-p)^{1-\omega(e)} \right\} q^{k(\omega)}, \quad \omega \in \Omega.$$

- $Z(p,q)$ is the usual partition function. $k(\omega)$ is the number of connected components of subgraph $G(\omega) = (V, E(\omega))$, where $E(\omega) = \{e \in E : \omega(e) = 1\}$.

Bg: RC Measure

- Note that when $q = 1$, Random-Cluster Measure is the usual percolation measure with parameter p; when $q = 2, 3, \ldots$, Random-Cluster Model relates the Ising (Potts) Model in the following way

$$\pi(\sigma_i = \sigma_j) - \frac{1}{q} = (1 - \frac{1}{q})\phi_{p,q}(i \longleftrightarrow j), \quad i, j \in V,$$

  where $\sigma \in \{1, 2, \ldots, q\}^V$, and the l.h.s. of the above equation is the corresponding order parameter of Ising(Potts) Model.

Bg:FKG Inequality of RCM

- When $q \geq 1$, $\phi_{p,q}$ satisfies the following famous "lattice FKG condition":

$$\phi_{p,q}(\omega_1 \vee \omega_2)\phi_{p,q}(\omega_1 \wedge \omega_2) \geq \phi_{p,q}(\omega_1)\phi_{p,q}(\omega_2)$$

  and then has the following FKG Inequality

$$\phi_{p,q}(A \cap B) \geq \phi_{p,q}(A)\phi_{p,q}(B),$$

  for all increasing $A, B \subset \Omega$.

Bg:FKG Inequality of RCM

- When $q \geq 1$, $\phi_{p,q}$ satisfies the following famous "lattice FKG condition":

$$\phi_{p,q}(\omega_1 \vee \omega_2)\phi_{p,q}(\omega_1 \wedge \omega_2) \geq \phi_{p,q}(\omega_1)\phi_{p,q}(\omega_2)$$

  and then has the following FKG Inequality

$$\phi_{p,q}(A \cap B) \geq \phi_{p,q}(A)\phi_{p,q}(B),$$

  for all increasing $A, B \subset \Omega$.
- What can we say for the case: $0 < q < 1$?

Bg:FKG Inequality of RCM

- When $q \geq 1$, $\phi_{p,q}$ satisfies the following famous "lattice FKG condition":

$$\phi_{p,q}(\omega_1 \vee \omega_2)\phi_{p,q}(\omega_1 \wedge \omega_2) \geq \phi_{p,q}(\omega_1)\phi_{p,q}(\omega_2)$$

  and then has the following FKG Inequality

$$\phi_{p,q}(A \cap B) \geq \phi_{p,q}(A)\phi_{p,q}(B),$$

  for all increasing $A, B \subset \Omega$.

- What can we say for the case: $0 < q < 1$?
- It seems that in this case, RCM is not positively associated. On the contrary, it maybe negatively associated!

Bg: Limits of RCM

- As p or q tends to 0, the Random-Cluster Measure has the following three types of limits:

Bg: Limits of RCM

- As p or q tends to 0, the Random-Cluster Measure has the following three types of limits:
- Fix $p = \frac{1}{2}$ and let q tend to 0, $\phi_{p,q} \Rightarrow \mu_{U,\mathcal{C}}$, the uniform distribution on $\mathcal{C}$.

Bg: Limits of RCM

- As p or q tends to 0, the Random-Cluster Measure has the following three types of limits:
- Fix $p = \frac{1}{2}$ and let q tend to 0, $\phi_{p,q} \Rightarrow \mu_{U,\mathcal{C}}$, the uniform distribution on $\mathcal{C}$.
- Let p, q and q/p tend to 0 at the same time, $\phi_{p,q} \Rightarrow \mu_{U,\mathcal{T}}$, the uniform distribution on $\mathcal{T}$.

Bg: Limits of RCM

- As p or q tends to 0, the Random-Cluster Measure has the following three types of limits:
- Fix $p = \frac{1}{2}$ and let q tend to 0, $\phi_{p,q} \Rightarrow \mu_{U,\mathcal{C}}$, the uniform distribution on $\mathcal{C}$.
- Let p, q and q/p tend to 0 at the same time, $\phi_{p,q} \Rightarrow \mu_{U,\mathcal{T}}$, the uniform distribution on $\mathcal{T}$.
- When $p = q$ and tends to 0, $\phi_{p,q} \Rightarrow \mu_{U,\mathcal{F}}$, the uniform distribution on $\mathcal{F}$.

Bg: Limits of RCM

- As p or q tends to 0, the Random-Cluster Measure has the following three types of limits:
- Fix $p = \frac{1}{2}$ and let q tend to 0, $\phi_{p,q} \Rightarrow \mu_{U,\mathcal{C}}$, the uniform distribution on $\mathcal{C}$.
- Let p, q and q/p tend to 0 at the same time, $\phi_{p,q} \Rightarrow \mu_{U,\mathcal{T}}$, the uniform distribution on $\mathcal{T}$.
- When $p = q$ and tends to 0, $\phi_{p,q} \Rightarrow \mu_{U,\mathcal{F}}$, the uniform distribution on $\mathcal{F}$.
- Recall that $\mathcal{C}$, $\mathcal{T}$ and $\mathcal{F}$ are the set of all connected subgraphs, spanning trees and spanning forests respectively.

Bg: UST is negatively associated

- Theorem A: For any finite graph G, the uniform spanning tree of G is negatively associated.

Bg: UST is negatively associated

- Theorem A: For any finite graph G, the uniform spanning tree of G is negatively associated.
- Theorem A suggests that, for p, q and also q/p small enough, the RCM should be negatively associated!

Bg: UST is negatively associated

- Theorem A: For any finite graph G, the uniform spanning tree of G is negatively associated.
- Theorem A suggests that, for p, q and also q/p small enough, the RCM should be negatively associated!
- Note that the proof of Theorem A can be found in the following papers:

[1] Benjamini, I., Lyons, R., Peres, Y., Schramm,O. *Uniform spanning forests,* Ann. Prob. 29 (2001), pp 1-65.

[2] T. Feder, M. Mihail *Balanced matroids*, Proceeding of the 24th ACM Symposium on the Theory of Computing, 1992, pp 26-38.

Bg: Conjecture on USF $\&$ UCS

- Conjecture B: For any finite graph G, the uniform spanning forest and connected subgraph of G are all negatively associated.

Bg: Conjecture on USF $\&$ UCS

- Conjecture B: For any finite graph G, the uniform spanning forest and connected subgraph of G are all negatively associated.
- The following is an attempt to solve conjecture B:

  Theorem C: If $G = (V, E)$ has eight or fewer vertices, or has nine vertices and eighteen or fewer edges, then the uniform forest $\mathbf{F}$ has the edge-negative-association property.

Bg: Conjecture on USF $\&$ UCS

- Conjecture B: For any finite graph G, the uniform spanning forest and connected subgraph of G are all negatively associated.
- The following is an attempt to solve conjecture B:

  Theorem C: If $G = (V, E)$ has eight or fewer vertices, or has nine vertices and eighteen or fewer edges, then the uniform forest $\mathbf{F}$ has the edge-negative-association property.
- Theorem C is due to G. R. Grimmett and S. N. Winkler, for details one may see the following [3]

[3] G. R. Grimmett and S. N. Winkler *Negative Association in Uniform Forests and Connected Graphs*, to appear in Random Structures and Algorithms, 2004.

The Problem

- Theorem C and its proof told us that giving a complete proof to Conjecture B is really a hard work.
  In a sense, the uniform measure maybe the simplest one on $\mathcal{F}$, but what about the others? In other word, there arises a problem:
  Does there really exist a probability measure on $\mathcal{F}$ or $\mathcal{C}$, which satisfies the edge-negative association property?

The Problem

- Theorem C and its proof told us that giving a complete proof to Conjecture B is really a hard work.
  In a sense, the uniform measure maybe the simplest one on $\mathcal{F}$, but what about the others? In other word, there arises a problem:
  Does there really exist a probability measure on $\mathcal{F}$ or $\mathcal{C}$, which satisfies the edge-negative association property?

- The answer is yes. For any $F \in \mathcal{F}$, the dirac measure δ_F is a good candidate. In addition, $\mu_{U,\mathcal{T}}$, the uniform probability measure on $\mathcal{T}$, as a probability measure on $\mathcal{F}$, is also edge negatively associated.
  But δ_F's and $\mu_{U,\mathcal{T}}$ are all singular in the sense that they are supported by a "small" subset of $\mathcal{F}$.

The Problem

- Let $\mathcal{M}^{+}(\mathcal{F})$ (resp. $\mathcal{M}^{+}(\mathcal{C})$) be the set of all probability measures on $\mathcal{F}$ (resp. $\mathcal{C}$) which put every element of $\mathcal{F}$ (resp. $\mathcal{C}$) positive charge.

  So, our problem should be the following:

  Does there really exist a probability measure in $\mathcal{M}^{+}(\mathcal{F})$ (resp. $\mathcal{M}^{+}(\mathcal{C})$), which satisfies the edge-negative-association property?

Main results

- **Theorem 1:** *Let* $\mathbf{p} = (p_e : e \in E) \in (0,1)^E$ *be a vector, and* $\mathbf{G}$ *be a random subgraph of* G. *Let* $\mathbf{G}_{\mathbf{p}}$ *be the random subgraph obtained by removing each edge* e *of* $\mathbf{G}$ *independently with probability* $1 - p_e$. *Then, the edge-negative-association property of* $\mathbf{G}$ *implies the same property of* $\mathbf{G}_{\mathbf{p}}$.

Main results

- **Theorem 1:** *Let* $\mathbf{p} = (p_e : e \in E) \in (0,1)^E$ *be a vector, and* $\mathbf{G}$ *be a random subgraph of* G*. Let* $\mathbf{G_p}$ *be the random subgraph obtained by removing each edge* e *of* $\mathbf{G}$ *independently with probability* $1 - p_e$*. Then, the edge-negative-association property of* $\mathbf{G}$ *implies the same property of* $\mathbf{G_p}$*.*

- **Corollary 2:** *Take* $\mathbf{G} = \mathbf{T}$*, the uniform spanning tree of* G*. For any* $\mathbf{p} \in (0,1)^E$*, define* $\mathbf{F} := \mathbf{G_p}$ *as in theorem 1, denote by* $\mu_{\mathbf{p}}$ *the distribution of* $\mathbf{F}$*. Then*
  - i, $\mathbf{F}$ *has edge-negative-association property, and* $\mu_{\mathbf{p}} \in \mathcal{M}^+(\mathcal{F})$*.*
  - ii, *Furthermore, for any* $\phi \neq E' \subset E$*, we have*

$$\mathbb{P}(E' \subset \mathbf{F}) \leq \prod_{e \in E'} \mathbb{P}(e \in \mathbf{F}).$$

Main results

- **Theorem 3:** *Let* $\mathbf{p} \in (0,1)^E$ *be a vector, and* $\mathbf{G}$ *be a random subgraph of* G. *Let* $\mathbf{G}^{\mathbf{p}}$ *be the random subgraph obtained by adding each edge* $e(\notin \mathbf{G})$*to* $\mathbf{G}$ *independently with probability* p_e. *Then, the edge-negative-association property of* $\mathbf{G}$ *implies the same property of* $\mathbf{G}^{\mathbf{p}}$.

Main results

- **Theorem 3:** *Let* $\mathbf{p} \in (0,1)^E$ *be a vector, and* $\mathbf{G}$ *be a random subgraph of* G*. Let* $\mathbf{G}^{\mathbf{p}}$ *be the random subgraph obtained by adding each edge* $e(\notin \mathbf{G})$*to* $\mathbf{G}$ *independently with probability* p_e*. Then, the edge-negative-association property of* $\mathbf{G}$ *implies the same property of* $\mathbf{G}^{\mathbf{p}}$*.*

- **Corollary 4:** *Take* $\mathbf{G} = \mathbf{T}$*, the uniform spanning tree of* G*. For any* $\mathbf{p} \in (0,1)^E$*, define* $\mathbf{C} := \mathbf{G}^{\mathbf{p}}$ *as in theorem 3, denote by* $\mu^{\mathbf{p}}$ *the distribution of* $\mathbf{C}$*. Then*
  - i, $\mathbf{C}$ *has edge-negative-association property, and* $\mu^{\mathbf{p}} \in \mathcal{M}^+(\mathcal{C})$*.*
  - ii, *Furthermore, for any* $\phi \neq E' \subset E$*, we have*

$$\mathbb{P}(E' \subset \mathbf{C}) \leq \prod_{e \in E'} \mathbb{P}(e \in \mathbf{C}).$$

About the proofs

- To prove Theorem 1 and Theorem 3, it suffices to check them directly.

About the proofs

- To prove Theorem 1 and Theorem 3, it suffices to check them directly.
- By Theorem 1 and Theorem A, Corollary 2 follows immediately.

About the proofs

- To prove Theorem 1 and Theorem 3, it suffices to check them directly.
- By Theorem 1 and Theorem A, Corollary 2 follows immediately.
- By Theorem 3 and Theorem A, Corollary 4 follows in an analogous way. Note that a little more care is needed when checking Corollary 4(ii).

An open problem

- Suppose $G = (V, E)$ be finite graph, we call G forest-separable, if there exists $F_1, F_2 \in \mathcal{F}$ such that $F_1 \cup F_2 = G$, $F_1 \cap F_2 = \phi$. Here we see F_1, F_2 as subsets of E, and call the non-ordered pair (F_1, F_2) a separation of G.

An open problem

- Suppose $G = (V, E)$ be finite graph, we call G forest-separable, if there exists $F_1, F_2 \in \mathcal{F}$ such that $F_1 \cup F_2 = G$, $F_1 \cap F_2 = \phi$. Here we see F_1, F_2 as subsets of E, and call the non-ordered pair (F_1, F_2) a separation of G.
- When G is a forest-separable graph with $|v| = n$, then $|E| \leq 2(n-1)$.

An open problem

- Suppose $G = (V, E)$ be finite graph, we call G forest-separable, if there exists $F_1, F_2 \in \mathcal{F}$ such that $F_1 \cup F_2 = G$, $F_1 \cap F_2 = \phi$. Here we see F_1, F_2 as subsets of E, and call the non-ordered pair (F_1, F_2) a separation of G.
- When G is a forest-separable graph with $|v| = n$, then $|E| \leq 2(n-1)$.
- If G is forest-separable, $\forall\ e, f \in E$ with $e \neq f$, let $I_{e,f}$ be the number of separations with e, f stay in the same separated forest(F_1 or F_2), denote by I the number of all separations of G.

An open problem

- We propose our problem as follows

  Conjecture D: For all forest-separable graph $G = (V, E)$, $I_{e,f} \leq \frac{1}{2} I$ for all $e, f \in E$ with $e \neq f$.

An open problem

- We propose our problem as follows

  Conjecture D: For all forest-separable graph $G = (V, E)$, $I_{e,f} \leq \frac{1}{2} I$ for all $e, f \in E$ with $e \neq f$.

- **Remark:** *Conjecture D implies Conjecture B (in USF case).*

******

Thank You !