Válasz Backhausz Ágnes bírálatára

Köszönöm a hasznos megjegyzéseket és észrevételeket. A konkrét kérdésekre adott válaszaim a következők.

1. A módszereket generált általánosított véletlen és konstruált általánosított kvázivéletlen gráfokon, továbbá valós éltbeli adatokon is kipróbáltuk, ezekről részletesebben a PhD diákommal közös [Bol-El15] és [Bol-El16] cikkekben írunk. Az EM-algoritmus alkalmazása során az eredmény nagy mértékben függött a kezdő klaszterezéstől, és gyorsan konvergált, ha a spektrális klaszterekből indultunk ki. Összeségében azt tapasztaltuk, hogy a konstruált általánosított kázivéletlen gráfsorozatokon sokkal nagyobb n esetén volt látványos a spektrális rés kialakulása, mint az általánosított véletlen gráfsorozatoknál; ez arra utalt, hogy előbbinél a konvergencia lassabb, amit 3. fejezetbeli számolásaim is alátámasztanak. A valós életbeli gráfoknál, ha n 'nagy' volt, akkor tényleg sikerült a spektrális rés alapján viszonylag kis diszkrepanciát elérni. Viszont, ha nincs sorozat, akkor a kis- és nagy-ordóknak nincs sok értelme, egyedi gráfból nehéz messzemenő következtetéseket levonni.

   A nagyméretű gráfokra vagy tömbökre alkalmazott algoritmusok legidőigényesebb része az SVD. Régebben, NKFP projektek kapcsán felmerült az SVD-algoritmus gyorsításának igénye. Erre az irodalomban tárgyalt véletlenített eljárásokat javasoltuk, melyek lényege, hogy alacsony rangú közelítés esetén a közelítés hibájánál nem nagyobb nagyságrendű perturbációk (melyek képesek ritkítani vagy digitalizálni a mátrixot) alkalmazhatók, azaz az egyszerűbbnek tűnő, de zajosított mátrioxot használjuk. Erre a célra alkalmaztuk a disszertációban tárgyalt perturbációs becsléseket, de az algoritmusok numerikus vonatkozásait a továbbiakban nem vizsgáltam.

2. Az álalánosított kvázivéletlen tulajdonságokkal kapcsolatos bizonyításoknál fontos volt a gráfsorozat általános tagjának egy lépcsős grafontól való eltérését becsülni vágás-normában. Ehhez használtam a vágás-norma tulajdonságait és a bírálatban idézett [3] cikk eredményeit. A nehézség inkább az volt, hogy a többrészes diszkrepancia csak a gráfból számolódik, és nem tartalmaz konstansokat. Így magát a lépcsős grafont is meg kellett konstruálni a klasztereken belüli és közti élsűrűségek segítségével, melyek nem szerepelnek a többrészes diszkrepancia definíciójában. Miután a részek térfogatával szintén normálok, könnyen lehet, hogy a jumble norma használatával egyszerűsíthetők a számolások. Ezt a normát sajnos a disszertáció írásakor még nem ismertem.

3. A disszertációban nem, de régebben írt Wiley könyvemben (2013) kitértem a sztochasztikus blokkmodellekkel kapcsolatos konzisztencia eredményekre. A disszertációbeli általánosított véletlen gráf modellben a $\mathbf{P}$ ún. valószínűség-mátrixot fixen tartom, és így tárgyalom a sajátértékek, spektrális klaszterek és k-részes diszkrepancia 1 valószínűséggel teljesülő tulajdonságait, ha $n \to \infty$. Nálam a klaszterek mindig rekonstruálhatók, ha a $\mathbf{P}$ mátrix rangja k.

   Sok cikk, pl. a bírálatban idézett [1] és [4] is, $\mathbf{P}$ elemeit n-el skálázza, és ilyen feltételek mellett bizonyítanak konzisztenciát, azaz ún. rekonstruálhatóságot. Erről bővebben írok a hiányolt hivatkozásoknál.

4. Ezt a kérdést Wiley könyvem utolsó fejezetében tárgyaltam részletesen, és ott az EM algoritmus teszteléséhez ténylegesen egy lépcsős grafonból generáltunk. A felírt likelihood-függvény nagyon hasonlít a [3] cikkbeli homomorfizmus-sűrűséghez. Ezt az algoritmust a disszertációban csak röviden, mint homogén blokkmodelt tárgyaltam (egyes cikkek ezt is inhomogénnek tekintik, ha több osztály van), és inhomogén blokkmodell alatt az α-modell k-osztályos vátozatára kifejlesztettet értettem. Az α-modell már egyosztályos esetben is általánosabb, mint az Erdős–Rényi véletlen gráf, ami akkor adódna, ha a csúcsokhoz rendelt összes α_i paraméter ugyanaz lenne. Ezek az exponenciális eloszláscsaládbeli modellek akkor alkalmazhatók jól, ha a fokszámsorozat elégséges statisztika. Az inhomogén blokkmodellben csak a részgráfokra és a páros részgráfokra kell ennek teljesülnie. Egyszóval, az EM-algoritmus bármilyen olyan szituációra módosítható, amelyben van

egy alacsony rangú grafon. A grafon rangja k, ha a sztochasztikus blokkmodellt generáljuk belőle, és részenként $2k$, ha a k-osztályos α-modellt (mindkét esetben a rang csak k-tól, és nem n-től függ).

A hiányolt hivatkozásokat részben ismertem, de akkor vettem csak be azokat a disszertáció irodalomjegyzékébe, ha a disszertációban az azokban található eredményeket közvetlenül használtam vagy összevetettem saját eredményeimmel. Ezt részletesebben a következőkben szeretném kifejteni.

- A [3] L. Lovász, B. Szegedy 2006-os cikkra, akárcsak Lovász László és társszerzőinek egyéb gráfkonvergenciával kapcsolatos cikkeire a Wiley-nál megjelent könyvebmen hivatkoztam, mert ott tárgyaltam a kiegyensúlyozott minimális vágások tesztelhetőségét, amihez használtam a gráfkonvergencia ekvivalens definícióit. Azonban ezt a társszerzőkkel közös eredményt nem vettem be a disszertációba, mivel a gráfkonvergencia ekvivalenciák rutin alkalmazásáról volt csak szó. Ugyanakkor Lovász László és társszerzőinek három cikkére hivatkoztam a disszertációban, a T. Sós Verával közös Általánosított kvázirandom gráfok cikk eredményeit pedig intenzíven használtam az általánosított kvázirandom tulajdonságok ekvivalenciájának bizonyításánál.

  A [2] arXiv cikket azonban nem ismertem és köszönöm a bírálónak, hogy felhívta rá a figyelmemet. Az ott definiált 'jumble norm' számomra hasznos lehet a többrészes diszkrepancia becslésekben, hiszen ott normálok a csúcs-részhalmazok térfogatával, hasonló normálás a jumble-normánál is van. A cikk szerzői az elnevezésben a Thomason által 1987-ben bevezetett 'jumbled graphs' fogalomra hivatkoznak. Az eredeti definíció szerint egy gráf (p, α)-jumbled ha bármely indukált H részgráfjában $|e(H) - p\binom{|H|}{2}| \leq \alpha|H|$. Ha p az élsűrűség, akkor kellően kis α az egyrészes diszkrepanciára emlékeztet. Thomason páros gráfokra általánosított diszkrepancia fogalmát használom is az ekvivalenciák bizonyításánál. Fontos, hogy az élsűrűség nálam nem szerepel a diszkrepancia definíciójában, de az arra tett feltételek mellett be tudom vinni a becslésekbe, akárcsak a fokszám feltételek mellett a részhalmazok számossága helyett a térfogatot. Szintén az ekvivalenciák bizonyításánál intenzíven használom a vágás-normát, de meggondolom, nem lehetne-e azt a jumble normával helyettesíteni.

- Az [1] arXiv cikk 2017-ben jelent meg (a disszertációmat 2016 júliusában adtam be), azonban az ennek eredményeit nagyrészt tartalmazó cikk utolért engem (2016 őszén megkaptam AMS reviewra):

  E. Abbe, C. Sandon: Community detection in general stochastic block models: fundamental limits and efficient algorithms for recovery, 2015 IEEE 56th Annual Symposium on FOCS.

  A cikk a sztochasztikus blokkmodellt tárgyalja adott k osztály esetén. Fő eredmény, hogy amenyiben egy gráf ebből a modellből jön, akkor algoritmusuk nagy valószínűséggel felismeri az osztályokat. Amit ők exact ill. almost exact recovery-nek neveznek, az valójában a becslés erős ill. gyenge konzisztenciája, amit valószínűségszámítási eszközökkel lehet bizonyítani.

  Míg én a $k \times k$-as $\mathbf{P}$ él-valószínűségmatrixot n növekedésével konstans módon tartom, addig ők ezt skálázzák a csúcsok n számával. Pl. a $\mathbf{P} = \mathbf{Q}/n$ vagy $\mathbf{P} = \ln n\, \mathbf{Q}/n$ eseteket vizsgálják, ahol a $\mathbf{Q}$ rögzített $k \times k$-as, nem-negatív elemű, szimetrikus mátrix.

- A [4] cikk éppen az ún. szimmetrikus blokkmodellt tárgyalja, ahol az osztályokon belüli él-valószínűségek mind p , az osztályok közöttiek pedig q. Ha $p = q$, akkor nyilván nem különíthetők el az osztályok: egy-osztályos eset, Erdős–Rényi gráf. A $k = 2$, $p = \frac{a}{n}$ és $q = \frac{b}{n}$ esetben a [4] cikk bebizonyítja, hogy az osztályok elkülönítése lehetetlen az $(a-b)^2 < 2(a+b)$ esetben és lehetséges, ha $(a-b)^2 > C(a+b)$ elég nagy C-vel. Ezekre az eredményekre [1] cikk is hivatkozik, de nem tesz hozzájuk lényegesen többet. Ezzel kapcsolatos még az 1966-os Kesten–Stigum küszöb, mely statisztikus fizikai szempontból fákra és többdimenziós Galton–Watson folyamatokra lett felállítva; ugyancsak vannak információelméleti vonatkozások is, és [1] hivatkozik Csiszár Imre 1963-as cikkére. Ezeket az eszközöket én nem használtam.

Szintén foglalkozott a skálázott blokkos modellel Coja-Oghlan 'planted partition model' és Bollobás–Janson–Riordan 'inhomgeneous random graps' címen, ezek a cikkek közelebb állnak az én megközelítésemhez (előbbi foglalkozott a spektrummal is), így ezekre a cikkekre hivatkoztam Wiley-nál megjelent könyvemben, de közvetlenül nem használtam eredményeiket a disszertációban, és ezt a bíráló sem hiányolta. A spektrális klaszterezés konzisztenciájáról szintén írt Lei és Rinaldo (Ann. Stat, 2015). Náluk a konzisztencia azt jelenti, hogy amennyiben gráfuk a blokkmodellből jön (adott k osztály mellett), akkor a (spektrális) klaszterezéssel kapott hibás osztálybasorolások száma osztva n-el 0-hoz tart, ha $n \to \infty$ (1 valószínűséggel). Ehhez a (néha közelítő) k-közép algoritmus célfüggvényének becslésére van szükség, amire általában még sok más paramétert is bevezetnek. Valójában ilyen majdnem biztos állításokat fogalmaz meg a disszertációmban az általánosított véletlen gráfokra kimondott Proposition 18, és látok arra esélyt, hogy az átfogalmazható a konzisztenciára.

A disszertációmban általánosított véletlen gráfnak nevezett objektum spektrális és diszkrepancia tulajdonságaira én akkor mondok ki tételeket, ha a $\mathbf{P}$ valószínűségmátrix fix és nem skálázódik n-el. Ezek az állítások valószínűségszámítási eszközökkel viszonylag könnyen kijönnek. Nehezebb volt az általánosított kvázirandom gráfokra bizonyítani spektrális és diszkrepancia tulajdonságokat, valószínűségszámítási eszközök nélkül. Azonban tapasztaltam, hogy egyirányú implikációk (pl. spektrális rés a k-adik sajátértéknél és a sajátaltérnél kis k-részes diszkrepanciát okoz) a sztochasztikus modelltől függetlenül is bizonyítható, és nagyméretű gráfokra az erősebb gráfkonvergencia és fokszámtulajdonságok nélkül is teljesül. Ezzel a diszkrepancia minimalizáló spektrális klaszterezést szerettem volna megalapozni, pusztán az adott gráfot, mint inputot figyelembe véve.

Szintén megjegyzem, hogy az [1], [4] cikkek a spektrummal keveset foglalkoznak, [1]-ben a 2-osztályos esetre tárgyalják újra az ismert Fiedler-vektor szerepét és idézik az általam is használt Weyl és Davis–Kahan típusú perturbációs tételeket.

Összefoglalva, Backhausz Ágnes javaslatára tanulmányozni fogom a jumble normát és annak alkalmazhatóságát a többrészes kvázirandom tulajdonságok ekvivalenciájának bizonyításánál. Szintén tervezem a különböző feltételekkel skálázott blokkmodellek vizsgálatát konzisztencia szempontjából, illetve az általánosított véletlen gráfokra kimondott tulajdonságok átfogalmazását a konzisztencia és rekonstruálhatóság nyelvén.

Budapest, 2017. június 8.

Bolla Marianna