Gauss-Markov véletlen mezők és a Sidorenko sejtés

Felsmann Dániel

A Gauss-Markov véletlen mezők a matematika több területén, például a statisztikában, a számítástudományban és a statisztikus fizikában is alapvetően fontos objektumok. A tulajdonságaik szorosan kapcsolódnak egy, a mezőt jellemző véges gráfhoz, azonban általában nem gráfelméleti megközelítésben szokás tárgyalni a tulajdonságaikat. A kutatómunkám során Csikvári Péter és Szegedy Balázs a témával foglalkozó cikkét ismertem meg, amelyben a szerzők a Gauss-Markov véletlen mezőket jellemző gráf vizsgálatával kapcsolják össze a területet a gráfelmélettel. A cikk egyik fontos eredménye például egy Sidorenko sejtéséhez hasonló determináns egyenlőtlenség.

Egy p dimenziós $X = (X_1, ..., X_p)$ valószínűségi változó többdimenziós normális eloszlású, ha $\sum_{i=1}^{p} \alpha_i X_i$ normális eloszlású tetszőleges $\alpha_i \in \mathbb{R}$ mellett. A többdimenziós normális eloszlást a μ várható értéke és Σ kovariancia mátrixa egyértelműen meghatározzák (a Σ mátrix elemei a megfelelő koordináták kovarianciái: $(\Sigma)_{ij} = \mathrm{cov}(X_i, X_j)$). Fontos kiemelni, hogy az így kapott mátrix mindig pozitív szemidefinit. Esetünkben egy adott $G = (V, E)$ véges gráf minden csúcsához egy-egy koordináta tartozik: $X = (X_v)_{v \in V}$, az X valószínűségi változó koordinátái pedig együttesen egy $|V|$ dimenziós normális eloszlást alkotnak.
Az $(X_v)_{v \in V}$ valószínűségi változó térbeli Markov tulajdonságú, ha $\forall S \subset V$ esetén teljesül, hogy az $(X_v)_{v \in S}$ marginális eloszlása feltéve az $(X_w)_{w \in V \setminus S}$ változókat megegyezik $(X_v)_{v \in S}$ eloszlásával, ha csak az $(X_w)_{w \in N_G(S)}$ változókat tesszük fel, ahol $N_G(S)$ azokat az S-en kívüli csúcsokat jelenti, amelyekbe megy él S-ből (S szomszédai). Ez a tulajdonság tehát azt fejezi ki, hogy $(X_v)_{v \in S}$ feltételesen független az S-el nem szomszédos csúcsokhoz tartozó változóktól. Ha a G gráf például egy út, akkor ez a definíció a klasszikus Markov-tulajdonsággal ekvivalens.

1. Definíció: Adott $G = (V, E)$ gráf esetén az $(X_v)_{v \in V}$ Gauss-Markov véletlen mező, ha többdimenziós normális eloszlású és rendelkezik a térbeli Markov tulajdonsággal.

A cikkben vizsgált esetben a gráf csúcsaihoz tartozó X_v valószínűségi változók mindegyike standard normális eloszlású, így az eloszlást a Σ kovariancia mátrix is meghatározza. További feltétel, hogy a mező homogén. Ez azt jelenti, hogy a gráf éleinek megfelelő (X_i, X_j), $(i, j) \in E(G)$ marginálisok eloszlása minden él esetén azonos. Ha Σ pozitív definit, akkor az eloszlás abszolút folytonos a Lebesgue mértékre nézve és sűrűségfvüggvénye $\mu = 0$ esetben:

$$f(\underline{x}) = \frac{1}{(2\pi)^{n/2} \det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2} \underline{x}^T \Sigma^{-1} \underline{x}\right)$$

A Gauss-Markov véletlen mezők egyes tulajdonságai megfeleltethetők a Σ kovariancia mátrix algebrai tulajdonságainak. A mező Markov-tulajdonsága például ekvivalens azzal, hogy $\det(\Sigma)$ maximális, illetve azzal is, hogy Σ^{-1} miden olyan eleme 0, ahol G-ben nincs él. Az f sűrűségfüggvényű X valószínűségi változó differenciál entrópiája

$$-\int f(\underline{x}) \ln f(\underline{x}) dx = \frac{n}{2} + \frac{n}{2} \ln(2\pi) + \ln \det(\Sigma)$$

amiből látszik, hogy a G gráfon az entrópiát maximalizáló normális elszlású változó Markov tulajdonságú.
A Markov tulajdonsághoz jutunk akkor is, ha a Σ kovariancia mátrix maximum likelihood becslését keressük a következő módon: legyen $X_1, X_2, ..., X_m$ egy $|V|$ dinemziós, 0 várható értékű normális eloszlásból vett minta. Tegyük fel, hogy az $S_G = \frac{1}{m} X_k X_k^T$ tapasztalati kovariancia mátrixot csak a gráf élei mentén tudjuk mérni. Erre az esetre vonatkozik Dempster alábbi tétele.

2. Tétel: Az előbbi esetben pontosan akkor létezik a Σ kovarianci mátrix maximum likelihood becslése, ha a csak részben kitöltött S_G mátrix kiegészíthető pozitív definitté. Ebben az esetben a $\hat{\Sigma}$ becslés az az egyértelmű mátrix, amelyre $(\hat{\Sigma}^{-1})_{ij} = 0$, ha $(i,j) \notin E(G)$.

A determinánst maximalizáló mátrix kereséséhez vezessük be a következő jelöléseket:

3. Definíció Adott $G = (V, E)$ gráf és $x \in (-1, 1)$ mellett jelölje $\mathcal{A}(G, x)$ az olyan $V \times V$ méretű M mátrixok halmazát, ahol

- M pozitív definit
- a főátlóban M minden eleme 1
- $(M)_{ij} = x$, ha $(i,j) \in E(G)$

Legyen $\mathbf{A}_G(x) \in \mathcal{A}(G, x)$ a determinánst maximalizáló mátrix és $\tau(G, x) = \det(\mathbf{A}_G(x))$.
Ha $\mathcal{A}(G, x)$ nem üres, akkor $\mathbf{A}_G(x)$ mindig egyértelműen létezik, mivel $\mathcal{A}(G, x)$ korlátos konvex halmaz és a határán nem lehet a maximum, illetve $M \mapsto \log\det(M)$ szigorúan konkáv. A cikk egyik fő célkitűzése a $\tau(G, x)$ függvény tulajdonságainak minél jobb megértése. Sajnos $\tau(G, x)$ értékére nem ismert általános képlet, azonban speciális gráfokra (például erősen reguláris, merevhúrú, teljes páros) explicit formulával kiszámítható. $\tau(G, x)$ hatványsorba fejthető, analitikus függvénye x-nek, ahol 0 körüli hatványsor egész együtthatós. Ennek, illetve az $\mathbf{A}_G(x)$ mátrixnak a közelítésére a szerzők be is mutatnak egy iteratív algoritmust.

A $\tau(G, x)$ függvény logaritmusa kifejezhető az extremális kombinatorikából ismert részgráf sűrűségek határértékeinek segítségével. A G gráf sűrűségét H-ban a $t(G, H)$ mennyiség méri, ez annak a valószínűsége, hogy egy véletlen $V(G) \to V(H)$ leképezés homomorfizmus. Sidorenko (illetve Erdős és Simonovits) sejtése azt mondja ki, hogy ha G páros gráf, akkor

$$t(G, H) \geq t(K_2, H)^{|E(G)|}$$

minden H nem üres gráf esetén.
Általánosabb megfogalmazásban azt is mondhatjuk, hogy a $G_1, G_2, ..., G_n$ gráfok teljesítenek egy multiplikatív egyenlőtlenséget az $\alpha_1, \alpha_2, ..., \alpha_n$ konstansokkal, ha

$$\prod_{i=1}^{n} t(G_i, H)^{\alpha_i} \geq 1$$

minden H nem üres gráf mellett. A $t(G, H)$ sűrűség általánosítható arra az esetre is, amikor G "sűrűségét" nem egy H gráfban, hanem egy W graphonban vizsgáljuk (a Sidorenko sejtés eredeti megfogalmazása is erre az alakra vonatkozik). A $\tau(G, x)$ függvény és $t(G, W)$ mennyiségek kapcsolata kifejezhető egy magas dimenziós gömbök segítségével definiált véletlen mátrix modellel, egy ez alapján felállított nagyeltérés tétel, valamint egy speciális W "gömbi" graphon segítségével. Ebből következnek a következő állítás is:

4. Tétel: Ha $G_1, ..., G_n$ teljesítenek egy multiplikatív egyenlőtlenséget $\alpha_1, ..., \alpha_n$ konstansokkal, akkor $\prod_{i=1}^{n} \tau(G_i, x)^{\alpha_i} \geq 1$ is teljesül $\forall x \in (-1, 1)$

Az előző állításnak Galvin és Tetali megfelelő tétele miatt közvetlen következménye a közetkező tétel:

5. Tétel: Ha $G = (V, E)$ d-reguláris páros gráf, akkor

$$\tau(G, x)^{1/|V(G)|} \leq \tau(K_{d,d}, x)^{1/2d}$$

A cikk egyik legfontosabb eredménye a következő, Sidorenko sejtéséhez hasonló állítás:

6. Tétel: Ha $x \in [0, 1)$, akkor tetszőleges G gráfra

$$\tau(G, x) \geq \tau(K_2, x)^{|E(G)|} = (1 - x^2)^{|E(G)|}$$

Belátható, hogy ha G páros gráf, akkor a $\tau(G, x)$ függvény x-ben páros, tehát ilyenkor $\tau(G, x) = \tau(G, -x)$. Ebben az esetben az előző tétel speciális eseteként kapjuk a Sidorenko sejtésének megfelelő állítás determinánsokra vonatkozó alakját:

7. Következmény: Ha G páros, akkor $\forall x \in (-1, 1)$

$$\tau(G, x) \geq (1 - x^2)^{|E(G)|}$$

Ezeket az eredményeket homogén Gauss-Markov mezőkre vonatkozó entrópia egyenlőtlenségként is értelmezhetjük. Kis számolás után látszik, hogy az $X = (X_v)_{v \in V}$ mező entrópiája $\mathbb{D}(X) = \frac{|V|}{2} \ln(2\pi e) + \frac{1}{2} \ln \tau(G, x)$. A csúcsoknak megfelelő koordinátákra $\mathbb{D}(X_v) = \frac{1}{2} \ln(2\pi e)$, míg az élek esetén $\mathbb{D}(X_i, X_j) = \ln(2\pi e) + \frac{1}{2} \ln \tau(K_2, x)$. Ezek segítségével az 6. tétel a következő entrópia egyenlőtlenségnek felel meg:

$$\mathbb{D}\left((X_v)_{v \in V}\right) - \sum_{(i,j) \in E} \mathbb{D}(X_i, X_j) + \sum_{v \in V} (\deg(v) - 1) \mathbb{D}(X_v) \geq 0$$

tehát a homogén Gauss-Markov véletlen mező differenciál entrópiája legalább akkora, mint az él entrópia $|E(G)|$-szeresének az és a csúcs entrópia $(|V(G)| - 2|E(G)|)$-szeresének a különbsége.

A 6. tétel egyenlőtlensége után természetesen merül fel a kérdés, hogy a kapott korlát mennyire éles. Ezzel kapcsolatban az derül ki, hogy a $\frac{\tau(G,x)}{(1-x^2)^{|EG)|}}$ függvény monoton nő az $\left[0, \frac{1}{\Delta - 1}\right]$ és $\left[\frac{1}{\bar{d}-1}, 1\right]$ intervallumokon, ahol Δ a gráf legnagyobb fokszáma, $\bar{d}$ pedig az átlagos fokszám. Ennek a vizsgálata a $\tau(G, x)$ függvény logaritmikus deriváltjával történik, amely szép alakban kifejezhető $\mathbf{A}_G(x)$ inverzének elemeivel.

A következő állítás szerint $\tau(G, x)$ viselkedése a $\left[0, \frac{1}{\Delta - 1}\right)$ és $\left(\frac{1}{\bar{d}-1}, 1\right]$ intervallumokon eltér egymástól:

8. Tétel: a) Ha $x \in \left[0, \frac{1}{\Delta-1}\right)$, és G legrövidebb körének hossza g, akkor

$$\left|\frac{\ln \tau(G,x)}{|E(G)|} - \ln(1-x^2)\right| \leq 2\frac{((\Delta-1)x)^g}{1-(\Delta-1)x}$$

b) Ha pedig $x \in \left[\frac{1}{\bar{d}-1}, 1\right]$, akkor létezik olyan G-től független, pozitív $\alpha(\bar{d}, x)$ függvény, amellyel

$$\left|\frac{\ln \tau(G,x)}{|E(G)|} - \ln(1-x^2)\right| \geq \alpha(\bar{d}, x)$$

A tételből látszik, hogy d reguláris gráfok esetén a $\tau(G,x)$ függvény viselkedésében egy fázisátmenet figyelhető meg a $\frac{1}{d-1}$ pontban. Az előző tétel b) részének következménye a következő állítás.

9. Következmény: Ha (G_n) olyan d reguláris gráfokból álló sorozat, amelyben a legrövidebb körök hossza $g(G_n) \to \infty$, akkor

$$\lim_{n\to\infty} \frac{\ln \tau(G_n, x)}{|V(G_n)|} = \frac{d}{2}\ln(1-x^2)$$

Hivatkozások

- Csikvári Péter és Szegedy Balázs: On Sidorenko's Conjecture for Determinants and Gaussian Markov Random Fields (még nem jelent meg)