Oksági Bayes hálók

P. Antal

2010. augusztus 29.

Tartalomjegyzék

1. Bevezetés

A fogalomalkotás (kategorizáció) mellett az oksági kapcsolatok felfedezése számít a megismerési folyamatok másik alapkövének. Fontosságát az mutatja, hogy míg például egy predikciós modell egy adott kontextusban érvényes és passzív megfigyelések alapján előrejelzésre szolgálhat, addig egy oksági kapcsolat (mechanizmus, törvényszerűség) egyrészt autonóm, azaz más kontextusban is előfordulhat mint stabil, moduláris elem, másrészt beavatkozások modellezésére is szolgálhat.

Ha feltesszük, hogy adottak passzív megfigyelésekhez tartozó minták $D_N = \{(X^{(1)}, \ldots, X^{(N)})\}$, akkor érdemes formálisan tovább bontani ezeket az indukciós célokat.

1. (értékegyüttesek tanulása) Gyakori mintázatokat (értékkonfigurációkat) keresünk.

2. (logikai kifejezések tanulása) Közelítőleg érvényes logikai formulákat tanulunk.

3. (statisztika következtetés) Megfigyelési predikciót - azaz $p(Y|X)$ jellegű következtetések elvégzését - lehetővé tevő modelleket tanulunk.

4. (oksági következtetés) Beavatkozások hatásának a predikcióját - azaz $p(Y|do(X))$ jellegű következtetések elvégzését - lehetővé tevő modelleket tanulunk.

5. (kontrafaktuális következtetés) A realizáltól („faktuálistól") eltérő („kontrafaktuális") - azaz elképzelt - beavatkozások hatásának a predikcióját - azaz $p(Y|X = x, do(X = x'))$ jellegű következtetések elvégzését - lehetővé tevő modelleket tanulunk.

Egy kényelmi feltevés még a jelen áttekintésben, hogy a változóink binárisak. Az előző hierarchiát részben felhasználva fontoljuk meg a különbőzőségeket.

1. (értékegyüttesek) Gyakori mintázatok, például együttelőfordulások felfedezése (azaz például annak felfedezése, hogy ha X előfordul, akkor többnyire Y is).

2. (statisztikailag szignifikáns asszociációk) Valószínűségi értelemben függő változók X, Y felfedezése adott megbízhatóság (konfidencia) és erősség (pontosság) mellett.

3. (oksági viszony) Az X, Y változók közötti statisztikai asszociáció lehet a kizárólagos következménye egy közös oknak $*$ mindennemű oksági kapcsolat nélkül (például direkt függésekkel $X \leftarrow * \rightarrow Y$). Cél ekkor olyan asszociák megtalálása, amelyek nem kizárólag rejtett változóknak a következménye („pure confounding"). Az asszociáció és oksági kapcsolat különbőzőségére egy egyszerű példa, ha ismert az $X \rightarrow Y$ oksági viszonya és $p(X, Y)$ valószínűségi kapcsolata például a generatív irányhoz tartozó $p(Y|X)$, $p(X)$ megadásával, akkor tudjuk azt, hogy $p(Y|do(X)) = p(Y|X)$ és $p(X|do(Y)) = p(X)$. Továbbá, hogy ha a (tetszőleges erősségű) $X - Y$ asszociációt egy közös ok $(*)$ okozza $(X \leftarrow * \rightarrow Y)$ — azaz tiszta zavarásról van szó („pure confounding") — , akkor a beavatkozásnak a megfigyelt asszociáció ellenére nincs hatása: $p(X|do(Y)) = p(X)$ és $p(Y|do(X)) = p(Y)$.

4. (direktség) Ha már abban bizonyosak is vagyunk, hogy X, Y változók között létezik oksági kapcsolat, adott esetben fontos lehet azt is tudni, hogy az egy aktuális változóhalmazhoz képest „direkt"-e, azaz más változók által közvetített vagy nem (máshogy fogalmazva a többiek ismerete ezt a kapcsolatot blokkolja vagy nem).

Az oksági következtetések modern formalizálásában úttörőt szerepet játszó J.Pearl-től származó hasonlat szerint az oksági és valószínűségi modellek kapcsolata hasonló a térbeli tárgyak és adott síkbeli ábrázolásuk kapcsolatához [**?**].

Az oksági kapcsolatok fogalmának tisztázása központi elem a filozófiában is, aminek kapcsán néhány idézetet és művet említünk meg.

1. Democritus: „I would rather discover a single cause than become king of the Persians."

2. Aristotle: „the material cause, the formal cause, the efficient cause, and the final cause"

3. D.Hume: „...there is a NECESSARY CONNECTION to be taken into consideration"

4. H.Reichenbach's: „Common Cause Principle: a correlation between events A and B indicates either that A causes B, or that B causes A, or that A and B have a common cause".

1.1. Oksági kritériumok

Az oksági kapcsolatok felismerésére több tudományágban is kritériumokat fogalmaztak meg. Az epidemiológiai területéről származok például a következő feltétel lista [**?**]:

1. erős statisztikai asszociáció,

2. X időben előzze meg Y-t,

3. plauzibilis magyarázat létezzen, úgy hogy nel legyenek alternatív, zavaró tényezőre is építő alternatív magyarázatok,

4. szükségesség (az ok megszüntetésével a hatás is szünjön meg),

5. elégségesség (az ok bekövetkeztével a hatás is erősödjön).

Egy másik Max Born-tól származó feltétel lista a következő:

1. Y előfordulása törvényszerű kapcsolatban függjön X előállásától,
2. az oknak időben megelőzőnek, vagy egyidejűnek kell lennie,
3. térbeli vagy fizikai törvényszerűségeken keresztüli szomszédosság

1.2. Megfigyelés, beavatkozás, spekuláció

A korábban informálisan használt mennyiségek definíciói a következők.

1.1. Definition. *Jelölje* $p(Y = y|X = x)$ *a (megszokott megfigyelési) feltételes valószínűségét* $Y = y$ *értéknek* $X = x$ *érték megfigyelése esetén.*

1.2. Definition. *Jelölje* $do(x)$ *az* X *változó* x *értékre történő beállításához tartozó beavatkozást és* $p(Y|do(x)$ *az ehhez tartozó beavatkozási eloszlást.*

1.3. Definition. *Jelölje* $p(Y = y|do(X = x), Y = y', X = x')$ *a kontrafaktuális valószínűségét annak, hogy* $Y = y$ *amikor* $X = x', Y = y'$ *és* $do(X = x)$.

Jelölje továbbá $(X \perp\!\!\!\perp Y|Z)_p$ X feltételes függetlenségét Y-tól Z feltétellel a p valószínűségi eloszlásban.

1.3. Oksági kapcsolat erősségének mérőszámai

Felhasználva a *do* szemantikát a statisztikai és oksági kapcsolatak kutatásának a legfőbb fogyasztójának számító epidemiológiában a következő mérőszámok használtak (egy bináris X (pl. kitettség) és Y (pl. betegség) között): rizikóbeli különbség vagy oksági hatás (δ), tulajdonítható/okozott rizikó (θ) és a már logisztikus regressziónál is megismert esélyhányados(Ψ).

$$\delta = p(y|do(x)) - p(y|do(\neg x) \tag{1}$$

$$\theta = \frac{p(y|do(x)) - p(y|do(\neg x)}{p(y|do(x))} \tag{2}$$

$$\Psi = \frac{p(y|do(x))/p(\neg y|do(x))}{p(y|do(\neg x))/p(\neg y|do(\neg x))} \tag{3}$$

Természetesen a standard epidemiológiai definíciója ezeknek nem a beavatkozásos do szemantikát használja, hanem az "adjustált" megfigyelési valószínűségéken alapulókat. Az „adjustálás" (vagy "kontrollolálás") a zavaró tényezők $\underline{Z}$ eliminálására szolgál úgy, hogy X hatását Y-ra a potenciális zavaró tényezők azonos értékei mellett vizsgáljuk (azaz feltételbe emeljük őket és "fixen tartjuk" őket.

$$p(y|do(x)) = \sum_z p(y|x, z)p(z) \tag{4}$$

1.4. Valószínűségi és oksági „paradoxonok”

Az asszociációk és oksági kapcsolatok felderítésénél jó ha tudatában vagyunk következő lehetőségeknek, amelyek a mindennapi logika szerint furcsának tűnhetnek.

1.4.1. Tisztán magasabbrendű függések

Tegyük fel, hogy X, Y, Z bináris változók, ahol X, Y függetlenek, egyenletes eloszlásúak, és Z pedig egy $Z = XOR(X, Y)$ logikai függvénnyel meghatározott. Ekkor $(X \perp\!\!\!\perp Z)$ és $(Y \perp\!\!\!\perp Z)$, de $(\{X, Y\} \not\perp\!\!\!\perp Z)$, azaz az együttesüktől már függ. Azaz a függések (asszociák) nem feltétlenül monotonak.

1.4.2. Függések iránya

Egy Markov lánc $\mathcal{X} = \{X_0, X_1, \ldots X_n\}$, amely a $p(X_i|X_{i-1})$ átmenet valószínűségekkel adott, átírható egy $p(X_{i-1}|X_i)$ alakba is, azaz lánc irányítása esetleges.

1.4.3. Intranzitív függések

Tegyük fel, hogy X, Y, Z közül legalább egy nem bináris változó, pl Y. Ekkor létezik egy olyan eloszlás $p(X, Y, Z)$, hogy $(X \not\perp\!\!\!\perp Y)$ és $(Y \not\perp\!\!\!\perp Z)$, de $(X \perp\!\!\!\perp Z)$, azaz a függések (asszociációk) nem feltétlenül tranzítívak. Azonban ha feltesszük, hogy az eloszlásunk függetlenségi viszonyai stabilak, azaz infinitizimális perturbációkra a kvalitív függési viszonyok nem változnak, akkor az intranzitív hármas függési modellt egyedül az úgynevezett v-struktúra magyarázhatja $(X \to Y \leftarrow Z)$, amelyben $(X \not\perp\!\!\!\perp Z|Y)$ (és nem $(X \perp\!\!\!\perp Z|Y)$). Azaz ekkor kivételesen adódik, hogy Y a következménye két őt befolyásoló független eseménynek, amelyek tapasztalataink szerint mindig korábbiak (innen ered a filozófiabeli „idő nyila” megnevezés).

1.4.4. Simpson paradoxona

Lehetséges olyan eloszlás, hogy $p(Y|X) > p(Y|\bar{X})$, de $p(Y|X, Z) < p(Y|\bar{X}, Z)$ and $p(Y|X, \bar{Z}) < p(Y|\bar{X}, \bar{Z})$, azaz az asszociáció hatása rétegenként (Z szerint) és összesítve ellenkező(!) lehet.

1.5. Okozati relációk tanulása oksági modellek felhasználásával

Amint a 1. Lista mutatja az adott absztrakciós szinthez, avagy változó halmazhoz tartozó direkt oksági kapcsolatok tanulása mind a rejtett változók, mind az asszociációs kapcsolatok tipikusan tranzitív jellege, és a többváltozós kapcsolatok miatt is kihívásokkal teljes. Annak megértéséhez, hogy milyen feltevések mellett is tudunk több változó együttes elemzése során oksági relációkat tanulni, szükségünk van a következő fogalmakra.

1.5.1. Valószínűségi Bayes hálók

1.4. Definition. *Egy eloszlás $p(X_1, \ldots, X_n)$ eleget tesz a globális Markov feltételnek G irányított, körmentes gráf viszonylatában, ha*

$$\forall\, X, Y, Z \subseteq V:\ (X \perp\!\!\!\perp Y|Z)_G \Rightarrow (X \perp\!\!\!\perp Y|Z)_p, \tag{5}$$

ahol $(X \perp\!\!\!\perp Y|Z)_G$ jelöli, hogy X és Y d-elválasztottak Z által, azaz hogy minden út p X és Y között blokkolt Z által a következőképpen

1. *vagy a p út tartalmaz egy Z-beli n csomópontot nem összetartó élekkel (azaz így $\rightarrow n \rightarrow$ vagy így $\leftarrow n \rightarrow$),*
2. *vagy a p út tartalmaz egy nem Z-beli n csomópontot összetartó élekkel (azaz így $\rightarrow n \leftarrow$), amelynek nincs leszármazottja Z-ben.*

1.5. Definition. *Egy irányított, körmentes gráf (DAG) G a $p(V)$ eloszlás viszonylatában Bayes háló, ha a valószínűségi változókat G csomópontjai reprezentálják és (G, p) eleget tesz a globális Markov feltételnek úgy, hogy G minimális (azaz él már nem hagyható el a feltétel megsértése nélkül).*

1.6. Definition. *Egy p eloszlást stabilnak*[1] *vagy hűségesnek hívunk, ha létezik egy DAG ami pontosan reprezentálja a p eloszlás függetlenségi relációit (azaz úgynevezett tökéletes térkép: $(X \perp\!\!\!\perp Y|Z)_G \Leftrightarrow (X \perp\!\!\!\perp Y|Z)_p$, $\forall\ X, Y, Z \subseteq V$). Egy p eloszlás stabil a G DAG-ra nézve, ha G pontosan reprezentálja a függetlenségeit.*

1.7. Definition. *Két DAG G_1, G_2 megfigyelési ekvivalens, ha pontosan ugyanazokat a d-szeparációs relációkat definiálják, azaz $((X \perp\!\!\!\perp Y|Z)_{G_1}) \Leftrightarrow (X \perp\!\!\!\perp Y|Z)_{G_2}$.*

Vegyük észre, hogy az ekvivalencia osztályok tartalmazhatnak $n!$ számú DAG-ot (amelyek bármelyike azt reprezentálja, hogy nincs egyetlen függetlenségi viszony sem) vagy csak 1 darabot (ami a teljes függetlenséget reprezentálja).

[1] For a different interpretation of this term in probability theory, see [?].

1.1. Theorem. *Két DAG G_1, G_2 megfigyelési ekvivalens, acsa ha az irányítatlan vázuk megegyezik és ugyanazon v-struktúrákat tartalmazzák (azaz konvergáló éleket, amelyek talpánál nincs él).*

1.8. Definition. *Az esszenciális gráf a megfigyelési ekvivalens DAG-ok halmazát reprezentálja egy részlegesen irányított DAG-gal (PDAG), amiben csak azok az élek irányítottak (az úgynevezett kényszerített élek), amelyek az ekvivalenciaosztálybeli DAG-oknál azonosan irányítottak, a többi él irányítatlansága az eldönthetetlenséget jelzi.*

1.5.2. Valószínűségi Bayes hálók oksági értelmezése

Az eszenciális gráfokban szereplő kényszerített élek kapcsán természetes módon vetődik fel az oksági értelmezésük, például ilyen élek triviálisan a (nem bezavart) v-struktúrák $X \rightarrow Y \leftarrow Z$, amelyek irányítása megfelel az emberi tapasztalatoknak az intranzitív hármasokkal kapcsolatban (nevezetesen, hogy X és Z független, Y-t időben megelőző okok.

Ellenérvként a következő feltevések gyengeségei lehetnek.

1. A direkt függések nem mindegyike oksági (hanem például szemantikai).
2. Adott eloszlás esetén a hozzátartozó megfigyelési ekvivalencia osztály egyértelműségéhez fel kell tételezni a stabilitást (ami logikai függéseknél nem teljesül, ahogy a paradoxonok között bemutattuk).
3. Az eszenciális gráf definíciója csak a minimális, konzisztens modellekre támaszkodik (egyáltalán nem véve figyelembe a kissé komplexebb konzisztens modellek irányítását).

Ezek tárgyalása meghaladja az összefoglaló kereteit, de az első ponttal kapcsolatban megjegyezzük, hogy az úgynevezett Oksági Markov feltétel elfogadásával ez kiküszöbölhető, ami leegyszerűsítva az olyan rejtett változók létét zárja ki, amelyek több általunk megfigyelt változóra is hatnak.

1.5.3. Okozati relációk tanulása lokális modellekkel

Azonban potenciálisan létező rejtett változók esetén is van lehetőség oksági kapcsolatok felfedezésére, amint a következő példa mutatja.

Kiindulásként emlékezzünk arra, hogy két változó passzív megfigyelése és esetleges statisztikai asszociációja semmit nem mond az oksági kapcsolatukról.

1.1. Example. *Az Oksági Markov feltétel (azaz rejtett közös okok kizárása) garantálja, hogy nem több mint három változó passzív megfigyelésével oksági relációt következtethetünk ki. Direkt függés esetén X, Y és Y, Z között, amikor nincs direkt függés X, Z között és nincs feltételes függetlenség $(X \perp\!\!\!\perp Z|\{Y, S\})$ (azaz feltételes függés áll fenn), akkor ezt a d-szeparáció szemantikája szerint egyértelműen egy $X \rightarrow Y \leftarrow Z$ v-structúrával kell kifejezni.*

1.2. Example. *Ha potenciális zavaró tényezők nincsenek a priori kizárva, akkor még legalább egy változó megfigyelése szükséges, hogy kizárjuk, hogy egy direkt függést tisztán egy zavaró tényező okoz. Folytatva az előző példát tegyük fel, hogy megfigyelünk egy negyedik W változót Y, W direkt függéssel és feltételes függetlenséggel $(W \perp\!\!\!\perp \{X, Z\}|Y)$ (stabil eloszlást feltételezve W függ X-től és Z-től is). Mivel Y függetlenséget okoz, a d-szeparációs szemantika egy $Y \rightarrow W$ élet követel meg (figyelembe véve a korábbi v-structúrát), hiszen egy tiszta zavaró tényező hatását $* \; Y \leftarrow * \rightarrow W \; Y$ mint okozat nem tudná blokkolni.*

Ezek az elméleti eredmények úgynevezett lokálisa tanulási algoritmusokban hatékonyan használhatóak oksági viszonyok felfedezésére, azaz megkerülik a költséges teljes modell alapú, eszenciális gráfokon alapú oksági következtetést. Mindazonáltal a teljes modell alapú módszerek lehetővé teszik a Bayes statisztika használatát is [**? ? ? ?**]. Ekkor formálisan is kijelenthető, hogy $X \rightarrow Y$ oksági kapcsolatnak az éppen rendelkezésre álló adatok alapján az a posteriori valószínűsége mekkora.

$$p(\mathrm{KenyszeritettEl(X, Y)}|\mathrm{Adat}) \tag{6}$$

1.5.4. Ellenérvek

Az okozati relációk tanulása kapcsán a korábbiak mellé a következő, nem csak Bayes háló specifikus ellenérvek is felhozhatóak [**? ?**].

1. (Rejtett zavaró tényezők, stabilitás, modell minimalitás.)
2. Kiválasztási bias, amikor a megfigyelés akár független események kombinációjátók függ, így okozva az adatokban egy akauzális függést.
3. Oksági modellek keveréke, azaz ha mind X befolyásolja Y-t és fordítva. Hasonló probléma a visszacsatolás.
4. Globális fizikai és szemantikai kényszerek a változók között.

5. Az asszociációk, a függetlenségek, a függések, és így az induktívan kikövetkeztetett oksági viszonyok viszonylagosak az elemzett változók halmazához képest, sőt még az értékükhöz képest is.

Mindazonáltal a többváltozós elemzések miatt és a Bayes statisztikai megközelítés miatt az oksági következtetésnek ma már egy széles spektruma van, szemben a korábbi két változós „correlation (asszociation) $\neq$ causation" zsákutcás igaz/hamis szemlélettel.

Hivatkozások

[] J. M. Bernardo and A. F. M. Smith. *Bayesian Theory.* Wiley & Sons, Chichester, 2000.

[] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis.* Chapman & Hall, London, 1995.

[] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. *Markov Chain Monte Carlo in Practice.* Chapman & Hall, London, 1996.

[] C. Glymour and G. F. Cooper. *Computation, Causation, and Discovery.* AAAI Press, 1999.

[] J. Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, 2000.

[] A. Rényi. *Probability Theory.* Akadémiai Kiadó, Budapest, 1970.

[] C. P. Robert. *The Bayesian Choice.* Springer-Verlag, New York, 2001.

[] P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search.* MIT Press, 2001.

[] M. Woodward. *Epidemiology: Study design and data analysis.* Chapman&Hall, 1999.