グラフィカルモデリングの紹介

因果推論を中心に

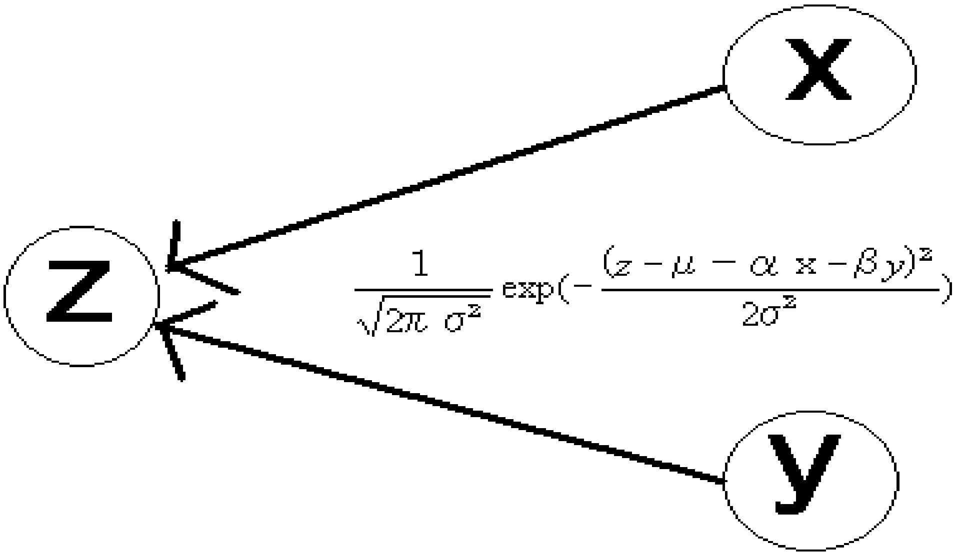
発表者　文学部研究科M2
近岡利昌

グラフィカルモデリングとは何か？

- 様々な確率変数の依存関係をグラフを用いてモデル化すること
- 確率変数を頂点、直接の依存関係を辺に対応させます

例えば簡単な回帰式の場合

$z = \mu + \alpha x + \beta y + e$

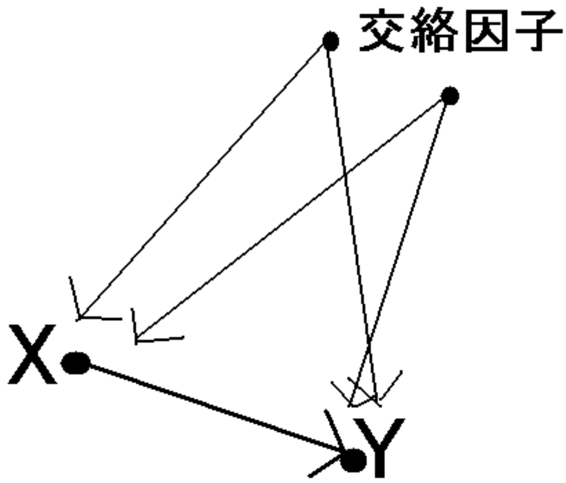


グラフィカルモデリングは何の役に立つか？

- 回帰分析、因子分析、SEM、信号検出理論、隠れマルコフモデル、パス解析等々、ほとんどの統計モデルを、このモデルのもとで統一的に表せます
- これら全部について触れるのは無理なので、今回はグラフィカルモデルを用いて<u>因果関係</u>について推論する方法を紹介します

因果関係をモデル化するには

- 通常XとYの間に見かけの因果関係があっても、<u>交絡因子</u>が存在する可能性があるので、因果関係について何か推論するためには介入実験が必要といわれています



グラフィカルモデリングを使うと

- グラフィカルモデリングを使うと変数間の因果関係、介入による変化、交絡因子などについて統一的に扱うことが出来ます。
- 変数間の因果関係を巡回無し有向グラフを使って表します。
- グラフと確率変数を対応付けるさいに重要なのが条件付独立という概念です。

条件付き確率

- P（A）： Aという事象が成り立つ確率
- P（A,B）： Aという事象とBという事象が同時に成り立つ確率、<u>同時確率</u>
- P（A｜B）： Bが成り立つときにAが成り立つ確率、<u>条件付き確率</u>
- *P（A,B)＝P（A／B)P(B)　　Bが成り立つときのAの条件付き確率にBの確率をかけるとAとBの同時確率に*

独立関係とは

- P（A｜B)＝P(A)となるとき、事象AとBは<u>独立</u>といい、A⊥Bと書きます
- AとBが独立であれば、

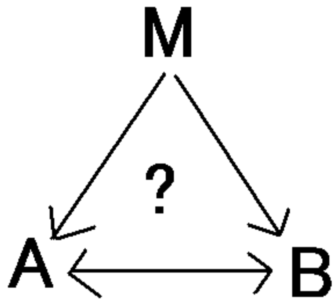
P（A,B)＝ P（A｜B)P(B)に上式を代入して、

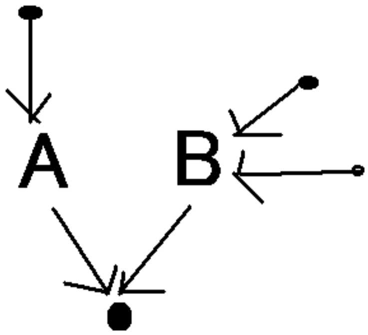
P(A,B) ＝P(A)P(B)

つまりBが成り立っても成り立っていなくても、Aが成り立つ確率は変わらないという意味です

グラフと独立関係の対応

- A⊥Bではない時、二つの変数の片方からもう片方に、もしくは別の因子から両方に因果関係が存在することが考えられます。(Reichenbach 1925)
- AとBの間の直接の矢印や、矢印を遡ってたどり着く共通の因子がないとき、A⊥B



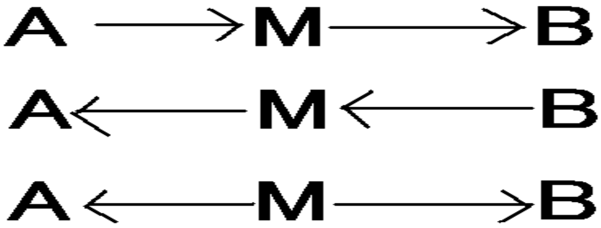


条件付き独立関係とは

- P（A｜M）＝P(A｜M,B）のとき、AとBはMのもと<u>条件付き独立</u>といい、A⊥B｜Mと書きます
- たとえば子どもの発育について研究しているとき、体重（A）と知能（B)に強い相関が認められたとします。しかし年齢（M)という因子をいれることでこの相関がほとんど消えてしまう場合など、AとBはMのもと<u>条件付き独立</u>といえます

グラフと条件付き独立の対応

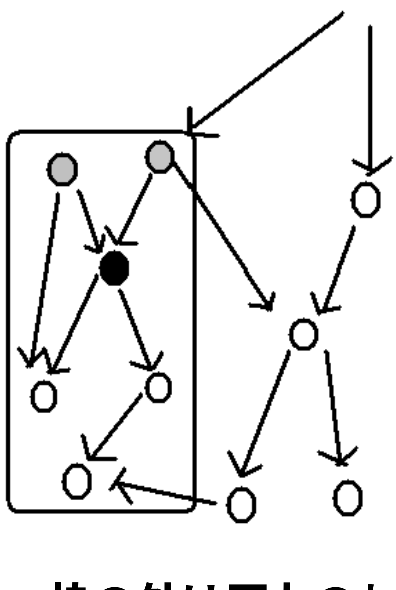
- A⊥B｜Mのとき、因果の向きは次の3つのどれかです



AB間のMを通らない因果の矢印や、M以外の共通原因の矢印はありません

因果マルコフ仮説

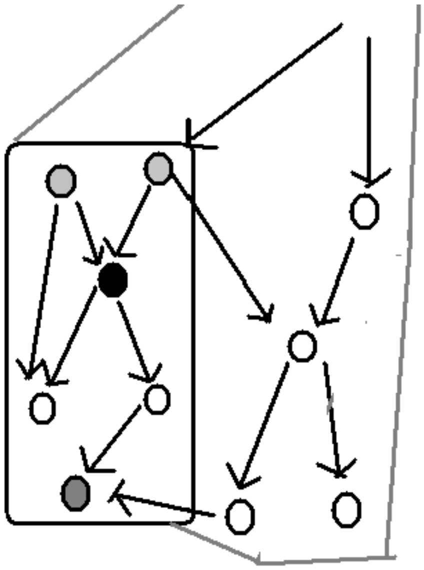
- これらの関係を一般化して言い表すと
- 「巡回無し有向因果グラフのいかなる変数も、直接の原因のもと、直接間接の結果を除いて、他の全ての変数と条件付独立である。」(因果マルコフ仮説) となります



枠の外は灰丸のもと黒丸と条件付独立

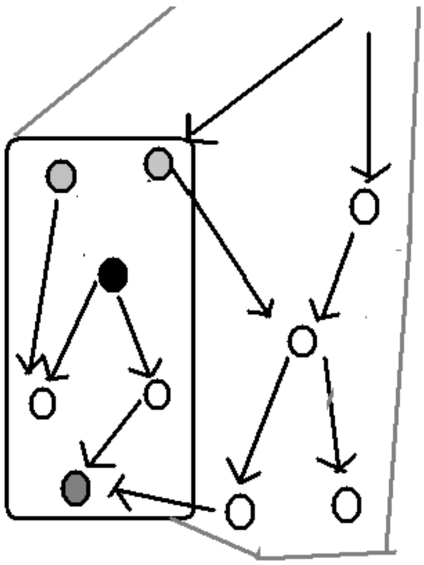
注意

- 条件付けする変数を増やすと、独立であった変数が非独立になってしまうことがあります。
- たとえば身長と外国語の点数の合計が250点台のものだけがクラスCに配属されるとします。身長と外国語の成績は独立かもしれませんが、もしXがクラスCに属するという条件がつくと、この二つは負の相関を示します



実験的介入をする場合

- 実験的介入を表すときはグラフを次のように書き換えます
- 例えば先のグラフの黒丸に介入を行うときは、黒丸の値を固定し、黒丸の原因と繋がる矢印を削除します

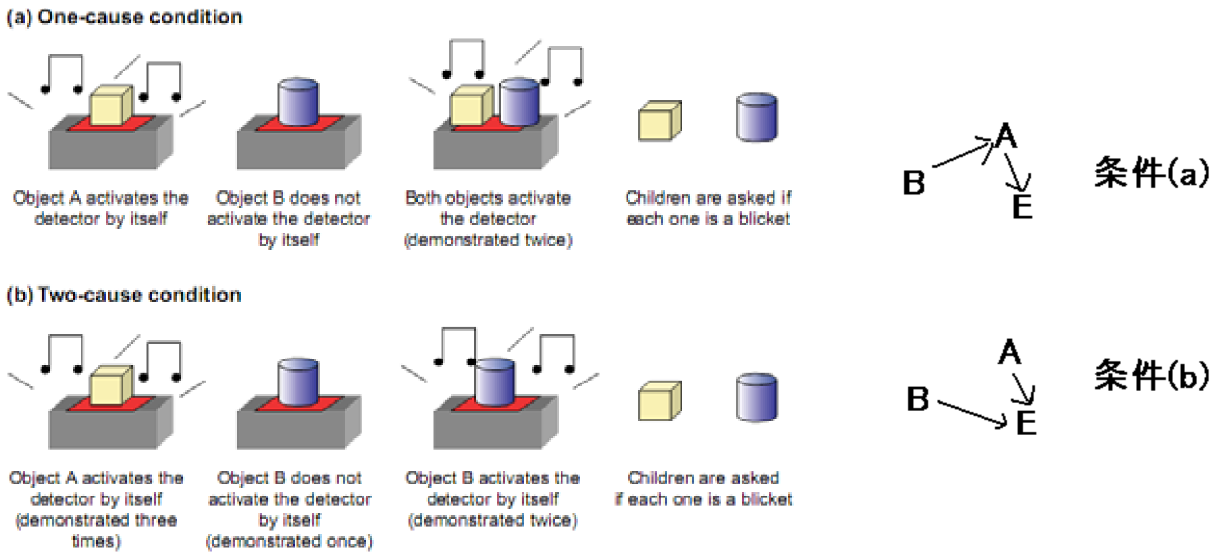


条件付確率と介入から因果グラフを推論する

- たとえばA⊥B｜Mとします。その場合考えられる因果構造は（1）A→M→B（2）A←M←B（3）A←M→Bの３つです。ここで変数間の独立・非独立は、変数間の因果関係の結果であると仮定します。
- Mに介入を行うとそれぞれのグラフが（1‘）A, M→B（2’）A←M, B（3’）A←M→Bとなります
- A⊥M,Bとなれば（1）、B⊥A,Mとなれば（2）、変数が独立にならなければ（3）が正しい因果構造として推論できます

研究事例紹介

- ２歳半の子どもの因果推論。(Gopnik et al., 2001)と４歳の子どもの因果推論（Sober et al., 2004)
- 子どもは「blicketが機械を動かすよ」と教示されます
- そして次の二条件を与えます



条件(a)では子どもは物体Aをblicketとして常に選ぶ
条件(b)では物体Aも物体Bも同じくらいにblicketとして選ぶ
どの条件でもそれぞれの物体と、機械の活動 (E)が共起する回数は同じ。しかし、条件(a)ではB⊥E｜AだがA⊥E｜Bではない。条件(b)ではB⊥E｜AでもA⊥E｜Bでもない。子どもの反応はグラフィカルモデルを使って記述できる

参考文献

- Gopnik, A. Schultz, L. (2004). Mechanisms of theory formation in young children. TRENDS in cognitive sciences vol.8, no.8, Aug, 371-377
- A, Gopnik. J, Tenenbaum (2007). Bayesian networks, Bayesian learning and cognitive development, Developmental Science (special section on Bayesian and BayesNet approaches to development) 10(3): 281-287