第1回人工知能学会DMSM
データマイニングと統計数理研究会(SIG-DMSM)
日程:2006/7/11(火) 於:統計数理研究所講堂

統計的因果推論と因果探索

大阪大学
狩野　裕(数理科学領域)
宮村　理(D3, 学振特別研究員)

目次

1. 序

無作為化実験

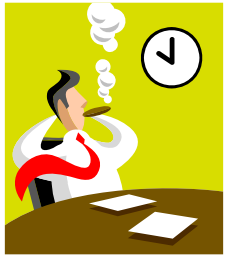
- 統計的因果推論の基本的道具
- 例「喫煙⇒肺がん」
  - ある動物を無作為に2群に分ける
    - 第1群には強制的に定期的に喫煙
    - 第2群には喫煙なし
    - 被験者(動物)が過去に背負ってきた履歴が確率的にバランス化
    - 環境要因をコントロール
  - 一定期間後に両群の発癌率を比較する
- キーワード
  - 無作為割付け(random assignment)

5

無作為化実験

- 動物での結果が直ちに人間へ適用できない
- 人間に対する実験は倫理的な問題
- 観察研究に頼らざるを得ない場合は少なくない

- だそく
  - 人間に対して無作為化実験可能である場合でも
    - 実験条件を遵守するとは限らない
      - ノンコンプライアンス(法令不遵守)
    - 不自然な状況での実験になる可能性



6

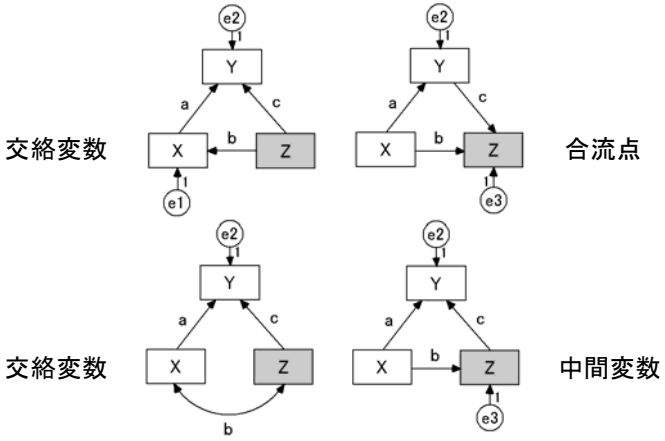
観察 (or 相関的) 研究

- 観察研究の方法
  - 喫煙者群と非喫煙者群に分ける
    - 受動的. 割付はしない
  - 両群を経時的に観察し発癌率を比較

- 観察研究には重大な問題点
  - (未観測)交絡変数 (confounding variable)
  - サンプルセレクション (sample selection)
  - 測定誤差 (measurement error)

7

第三変数とその役割



2. 因果探索

8

9

因果探索とは

- 多くの変数間に存在すると考えられる「原因＋結果」の関係を探索的に同定する
  - 相関的研究 or 観察研究
  - 異なった因果モデルを何らかの指標（統計量）で比較
- 同値（同等）モデルの存在
  - 評価が同等なモデルが多数発生
  - 統計量の大きさでモデルを区別できない
  - いくつかのアプローチが存在
- いくつかの困難
  - 統計的問題
    - 交絡変数，サンプルセレクション，測定誤差
  - 変数の加除に対する不安定性
  - 計算機・アルゴリズム上の問題
    - 大規模ネットワーク推定，非線形性
    - アルゴリズムの変更に対して解が不安定

10

因果の方向を決める
-- 同値モデルへの対策 --

- 第三変数との関わりで決める
  - 自然に決められる場合
  - 能動的に決める場合
    - 操作変数（instrumental variable）
- 線形モデル＋非正規誤差
- 非線形モデル＋正規誤差
- 経時データの利用

- （それでもなお）たくさんの非常に多くの同等に良いモデルが発生
  - グラフマイニング，バスケット分析

同値モデルと
第三変数による因果探索

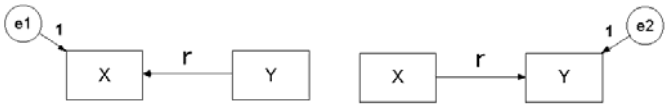
11

12

相関係数から因果の方向は決まらない
---同値モデルの問題---

相関構造

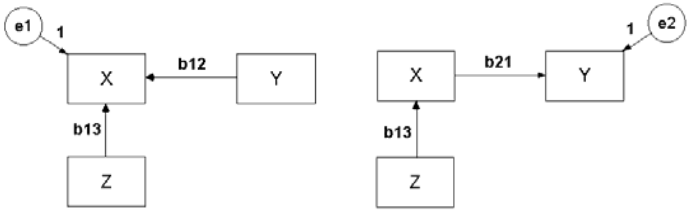
	X	Y
X	1	r
Y	r	1



- データから区別できないモデルを同値モデルという
- 「区別できない」とは適合度の値が同一であること

13

因果の方向を決める



相関構造

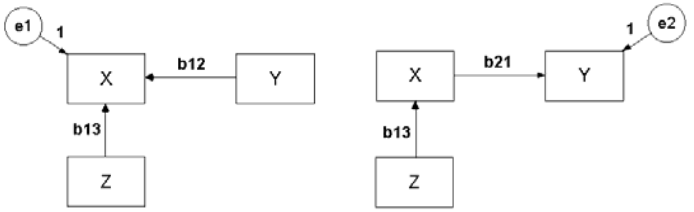
	X	Y	Z
X	1	b12	b13
Y		1	0
Z			1

相関構造

	X	Y	Z
X	1	b21	b13
Y		1	b21b13
Z			1

14

因果の方向を決める：適合度との関係

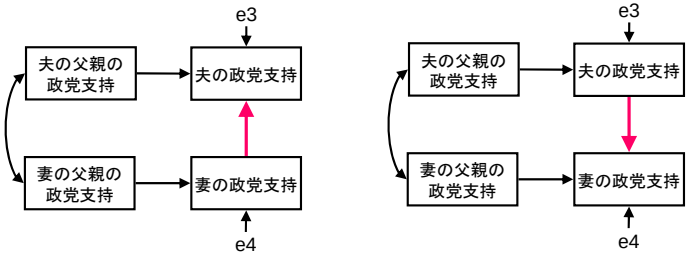


適合度が低い　　適合度が高い

X→Y の因果関係が示唆される

15

政治的社会化モデル



出典：Asher(1976). Causal Modeling. Sage

線形モデル＋非正規誤差

16

17

因果構造とBSS

$$\begin{cases} X_1 = b_{12}X_2 + e_1 \\ X_2 = b_{21}X_1 + e_2 \end{cases}$$

$$\Longleftrightarrow$$

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$\Longleftrightarrow$$

$$x = Be$$

- 未観測の誤差(信号)e を未知の混合行列によって線形混合された観測 x から，Bを推定し信号を復元する→Blind Source Separation

18

BSS と ICA

- $x = Be$において，信号に「非正規＋独立」を仮定すると独立成分分析(ICA)のモデルになる
- 信号が非正規・独立であれば，モデルは識別可能
- b_{ij}の推定・検定により因果の方向を検討できる
- 推定方法
  - FastICA等のICA algorithm の利用
  - 高次モーメントを用いたモーメント推定法

19

実例1

- 母集団とサンプル
  - 大学生
  - 標本サイズ n=222
- X1: 高校時代の詐欺犯罪の回数
- X2: 過去1年間の詐欺犯罪の回数
- 高次モーメントを用いる方法で方向を決定
- 出典
  - Shimizu and Kano (in press) to appear in JSPI
  - Shimizu et al (2006). 拡張版 to appear in CSDA

20

実例1：結果

Cumulants of observed variables

$\mathrm{cum}(x_1x_1x_1)$	$\mathrm{cum}(x_2x_2x_2)$
1.39	0.95
$\mathrm{cum}(x_1x_1x_1x_1)$	$\mathrm{cum}(x_2x_2x_2x_2)$
2.79	1.18

$$\mathrm{cum}(x_i x_j x_k) = E(x_i x_j x_k)$$

$$\mathrm{cum}(x_i x_j x_k x_l) = E(x_i x_j x_k x_l) - E(x_i x_j)E(x_k x_l) - E(x_i x_k)E(x_j x_l) - E(x_i x_l)E(x_k x_j)$$

Estimated path coefficients, standard errors and model fit indices

	b_{12} or b_{21}	T_2 (df)	p value
Model 1′: $x_1 \leftarrow x_2$	0.59 (0.40)	14.64 (5)	0.01
Model 2′: $x_1 \rightarrow x_2$	0.65 (0.12)	3.21 (5)	0.67

21

Some details

$$X_1 = b_{12}X_2 + e_2$$

$$E\begin{bmatrix} m_{20} \\ m_{11} \\ m_{02} \end{bmatrix} = \begin{bmatrix} b_{12}^2 & 1 \\ b_{12} & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} E(x_2^2) \\ E(e_1^2) \end{bmatrix} \quad \text{or} \quad E[m_2] = \sigma_2(\tau_2)$$

$$E\begin{bmatrix} m_{30} \\ m_{21} \\ m_{12} \\ m_{03} \end{bmatrix} = \begin{bmatrix} b_{12}^3 & 1 \\ b_{12}^2 & 0 \\ b_{12} & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} E(x_2^3) \\ E(e_1^3) \end{bmatrix} \quad \text{or} \quad E[m_3] = \sigma_3(\tau_3)$$

$$E\begin{bmatrix} m_{40} \\ m_{31} \\ m_{22} \\ m_{13} \\ m_{04} \end{bmatrix} = \begin{bmatrix} b_{12}^4 & 6b_{12}^2 & 1 \\ b_{12}^3 & 3b_{12} & 0 \\ b_{12}^2 & 1 & 0 \\ b_{12} & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} E(x_2^4) \\ E(x_2^2)E(e_1^2) \\ E(e_1^4) \end{bmatrix} \quad \text{or} \quad E[m_4] = \sigma_4(\tau_4)$$

22

Some details

Estimation

$$\min_{\tau} \left\| \begin{bmatrix} m_2 \\ m_3 \\ m_4 \end{bmatrix} - \begin{bmatrix} \sigma_2(\tau_2) \\ \sigma_3(\tau_3) \\ \sigma_4(\tau_4) \end{bmatrix} \right\|^2$$

Test statistic

$$T = n \min_{\tau} \left\| \begin{bmatrix} m_2 \\ m_3 \\ m_4 \end{bmatrix} - \begin{bmatrix} \sigma_2(\tau_2) \\ \sigma_3(\tau_3) \\ \sigma_4(\tau_4) \end{bmatrix} \right\|^2$$

23

実例2：親子の身長

Third-order cumulants of observed variables

$\text{cum}(x_1x_1x_1)$	$\text{cum}(x_1x_1x_2)$	$\text{cum}(x_1x_2x_2)$	$\text{cum}(x_2x_2x_2)$	
-0.04	0.36	-0.15	-0.03	

Fourth-order cumulants of observed variables

$\text{cum}(x_1x_1x_1x_1)$	$\text{cum}(x_1x_1x_1x_2)$	$\text{cum}(x_1x_1x_2x_2)$	$\text{cum}(x_1x_2x_2x_2)$	$\text{cum}(x_2x_2x_2x_2)$
0.05	-0.17	0.31	-0.85	0.77

Note: $\text{cum}(x_ix_jx_k) = E(x_ix_jx_k)$,

$\text{cum}(x_ix_jx_kx_l) = E(x_ix_jx_kx_l) - E(x_ix_j)E(x_kx_l) - E(x_ix_k)E(x_jx_l) - E(x_ix_l)E(x_kx_j)$.

	b_{12} or b_{21}	T (df)	p value
x_1: Son's height ← x_2: Father's height	0.19 (0.18)	4.90 (5)	0.43
x_1: Son's height → x_2: Father's height	0.16 (0.17)	11.81(5)	0.04
x_1 and x_2 are independent	0*	38.10 (6)	0.00

Note: 0* parameters fixed at 0

非線形モデル＋正規誤差

24

25

同値モデル

$$X_1 = \beta_{12} X_2 + e_1$$

versus

$$X_2 = \beta_{21} X_1 + e_2$$

$$N(x_2|\beta_{12}x_2, \sigma_1^2)N(x_2|0, \sigma_2^2)$$

is equivalent to

$$N(x_1|\beta_{21}x_1, \sigma_2^2)N(x_1|0, \sigma_1^2)$$

26

非線形モデル

$$X_1 = \beta_{12} X_2 + \gamma_{12} X_2^2 + e_1$$

versus

$$X_2 = \beta_{21} X_1 + \gamma_{21} X_1^2 + e_2$$

$$N(x_1|\beta_{12}x_2 + \gamma_{12}x_2^2, \sigma_1^2)N(x_2|0, \sigma_2^2)$$

differs

$$N(x_2|\beta_{21}x_1 + \gamma_{21}x_1^2, \sigma_2^2)N(x_1|0, \sigma_1^2)$$

注意：結果変数は非正規分布

27

少し一般化

- 基底関数をもちいた各種回帰モデル
  （e.g., Imoto, et. al, 2002）

$$X_1 = \sum_{m=0}^{M} \gamma_m^{(12)} b_m(X_2) + e_1$$

differs

$$X_2 = \sum_{m=0}^{M} \gamma_m^{(21)} b_m(X_1) + e_2$$

- e.g., Bayes Approach: モデルの事後確率

$$f_{12}(x_1, x_2|\lambda) = \int \prod_{\alpha=1}^{n} N\left(x_{1\alpha}| \sum_{m=0}^{M} \gamma_m^{(12)} b_m(x_{2\alpha}), \sigma_1^2\right) N(x_{2\alpha}|0, \sigma_2^2) \times \pi_n(\gamma, \sigma_1^2, \sigma_2^2|\lambda) d\gamma d\sigma_1^2 d\sigma_2^2$$

いくつかの困難

28

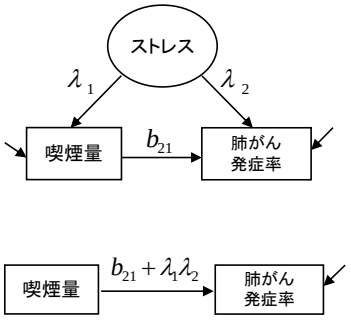
29

交絡変数

- 原因変数と結果変数を結ぶ変数
- 分野によって呼称が違う
  - 第三変数，剰余変数，二次変数，媒介変数，共変量
- 観察研究のアキレス腱

30

未観測交絡変数

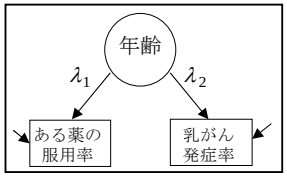


Cor(喫煙量，肺がん発症率）
$=b_{21}+\lambda_1\lambda_2$

Ronald Fisher in 1956

31

未観測交絡変数

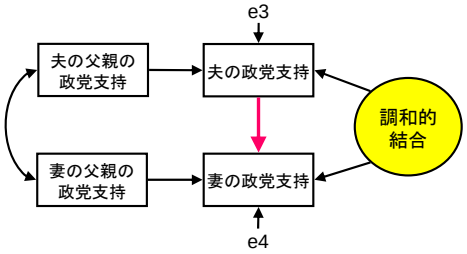


出典：Jick et al (1974).
Reserpine and breast cancer
Lancet, September 21.

薬の服用を減少させても，（年を若返らせることができない限り）乳がん発症率は減少しない

32

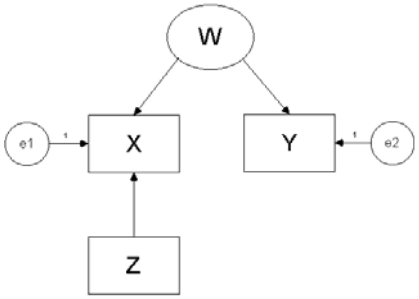
交絡変数はこわい



盛山(1986，行動計量学)

33

因果方向決定にも影響



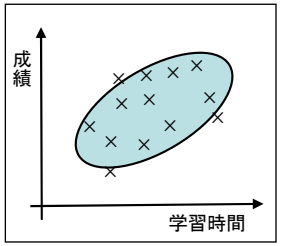
	X	Y	Z
X	1	*	*
Y	*	1	0
Z	*	0	1

誤ってY→Xと結論してしまう

34

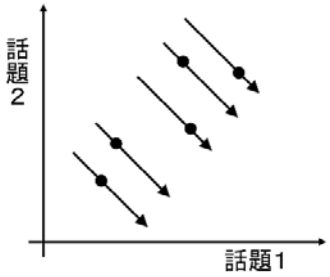
個体内変動と個体間変動

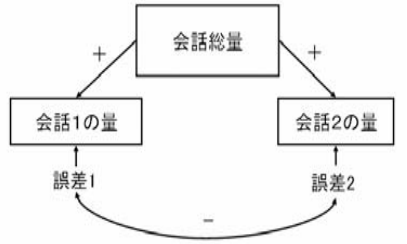
- 1時点における相関研究は個体間の変動にもとづく相関分析
  - Xが小さい個体はYも小さい（傾向がある）
  - Xが大きい個体はYも大きい（傾向がある）
- Xが小さい個体でXを大きくするとYも大きくなる（傾向がある）のか？
  - 個体内の変動
  - 個体間の変動が個体内の変動を近似する必要性
    - エルゴード性
  - 未観測交絡変数の存在はエルゴード性が成り立たない例



35

例：親子の会話

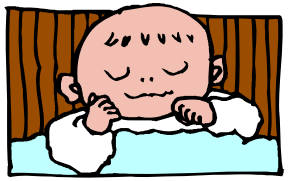


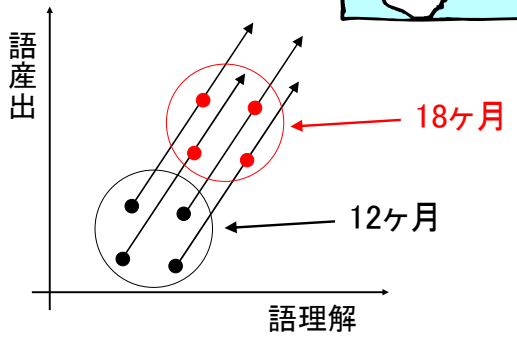


出典：南風原＋小松(1999)

36

例：語理解と語産出





出典：南風原＋小松(1999)

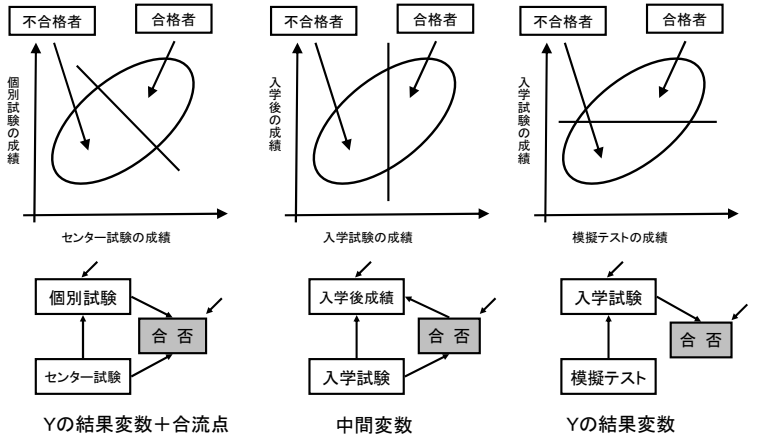
37

サンプルセレクション

- 得られた標本が想定した母集団からの無作為標本でない場合がある
  - 適正な因果分析が実行できない！
- Yの結果変数で選択
  - 欠測指標はYに関係
  - Nonigonrable missing
- Xの結果変数で選択
  - 欠測指標はXに関係
  - Missing at random (MAR)

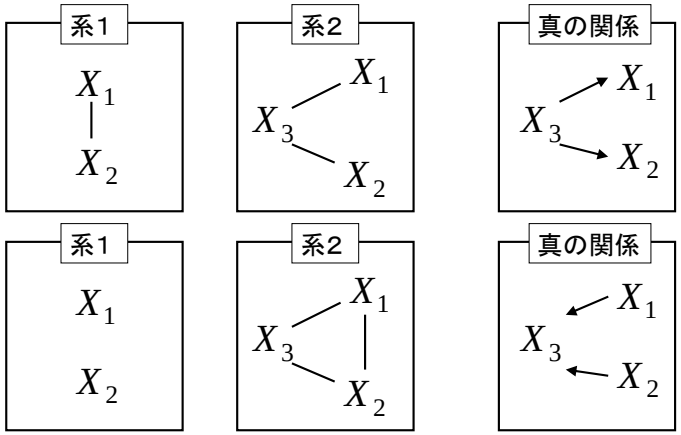
38

具体例



39

変数の加除による変化



40

因果探索と変数の重要性

- 加除によって推定結果が大きく変化する変数は重要
  - 植野真臣（1996）
    - Bayesian networks
  - Rakotomamonjy(2003)
    - SVMにおける変数選択
- 多くの同等に良いモデルにおいて安定して出現する変数とそのパターンは重要

3. SEM, DAG and AG

41

42

表か裏か

- 因果分析は相関行列に基づくべきか，偏相関行列に基づくべきか？
- この問には議論がある
  - 偏相関係数は値が小さくなり使いにくい？

- （真の）因果構造に依存
  - どの変数間の関係か
  - どの変数で条件をつけるべきか

43

SEM_1

- Structural Equation Modeling
  （構造方程式モデリング）
  - 検証的因果推論の道具
  - パス解析モデル＋因子分析モデル
- 原因変数と結果変数の関係を基本的には線形モデルで記述
  - 非線形モデルもある
- 事前仮説に基づく因果モデルをデータとの整合性を検討することで検証する

44

SEM_2

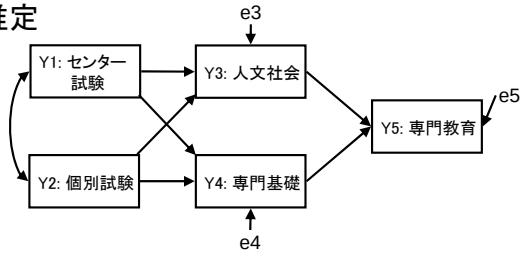
- 測定誤差の調整
  - 潜在変数と多重指標の導入により誤差を分離する
  - 今回は議論しない

45

パス解析モデル

- (観測)変数間の因果モデル
  - 複数個の(線型)回帰モデル
- 推測
  - 適合度の吟味
  - パス係数の推定
  - 効果の分解

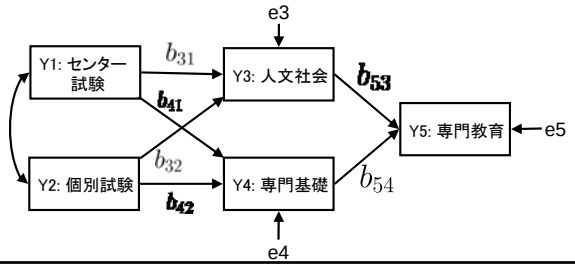
注:選抜効果は考えない



46

構造方程式

$$Y_3 = b_{31}Y_1 + b_{32}Y_2 + e_3$$
$$Y_4 = b_{41}Y_1 + b_{42}Y_2 + e_4$$
$$Y_5 = b_{53}Y_3 + b_{54}Y_4 + e_5$$



47

誘導形

$$\boldsymbol{y} = B\boldsymbol{y} + \Gamma\boldsymbol{x} + \boldsymbol{e}$$

$$\begin{bmatrix}\boldsymbol{y}\\ \boldsymbol{x}\end{bmatrix} = \begin{bmatrix}B & \Gamma\\ O & O\end{bmatrix}\begin{bmatrix}\boldsymbol{y}\\ \boldsymbol{x}\end{bmatrix} + \begin{bmatrix}\boldsymbol{e}\\ \boldsymbol{x}\end{bmatrix}$$

$$\begin{bmatrix}\boldsymbol{y}\\ \boldsymbol{x}\end{bmatrix} = \begin{bmatrix}I-B & -\Gamma\\ O & I\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{e}\\ \boldsymbol{x}\end{bmatrix}$$

48

共分散構造とパラメータ

$$\begin{bmatrix}\boldsymbol{y}\\ \boldsymbol{x}\end{bmatrix} = \begin{bmatrix}I-B & -\Gamma\\ O & I\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{e}\\ \boldsymbol{x}\end{bmatrix}$$

$$V\begin{bmatrix}\boldsymbol{y}\\ \boldsymbol{x}\end{bmatrix} = \begin{bmatrix}I-B & -\Gamma\\ O & I\end{bmatrix}^{-1}\begin{bmatrix}V(\boldsymbol{e}) & O\\ O & V(\boldsymbol{x})\end{bmatrix}\begin{bmatrix}I-B' & O\\ -\Gamma' & I\end{bmatrix}^{-1} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$$

- 推定すべきパラメタ θ
  - パス係数
  - 独立変数の分散・共分散

49

統計的推測

- 尤度

$$L(\mu, \Sigma(\theta)) := \prod_{\alpha=1}^{n} N_p\left(\begin{bmatrix} y_\alpha \\ x_\alpha \end{bmatrix} \middle| \mu, \Sigma(\theta)\right)$$

- 最尤推定

$$(\hat{\mu}, \hat{\theta}) := \underset{(\mu,\theta)}{\operatorname{argmax}}\ L(\mu, \Sigma(\theta))$$

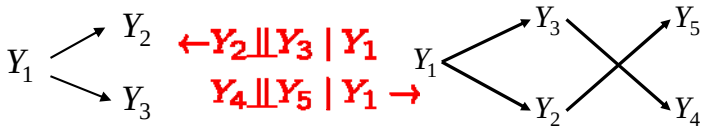
- 適合度検定

$$H_0 : V\begin{bmatrix} y_\alpha \\ x_\alpha \end{bmatrix} = \Sigma(\theta) \quad \text{vs.} \quad H_1 : V\begin{bmatrix} y_\alpha \\ x_\alpha \end{bmatrix} = \Sigma$$

50

DAG (Directed Acyclic Graph)

- DAG: G = (V, E)
  - V: 頂点集合; E: 辺集合
- 矢線が巡回しない．逐次型
- グラフと条件付独立性とが対応がつくように同時分布を定義
  - 隣接しない⇔条件付独立



51

DAGと同時密度との関連付け

- 同時密度の因数分解

$$f(y_1, \cdots, y_p) = \prod_{i=1}^{p} f_i(y_i | pa(y_i))$$

52

条件付独立

- 変数X,Yが辺で直接結ばれていなければ，なんらかの変数集合Sが存在し，Sを与えた下でX,Yは条件付き独立となる
  - $X \perp\!\!\!\perp Y | S$
- Sの同定方法
  - 大域的マルコフ性
    - モラルグラフ
    - 有向分離（d-separation）
  - （局所的マルコフ性）

53

DAGと条件付独立

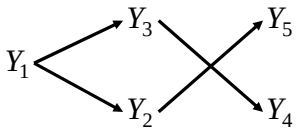
$$f(y_1,y_2,y_3,y_4,y_5)=\prod_{i=1}^{5} f_i(y_i|pa(y_i))$$
$$=f_1(y_1)f_2(y_2|y_1)f_3(y_3|y_1)f_4(y_4|y_3)f_5(y_5|y_2)$$

$\therefore$

$$f(y_1,y_4,y_5)=\int\int f(y_1,y_2,y_3,y_4,y_5)dy_2dy_3$$
$$=g(y_1,y_4)h(y_1,y_5)$$

$\therefore$

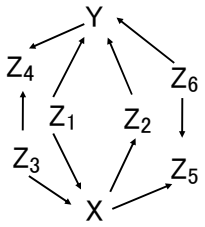
$Y_4 \perp\!\!\!\perp Y_5|Y_1$



54

有向分離(d-separation)

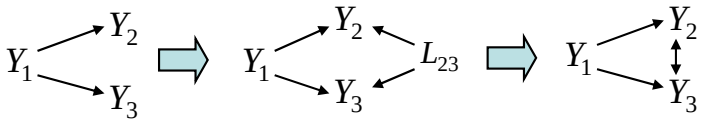
- Y_αとY_βを条件付独立にするSについての条件
  - Y_αとY_βの間の有向道(間接効果)は殺す
  - 合流点で条件付けるときは慎重に
- d-separation
  - XとYを結ぶ各道において，以下のどちらかが成立
    - 合流点があるとき，
      zは合流点とその子孫を含まない
    - 非合流点があるとき，
      zは少なくとも1つの非合流点を含む



55

DAGの拡張

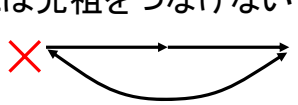
- DAGはやや制約的
  - 偏相関は存在しないという仮定
    - 子供(従属変数)どうしの相関は親が説明
  - 測定されていない交絡変数の問題
- Ancestral Graphへ



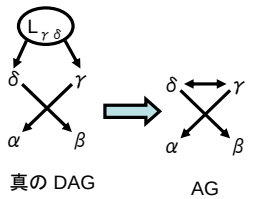
56

AG

- Ancestral Graph
  - Richardson and Spirtes (2002). Ann. Statist.
- Lを観測すれば DAG になる
- Lは先祖をつなげない



- AGの性質は未だ十分に解明されていない
  →だから面白い



57

AGと条件付独立

- 辺で直接結ばれていない変数Y_α, Y_βについて
  - ある観測変数の集合を与えた下で条件付独立になる場合
    - Maximal AG という
    - $Y_\alpha \perp\!\!\!\perp Y_\beta \mid \{Y_\gamma, Y_\delta\}$

条件付独立

  - どんな観測変数の集合を選んでも条件付独立にならない場合
    - 代数的制約が発生（正規性）

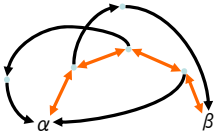
$$\sigma_{\alpha\beta} = \frac{\sigma_{\alpha\gamma}\sigma_{\gamma\beta}}{\sigma_{\gamma\gamma}} + \frac{\sigma_{\alpha\delta}\sigma_{\delta\beta}}{\sigma_{\delta\delta}} - \frac{\sigma_{\alpha\gamma}\sigma_{\gamma\delta}\sigma_{\delta\beta}}{\sigma_{\gamma\gamma}\sigma_{\delta\delta}}$$

条件付独立でない

58

隣接しない≠条件付独立

α, β 間の primitive inducing path:
- 双方向辺のみで構成されている
- 全ての頂点$\in$an(α)$\cup$an(β)



MAG:
α, β 間に primitive inducing path があるならば $\alpha \leftrightarrow \beta$ も必ず存在するような AG

Non-maximal な AG:
α, β 間に primitive inducing path ＋隣接しない
⇔ 条件付独立を意味しない制約

an({α, β}) の要素は条件付けなければならないが，同時に合流点でもあるので条件付けが行えない．

59

AGにおける代数的制約と Bi-partial covariance

- Bi-partial covarianceの定義
  - 互いに素なインデックス集合：$\{\alpha, \beta\}, A, B, C$

$$e_\alpha = Y_\alpha - \left(\sum_{a \in A} \tilde{\gamma}_a Y_a + \sum_{c \in C} \tilde{\gamma}_c Y_c \right)$$

$$e_\beta = Y_\beta - \left(\sum_{b \in B} \tilde{\gamma}'_b Y_b + \sum_{c \in C} \tilde{\gamma}'_c Y_c \right)$$

$$\Longrightarrow \sigma_{(\alpha,A),(\beta,B)\cdot C} := \mathrm{Cov}(e_\alpha, e_\beta)$$

- Timm and Carlson (1976). Psychometrika.
- Miyamura and Richardson (2006) in preparation.

60

Properties of Bi-partial covariance

- 陽表現

$$\sigma_{(\alpha,A),(\beta,B)\cdot C} = \begin{vmatrix} \sigma_{\alpha\beta\cdot C} & \Sigma_{\alpha,A\cdot C} & \Sigma_{\alpha,B\cdot C} \\ \Sigma_{A,\beta\cdot C} & \Sigma_{A,A\cdot C} & \Sigma_{A,B\cdot C} \\ \Sigma_{B,\beta\cdot C} & 0 & \Sigma_{B,B\cdot C} \end{vmatrix}$$

where $\Sigma_{\alpha,A\cdot C} = \Sigma_{\alpha,A} - \Sigma_{\alpha,C}\Sigma_{C,C}^{-1}\Sigma_{C,A}$

- 補足：A＝ϕのときは偏共分散に比例

$$\begin{aligned} \sigma_{(\alpha,\phi),(\beta,B)\cdot C} &= \left|\Sigma_{B,B\cdot C}\right| \times \left(\sigma_{\alpha\beta\cdot C} - \Sigma_{\alpha,B\cdot C}\Sigma_{B,B\cdot C}^{-1}\Sigma_{B,\beta\cdot C}\right) \\ &= \left|\Sigma_{B,B\cdot C}\right| \times \left(\sigma_{\alpha\beta} - \Sigma_{\alpha,B\cup C}\Sigma_{B\cup C,B\cup C}^{-1}\Sigma_{B\cup C,\beta}\right) \\ &= \left|\Sigma_{B,B\cdot C}\right| \times \sigma_{\alpha\beta\cdot B\cup C} \end{aligned}$$

61

AG と bi-partial covariance

- α, β に対して A, B, C を以下のように選ぶ

  A = an(α)＼(an(β) ∪{α, β})
  B = an(β)＼(an(α) ∪{α, β})
  C = (an(α)∩ an(β))＼{α,β}
  ここで an(α) はαの先祖集合

AG で隣接しない＝「$\sigma_{(\alpha,A),(\beta,B).C}=0$」

$$\sigma_{(\alpha,\gamma),(\beta,\delta)}=\begin{vmatrix}\sigma_{\alpha\beta} & \sigma_{\alpha\gamma} & \sigma_{\alpha\delta}\\ \sigma_{\gamma\beta} & \sigma_{\gamma\gamma} & \sigma_{\gamma\delta}\\ \sigma_{\delta\beta} & 0 & \sigma_{\delta\delta}\end{vmatrix}=0 \quad\Leftrightarrow\quad \sigma_{\alpha\beta}=\frac{\sigma_{\alpha\gamma}\sigma_{\gamma\beta}}{\sigma_{\gamma\gamma}}+\frac{\sigma_{\alpha\delta}\sigma_{\delta\beta}}{\sigma_{\delta\delta}}-\frac{\sigma_{\alpha\gamma}\sigma_{\gamma\delta}\sigma_{\delta\beta}}{\sigma_{\gamma\gamma}\sigma_{\delta\delta}}$$

62

本節のまとめ_1

- SEM: Structural Equation Modeling
  - 矢線は直接効果を記述
  - 独立関係よりも「効果」に重点
    - 直接効果，間接効果，総合効果
  - 従属変数間の相関は親変数による擬相関と誤差共分散によって説明
  - DAGでは誤差共分散を表現できない
    →AGによって実現

63

本節のまとめ_2

- DAG: Directed Acyclic Graph
  - 探索的要素が強い
  - グラフが同時分布を規定する
    - 分布形は問わない
  - （条件付）独立性をキーワードとする
    - 矢線がない変数間には何らかの条件付独立が成立
      - 同時分布の因数分解＋有向分離
  - DAGで線形構造方程式モデルを考えるときはSEMと同等と思ってよい
    - 思想的には異なる
- AGとMAGは未観測潜在変数を想定
  - 条件付独立やbi-partial covarianceと関連
  - 今後の発展を待ちたい

まとめと文献

64

65

まとめ

- 因果の方向を定める
  - 困難な問題と考えられている
  - いくつかの統計的方法論が存在
    - パス解析（操作変数法）
    - 非正規性の積極利用（ICA）
    - 非線形性の利用
    - 経時データの利用
- 方向は既知の下で
  - 因果関係が存在するかどうか
  - 因果関係が存在するとき，その大きさを如何に「正確に」評価するか
    - 無作為割付けの伴う実験的研究がよいが，観察研究に頼らざるを得ない場合も多い

66

まとめ

- 多変数間の因果ネットワークの推定
  - 条件付独立の情報は有益
  - 条件を付けることの基本的性質をふまえて探索
  - DAG→AGの発展を待ちたい
- 調査データに基づく因果推論は非常に脆弱なものと認識
  - 交絡変数，サンプルセレクション，測定誤差
  - 交絡変数から確率的に独立な無作為割付のもとでの実験がベスト
- 実験が不可能な場合，交絡変数の影響を殺す最大限の努力が必要
  - 交絡変数を観測しモデリング
  - 交絡変数の値が同じ仮想的な観測値をつくり比較
    - Rubinの反事実モデル

67

参考文献

- Bollen, K. A. (1989). Structural Equations with Latent Variables. Wiley.
- Bullock, H. E., Harlow, L. L. & Mulaik, S. A. (1994). Causal issues in structural equation modeling research. Structural Equation Modeling, 1, 253–267.
- Holland, P. W. (1986). Statistics and causal inference (with discussion). Journal of the American Statistical Association, 81, 945–970.
- Hirano, K., Imbens, G. & Ridder, G. (2003). Efficient estimation of average treatment effect using the estimated propensity score. Econometrica, 71, 1161–1189.
- Imoto, S, et al. (2002). Estimation of genetic networks and functional structures between genes by using Bayesian network and nonparametric regression. Pacific Symposium on Biocomputing, **7**, 175–186.
- Imoto, S. et al. (2004). Combining microarrays and biological knowledge for estimating gene networks via Bayesian networks. Journal of Bioinformatics and Computational Biology, 2(1), 77–98.
- Mulaik, S. A. & James, L. R. (1995). Objectivity and reasoning in science and structural equation modeling. In Structural Equation Modeling: Concepts, Issues, and Applications, (Hoyle, H., Ed.), pp.118–137. Sage Publications: CA.
- Rakotomamonjy, A. (2003). Variable selection using SVM-based criteria. Journal of Machine Learning Research 3, 1357–1370.
- Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70, 41–55.
- Rosenbaum, P. R. (2002). Observational Studies. 2nd ed. Springer.
- Shimizu, S., Hyvärinen, A., Hoyer, P. O. and Kano, Y. (2006). Finding a causal ordering via independent component analysis. Computational Statistics and Data Analysis. Vol.50, No.11, 3278–3293
- Shimizu, S. and Kano, Y. (in press). Use of non-normality in structural equation modeling: Application to direction of causation. Journal of Statistical Planning and Inference. (accepted on January 7, 2006)

68

- 井元清哉(2006). マイクロアレイデータからの遺伝子間因果に関する知識発見. 日本統計学会75周年記念シンポジウム. 日本統計学会75周年記念第1回研究集会講演報告集. 81–91.
- 岩崎 学(2002). 不完全データの統計解析. エコノミスト社
- 植野真臣(1996). 意思決定アプローチによるBayesian networkの因果モデル構築. 人工知能学会誌, 11/5, 49–58.
- 狩野 裕 (2002).「構造方程式モデリング，因果推論，そして非正規性」竹内啓（編著）多変量解析の展開 Part II. 岩波書店
- 佐藤俊哉・松山 裕 (2002).「疫学・臨床研究における因果推論」竹内啓（編著）多変量解析の展開 Part III. 岩波書店
- 竹内啓(1986). 因果関係と統計的方法. 行動計量学, 14, 85–90.
- 豊田秀樹(1998). 共分散構造分析[入門編]. 朝倉書店
- 星野・繁桝(2004). 傾向スコア解析法による因果効果の推定と調査データの調整について. 行動計量学, 31, 43–61.
- 星野崇宏(2005). 欠測群の周辺分布の母数に対する傾向スコアを用いた重み付きM推定量の提案と介入効果研究への応用. 行動計量学, 32, 121–132.
- 星野崇宏(2006). 傾向スコアを用いた準実験及び観察研究からの因果分析について. 2006年応用統計学会チュートリアルセミナー資料集. Pp.11–38
- 宮川雅巳 (1997). グラフィカルモデリング. 朝倉書店
- 宮川雅巳 (2004). 統計的因果推論. 朝倉書店
- 宮川・黒木(1999). 因果ダイアグラムにおける介入効果推定のための共変量選択. 応用統計学, 28, 151–162.
- 鷲尾 隆(2006). データマイニングとその統計的因果推論への適用. 日本統計学会75周年記念第1回研究集会講演報告集. 92–99.