

疎構造学習のグラフィカルラッソに基づく
変化検知に関する基礎研究
ー生活習慣病データを対象としてー

社会システム計画研究室
都14-0158
金 ヨンギョン

グラフィカルラッソによる疎構造分析

精度行列 $\Lambda = (\lambda_{ij}) \Rightarrow$ 隣接行列

$$\Lambda^* = \begin{pmatrix} * & * & * & * & * & 0 \\ * & * & * & * & * & 0 \\ * & * & * & 0 & 0 & 0 \\ * & * & 0 & * & 0 & 0 \\ * & * & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}$$

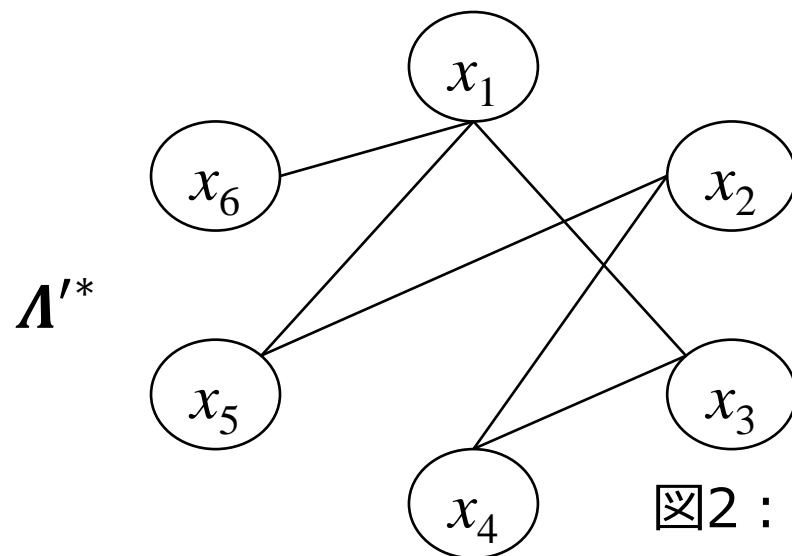


図2：構造変化例

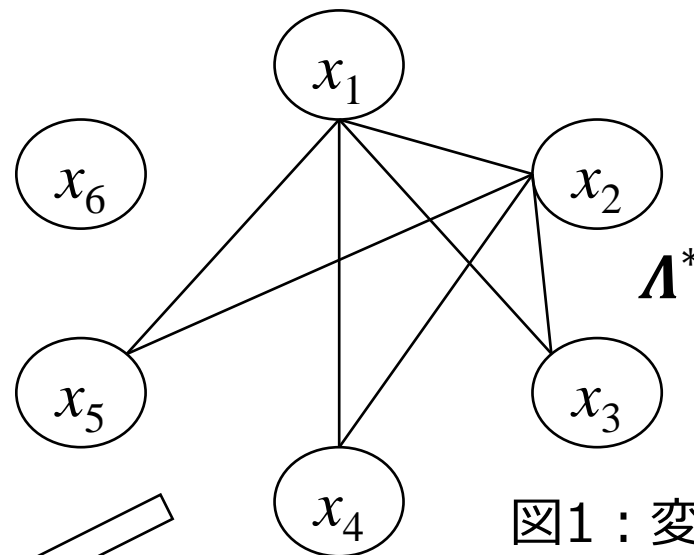


図1：変数間の相関グラフ

$$\Rightarrow \Lambda^* = \arg \max_{\Lambda} (\ln \det \Lambda - \text{tr}(S\Lambda) - \rho \|\Lambda\|_1)$$

S : データの標本分散共分散行列
 $\rho > 0$: 正則化パラメータ

研究目的

- ◆ 生活習慣病データを対象とし、
グラフィカルラツソを用いて疎構造学習を行い、
 - データ集合間の構造変化
 - 要因（項目）ごとの変化度

糖尿病データの概要

◆ 糖尿病(Pima Indians Diabetes Database)

- 1990年のアメリカ国立糖尿病学会
- 年齢21歳以上女子の**768人**(重複なし)のデータ使用

◆ 検査項目(8項目)

- (1) 妊娠回数 (2) 経口耐糖能検査における血清血糖値 (3) 拡張期血圧
(4) 三頭筋皮下脂肪厚 (5) インシュリン (6) 体格指数
(7) 糖尿病家系関数 (8) 年齢

認知症データの概要

◆ 認知症(Magnetic Resonance Imaging Comparisons of Demented and Nondemented Adults)

- 2010年のワシントン大学の認知症研究所
- 150人(重複あり)が373回のMRI撮影

◆ 検査項目(11項目)

- (1) 訪問回数 (2) MR 遅延時間 (3) 年齢 (4) 教育期間
- (5) 社会的立場 (6) 精神短期間検査 (7) 行動検査 (8) 頭蓋内部体積
- (9) 脳体積 (10) ASF (11) 性別

認知症の基礎分析：t-検定

要素	t 値	p 値
訪問回数	-1.85	0.03
MR 遅延時間	-2.35	0.01
年齢	-0.11	0.45
教育期間	-3.79	0.00
社会的立場	1.16	0.12
精神短期間検査	-11.58	0.00

要素	t 値	p 値
行動検査	23.43	0.00
頭骨内部体積	-0.82	0.21
脳体積	-6.33	0.00
ASF	0.63	0.27
性別	-4.38	0.00

グラフィカルラッソによる構造分析

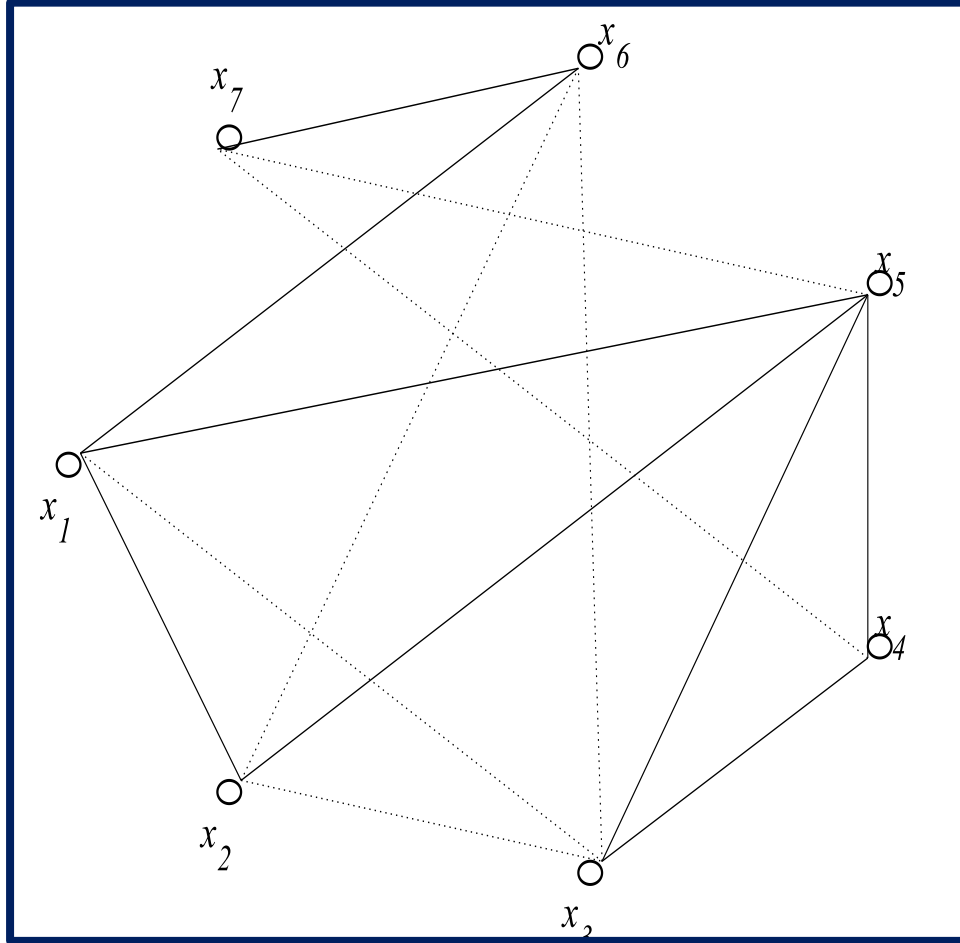


図3：健康者データに対する相関グラフ

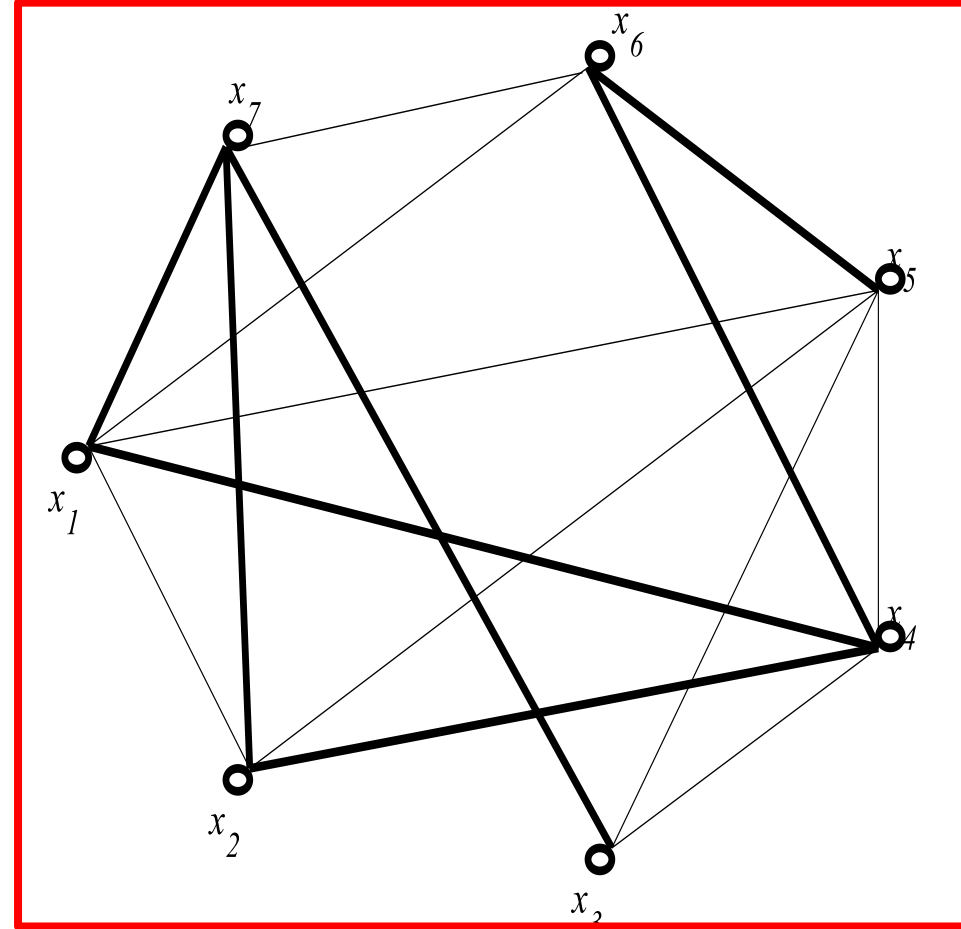


図4：認知症患者データに対する相関グラフ

	要素
X_1	訪問回数
X_2	MR 遅延時間
X_3	教育期間
X_4	精神短期間検査
X_5	行動検査
X_6	脳体積
X_7	性別

まとめと今後の課題

- t検定により、精神短期間検査と行動検査の平均差が大きい。
- 行動検査0である人は脳体積が大きい。
- 女子健常者は病院によく訪問し、MRI 時間が遅くなる。
- 男子認知症者は病院によく訪問しており、精神短期間検査は女子認知症者より高い。

精神短期間検査は認知症診断の検査項目間との影響性が高いと考えられる。

今後、分析した項目が診断に重要な要因になるかを検討する必要がある。

参考文献

- データヘルス計画HP 厚生労働者
- 井出剛：疎な相関グラフの学習による相関異常の検出, 2009
http://ide-research.net/papers/2009_DMSM_Ide.pdf
- 糖尿病(1990年)と認知症(2010年)のデータ Kaggle
<https://www.kaggle.com/>