

Chapter 7: Using Correspondence Analysis for Joint Display of Affiliation Networks

濱崎雅弘 (hamasaki@ni.aist.go.jp)
産業技術総合研究所 情報技術研究部門

SNS輪読会
2005年7月27日(水)

2005-07-27 SNS輪読会 濱崎雅弘

自己紹介

- 氏名: 濱崎雅弘
- 所属: 産業技術総合研究所 情報技術研究部門
実世界指向インタラクショングループ
- 専門:
 - 情報共有, オンラインコミュニティ支援など
 - Web上に情報はたくさんあるが結局まずは人に聞いた方が早かったことが多かったのでじゃそれを上手く支援する技術について色々考えたいと思ってこういう所に辿り着きました
 - 基本的にAI(Artificial Intelligence)を作るよりはIA(Intelligent amplifier)を作りたい人。だが、人工知能学会との縁が長かったせいかどうかすればAIはできるのかというのも興味がある

2005-07-27 SNS輪読会 濱崎雅弘

概要

- アクターとイベントの二種類のノードが含まれるアフィリエーションネットワークを、どうやって書いたら、綺麗に、わかりやすくできるか?
- 本章ではコレスポネンス分析という手法を用いてアフィリエーションネットワークを一つの空間に記述する

2005-07-27 SNS輪読会 濱崎雅弘

フィリエーションネットワーク

- アフィリエーションネットワークはアクターとイベントの二種類のノードから構成されるネットワーク
 - アクター: meaningful social unit
 - イベント: collection of actors
- アフィリエーションネットワークには2種類のノードがあるので普通のSocial Networkの表現手法は適さない
 - ではどのような手法で表現するか?

2005-07-27 SNS輪読会 濱崎雅弘

コレスポネンス分析

- 要素間の相関係数を高くする次元を設定する



- 特異値分解を用いて次元を圧縮する

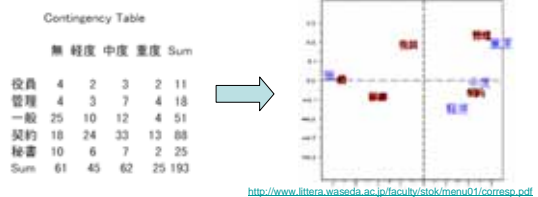
$$A = U D V^T$$

$$A = \lambda_1 u_1 v_1^T + \lambda_2 u_2 v_2^T + \dots + \lambda_r u_r v_r^T$$

<http://case.fz.nms.okayama-u.ac.jp/statedu/hw2-book/node182.html>
2005-07-27 SNS輪読会 濱崎雅弘

コレスポネンス分析の利用法

- 二組の質的変数の関係を分析する



クロス集計表をもとに、列要素と行要素のカテゴリ同士をそれぞれ同じ空間上にマッピングする

2005-07-27 SNS輪読会 濱崎雅弘

コレスポネンス分析が示すもの

- 次元の圧縮ができるので多次元データを見やすくできる
- 異なる要素間で高い相関係数を示す軸を用いているので、同じ空間上にデータを乗せられる
- **同じ要素の2点間の距離の2乗は 2乗距離を示す**
 - 2乗距離とは傾り方の似ている度合い
 - 似ているほど2点間の距離は小さくなる
 - Ex. ActorとActorの距離, EventとEventの距離
- **異なる要素の2点間の距離は意味無し**
 - 近くにあるからにはなんらかの関係性があるのだからけど、2点間の距離には数学的な解釈はなされていない
 - Ex. ActorとEventの距離

2005-07-27 SNS輪読会 濱崎雅弘

アフィリエーションネットワークをどのように表現するか

- 異なる要素間の距離には厳密な意味はない
- 各要素のスケージングの仕方によってどのように変わるか
 - Principal Coordinate (主座標: 分散 = 固有値)
 - Standard Coordinate (標準座標: 分散 = 1)
 - Carroll, Green, and Schaffer Coordinates (分散 = 固有値の平方)

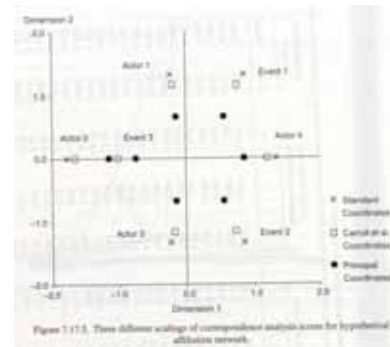
2005-07-27 SNS輪読会 濱崎雅弘

スケージング手法による表現の違い

- Fig. 7.16.3
 - Actor: standard coordinates
 - Event: principal coordinates
 - Actor1とevent3が同じ位置 (Actor1はEvent3にしか属していないため)
 - ActorはEventの囲みの中にいる
- Fig. 7.16.4
 - Actor: principal coordinates
 - Event: standard coordinates
 - Event はActorの囲みの中にある (Fig. 7.16.3と逆の傾向)
- Fig. 7.16.5
 - Actor: Carroll, Green, and Schaffer Coordinates
 - Event: Carroll, Green, and Schaffer Coordinates
 - ActorとEvent両方が輪郭を作る

2005-07-27 SNS輪読会 濱崎雅弘

スケージング手法ごとのプロット位置の比較



2005-07-27 SNS輪読会 濱崎雅弘

まとめ

- アフィリエーションネットワークは2種類のノードがあるので既存のSocial Networkの表現手法とは異なる手法が必要
- コレスポネンス分析は2種類の要素を同一空間上にマッピングできるのでアフィリエーションネットワークの2種類のノードを両方同時に表現可能である
 - 同一要素間の距離の2乗は 2乗値
 - 異なる要素間の距離は関係ない
 - ただし異なる要素間の位置関係から傾向を読み取るのは可
- 各要素の縮尺の設定次第で見え方が変わる
 - 標準座標
 - 主座標
 - Actor & Event: Carroll, Green, and Schaffer座標

2005-07-27 SNS輪読会 濱崎雅弘

参考資料

- コレスポネンス分析
 - <http://www.interscope.co.jp/method/m05.html>
 - <http://www.littera.waseda.ac.jp/faculty/stok/menu01/corresp.pdf>
- 特異値分解
 - <http://staff.aist.go.jp/toru-nakata/Gauss/Gauss2.html>
 - <http://case.f7.ems.okayama-u.ac.jp/statedu/hbw2-book/node182.html>

2005-07-27 SNS輪読会 濱崎雅弘

スケーリング手法の比較 (Fig. 7.16.3)

- Actor: standard coordinates, Event: principal coordinates
- Actor1とevent3が同じ位置 (Actor1はEvent3にしか属していないため)
- ActorはEventの囲みの中にある

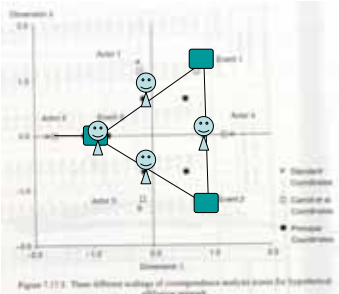


Figure 7.16.3. Three different subplots of correspondence analysis plots for hypothetical alliance network.

スケーリング手法の比較 (Fig. 7.16.4)

- Actor: standard coordinates, Event: principal coordinates
- Actor1とevent3が同じ位置 (Actor1はEvent3にしか属していないため)
- ActorはEventの囲みの中にある

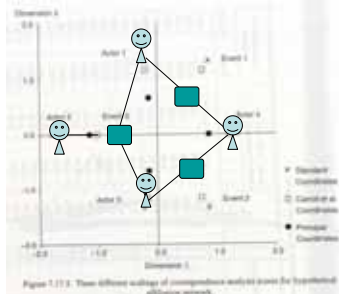


Figure 7.16.3. Three different subplots of correspondence analysis plots for hypothetical alliance network.

スケーリング手法の比較 (Fig. 7.16.5)

- Actor & Event: Carroll, Green, and Schaffer Coordinates

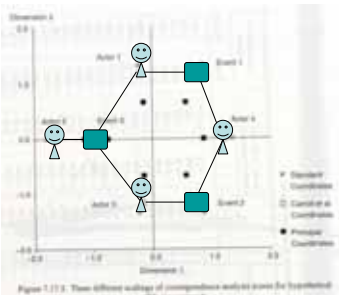


Figure 7.16.3. Three different subplots of correspondence analysis plots for hypothetical alliance network.