

疎構造学習による放牧牛の社会性ネットワークの推定 Estimation of social networks of cattle by sparse structure learning

永岡 優¹⁾ 小松 瑞果¹⁾ 大川 剛直¹⁾ 大山 憲二²⁾
Yu Nagaoka¹⁾ Mizuka Komatsu¹⁾ Takenao Ohkawa¹⁾ Kenji Oyama²⁾

1 はじめに

近年、IoT (Internet of Things) の発展に伴い、農畜産業界では ICT を導入するスマートアグリが盛んに進められている。本研究では、放牧牛から取得したデータをもとにその状態把握に取り組んでいる。一般に、放牧牛は社会性を有するという背景がある。そこで、動物の社会性の分析において有効性が示されつつある、グラフ理論に基づく社会性ネットワーク分析のアプローチに注目した。

動物に対する社会性ネットワーク分析では、対象の動物や検証する仮説に基づき分析者が経験的な方法でネットワークを構築することが一般的である。しかし、このような方法では、構築されるネットワークが分析者により変化しえるため、結果に大きな影響を与えることが課題である。そこで本研究では、放牧牛を対象としたデータ駆動型のネットワークの推定を提案する。さらに、本手法によって得られるネットワークの、社会性分析における有効性を検証するため、実際の観察記録と照合した評価を行う。

2 関連研究

2.1 社会性ネットワークの構築

動物の社会性ネットワーク分析 [1] [2] とは、動物の個体を頂点 (node)、個体同士のアソシエーションを辺 (edge) として表現したネットワーク (グラフ) を構築し、ネットワークから算出される統計量を分析することで、ある集団に関する仮説を検証するアプローチである。「アソシエーション」とは、分析者が検証する仮説に応じて設定する二個体の関係を指す。辺を定める閾値やアソシエーション指標は、対象とする動物や仮説に応じて設定する。このような、経験的に構築されたネットワークは、分析者が定める閾値や指標などに依存する。そこで、分析者に依らず、データに対して一意に定まり、かつ、仮説検証に応用可能なネットワーク構築手法が望まれている。

2.2 構造学習および Graphical Lasso

構造学習 [4] とは、多変量のサンプルが与えられた際に、それらが独立同分布に従うという仮定のもとグラフ構造を推定する機械学習のタスクである。ここでは、与えられたサンプルが多変量正規分布に従うという仮定したのもと、これを説明する。多変量正規分布は、式 (1) のように表せる。

$$\mathcal{N}(x | \mathbf{0}, \Lambda^{-1}) \equiv \frac{|\Lambda|^{\frac{1}{2}}}{(2\pi)^{\frac{M}{2}}} \exp\left(-\frac{1}{2}x^T \Lambda x\right) \quad (1)$$

1) 神戸大学大学院 システム情報学研究科. Graduate School of System Informatics, Kobe University

2) 神戸大学大学院 農学研究科附属食資源教育研究センター. Food Resources Education and Research Center, Graduate School of Agricultural Science, Kobe University

ここで、 M は特徴量の数、 x は標準化データ、 $\Lambda \in \mathbb{R}^{M \times M}$ は分散共分散の逆行列であり精度行列という。精度行列 Λ は特徴量同士の直接相関を表す行列、すなわち特徴量間のネットワークを定める行列に相当する。構造学習では、与えられたデータに対し Λ 当てはめることで、最適な Λ を推定する。

Graphical Lasso [5] は構造学習のなかでも、スパース (疎) なグラフ、すなわち要素に 0 を多く含む Λ を推定する手法である。グラフの疎性は、高い解釈性とノイズに対する頑健性という二つのメリットをもたらす。そのため、遺伝子発現ネットワークの推定などで用いられている。Graphical Lasso では、疎なネットワークを推定すべく損失関数に L_1 正則化項を導入し最尤推定を行う。損失関数 $l_\alpha(\Lambda)$ を式 (2) に示す。

$$l_\alpha(\Lambda) = -\log |\Lambda| + \text{tr}(\Lambda S) + \alpha \sum_i \sum_j |\Lambda_{i,j}|. \quad (2)$$

ここで、 S はサンプルの分散共分散行列、 α はペナルティ係数を表す。

2.3 Graphical Lasso による推定の検定

Coloigner ら [6] では、Graphical Lasso により推定されたネットワークの検定を行っている。この研究では、鎌状赤血球症 (SCD) の患者および健常者の fMRI 画像データを入力として、脳結合ネットワークの推定を行う。その際、SCD の脳結合ネットワークへの影響を検証するため、サンプルを「SCD 患者」と「健常者」の 2 群に分割し、並べ替え検定を行っている。

3 提案手法

放牧牛に関して、社会的紐帯をもつ個体同士で、先導・追従・距離の保持など行動の同期がみられることが知られている [3]。そこで、本研究では位置の近接性に注目したネットワーク推定を行い、社会性の仮説検証への有効性を評価する。具体的には、各個体の時系列位置データを入力として、Graphical Lasso を適用することで放牧牛の社会的関係を記述することを提案する。Graphical Lasso の採用理由には、社会的ネットワーク分析の構築においても解釈性および頑健性が重要であることが挙げられる。本研究では各個体の放牧場における x 座標、 y 座標における精度行列 Λ をそれぞれ推定する。

3.1 Graphical Lasso によるネットワーク推定およびインタラクションとの関係の検証

本研究では、Graphical Lasso により推定した精度行列の変化と、個体同士のインタラクション有無と関係を有するという仮説を検証する。本研究では「計測期間中にインタラクションが発生したデータ群 (以下「実験群」とする)」と「対照群」を用意する。実験群をさらに「インタラクション発生日より前 (前期)」と「発生日以後 (後期)」に分割する。実験群の「前期-前期間および後期-後期間 (同種間)」と「前期-後期間 (異種間)」間について精度行列 Λ の類似度を算出する。この類似度が、各日時のデータと「全期間のデータ (通期)」の

データの類似度に対して大きかった期間の比率を、「同種間」と「異種間」で比較する。ここで、実験群と対照群で「同種間」「異種間」で類似していた期間の比率に有意な差が認められれば、Graphical Lasso により推定したグラフの変化とインタラクションの有無との関係が確かめられる。

3.2 ネットワーク類似度の評価

推定した精度行列間の類似度の算出には対数尤度比を用いる。尤度とは、統計モデルのパラメータ（ここでは精度行列 Λ ）から観測データが得られる確率であり、対数尤度比はモデル間の適同度の差を表す。尤度 L および対数尤度比 R は算出方法は式 (3) (4) に示す。

$$L(\mathbf{0}, \Lambda^{-1} | \mathcal{D}) = \prod_{i=1}^N f_X(x | \mathbf{0}, \Lambda^{-1}) \quad (3)$$

$$R = (\log L_{1,x} + \log L_{1,y}) - (\log L_{2,x} + \log L_{2,y}) \quad (4)$$

$f_X(x | \mathbf{0}, \Lambda^{-1})$ は式 (1) に基づく確率密度関数、 $L_{k,x}$, $L_{k,y}$ はそれぞれ入力データ k に対する x 座標と y 座標における精度行列である。本研究では、位置データを使用するため、精度行列 Λ_x , Λ_y および対数尤度 L_x , L_y を算出し、対数尤度比 R を求める。

4 実験

実験は、神戸大学大学院農学研究科附属食資源教育研究センターの 80 m x 140 m の広さの放牧場で実施した。日時は 2023 年 4 月 29 日から 5 月 2 日のうちの 4 日間である。および 2023 年 12 月 16 日から 12 月 19 日の 4 日間。首輪型デバイスを装着した放牧牛 1 頭につき、位置情報（経度および緯度）データを 1 秒間隔で取得し、単位を m に変換する。個体数 26 である ($n = 26$)。データの欠損については IterativeImputer により補完し、50 秒ごとにサンプリングすることでデータの平滑化を行う。また 12 月 16 日から 12 月 19 日の期間中、12 月 18 日 8 時 45 分に個体番号 7, 15, 25 間に異常接近およびマウンティングが発生している。そこで、この期間のデータを実験群とする。このとき、牛の放牧のスケジュールは定常であるため、データを 4 時間ごと（計 24）に分割し、同じ時間帯に関して比較を行う。このデータを入力として、Graphical Lasso を適用し、個体間の位置の近接性を表す精度行列 Λ を推定する。

5 実験結果

図 1 に、左から、時間帯 8 : 00 から 12 : 00 間の 12 月 16 日から 12 月 19 日および通期における精度行列 Λ を可視化した。 x 座標の精度行列は赤、 y 座標の精度行列は青で示され、色が濃いほど値が大きい、すなわち個体間の関係が強いことを示す。また、灰色は要素が 0、すなわち個体間に関係がないことを示す。図 1 では、大部分で位置の同期がみられないことが強調して記述されていることから、解釈性に優れていることがわかる。また、12 月 18 日（インタラクション発生日）に注目すると、個体 7-15, 7-25, 15-25 間の要素が濃い紫色、つまり位置の相関が大きいことがわかる。

図 1 に、全期間における個体 7-15-25 間の精度行列の要素の値の推移を示す。 25-7 間は青、 7-15 間は茶、 15-25 間は緑で示され、細線は x 座標または y 座

標の精度行列の要素、太線はそれらの積を表す。インタラクションが発生した時間帯（1218 から 0812）でいずれも大きな値であるほか、その前後でも値が大きいことがわかる。

最後に、実験群と対照群における前期・後期の対数尤度比の比較に関する結果を確認する。実験群では、異種間の対数尤度比が通期との対数尤度比を下回った割合が 0.708、同種間では 0.563 と、0.145 の差があった。対して、対照群では、異種間では 0.667、同種間では 0.625 と、0.042 の差があった。実験群と対照群で有意な差があることから、Graphical Lasso による空間近接性ネットワークの推定はインタラクションの有無と関係を有するという仮説について検証できた。

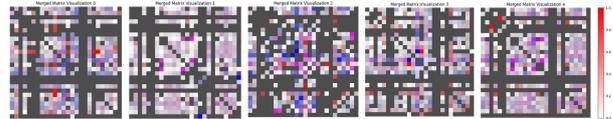


図 1 0800-1200 間の精度行列の比較

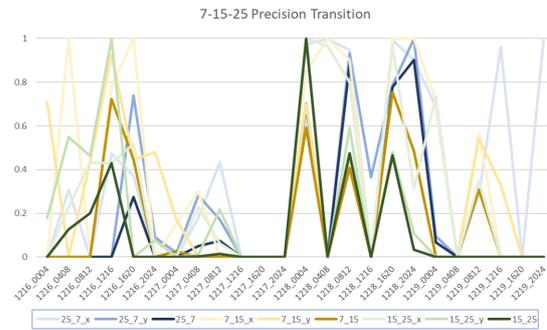


図 2 個体 7-15-25 の精度行列の値の推移

6 まとめ

各牛の時系列位置データから、疎構造学習手法である Graphical Lasso を適用することで牛同士の空間近接性ネットワークを推定し、解釈性に富むネットワークの構築することができた。また、今回の 2 群検定により、インタラクションの検出に関して一定の示唆が得られた。

謝辞

本研究の一部は JSPS 科研費 21H04914 の助成による。

参考文献

- [1] Darren P. Croft et al. 動物の社会ネットワーク分析入門。島田 将喜訳。東海大学出版部 (2019)。原書名 Exploring Animal Social Networks, (2008)。
- [2] Damien R. Farine, and Hal Whitehead. “Constructing, conducting and interpreting animal social network analysis.”, Journal of animal ecology, Vol. 84, No. 5, pp. 1144-1163 (2015)。
- [3] 佐藤 秀介ほか。動物行動図説一家畜・伴侶動物・展示動物一。朝倉書店 (2011)。
- [4] 井出 剛, 杉山 将。異常検知と変化検知。講談社。(2017)。
- [5] Jerome H. Friedman et al. “Sparse inverse covariance estimation with the graphical lasso.”, Biostatistics, Vol. 9, No. 3, pp. 432-441 (2008)。
- [6] Coloiner J, et al. “Graph Lasso-Based Test for Evaluating Functional Brain Connectivity in Sickle Cell Disease. Brain Connect.”, Brain Connect Vol. 7, No. 7, pp. 443-453 (2017)。