

放牧牛のネットワーク分析における疎構造学習の適用 Application of Sparse Structure Learning in Network Analysis of Cattle Movement

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. その一例として, 家畜の状態の推定や異常検知の自動化が挙げられる. 本研究では, 放牧牛から取得したデータをもとに放牧牛の状態の把握に取り組む. 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.

動物の社会性を分析する際には, 対象の動物や検証する仮説に基づき分析者が指標を設定してネットワークを構築することが一般的である. しかし, このような方法は, 構築されるネットワークが分析者の定める指標に強く依存し, 分析から得られる結論に大きな影響を与える点が問題となっている. 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し, 分析者の恣意性を排除したアプローチを目指す. 具体的には, ネットワークの疎構造に着目する, 疎構造学習を導入する. 生化学などの特定の分野において応用が進みつつある疎構造学習について, 本研究では, 放牧牛同士の関係の解析へ新たに応用する.

2 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より具体的に定義される二個体の関係を表す指標である. 例として, ここでは各個体の空間近接性をアソシエーションとしたネットワークの構築について述べる. 動物の個体数を M , サンプル長を T とする. ある時刻において, 個体 i と j が近接した状態であると判定する閾値 r を設定する. 個体 i と j が距離 r 内に存在するサンプル数を $T_{i,j,r}$ とすると, アソシエーション指標 $A_{i,j,r}$ は式 (1) のように表される. ここでは, 例として Simple ratio と呼ばれるアソシエーション指標を示す.

$$A_{i,j,r} = \frac{T_{i,j,r}}{T} \quad (1)$$

$A_{i,j,r}$ は各辺の重み, すなわちネットワークを定める隣接行列における, i 行 j 列の要素に相当する ($i, j = 1, \dots, M$). 閾値やアソシエーション指標は, 対象とする動物や検証する仮説に応じて設定することが一般

的である. 一方で, このような, いわば経験的な構築においては, 異なるデータを対象とした際のネットワーク一貫性, および仮説の検証に適切な構築の妥当性の担保が懸念される. そのため, 分析者の恣意性に依らず, データに対して一意にネットワークが定まる構築が望まれている. また, ネットワークの構築の際に慣例的に行われることとして, 特定の辺の削除 (フィルタリング) が挙げられる. このように, ネットワークが疎になるように整形することで, ネットワークの解釈性やロバスト性の向上が期待されている.

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, 「速度が同期する個体同士はアソシエーションが強い」という仮説を立て, これをデータ駆動的に扱うことで, 放牧牛に対するネットワークを構築する. 具体的には, 比較的取得が容易と考えられる放牧牛の各個体の時系列速度データを入力として, 各個体同士の速度の相関関係を表すグラフを疎構造学習により推定することを提案し, 放牧牛の社会的関係を記述する.

3.1 疎構造学習手法 “Graphical Lasso”

構造学習 [4] とは, 独立同分布に従う多変量のサンプルが与えられた場合, サンプルの変数間の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仮定した場合の構造学習を説明する. 多変量正規分布は, 式 (2) のように表せる.

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

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疎構造学習と呼ばれる手法のひとつである. 疎構造学習は, 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. これにより, 解釈性に富み, 且つノイズに対してロバストなネットワークの推定が可能となる.

Graphical Lasso は, 精度行列 Λ の推定にあたり, 損失関数に L_1 正則化項を導入する最尤推定を行う. 損失関数 $l_\alpha(\Lambda)$ を式 (3) に示す.

$$l_\alpha(\Lambda) = -\log |\Lambda| + \text{tr}(\Lambda S) + \alpha \sum_i \sum_j |\Lambda_{i,j}| \quad (3)$$

S は多変量からなる標本 (入力データ) 分散共分散行列, α はペナルティ係数を表す.

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

2) 神戸大学 農学研究科食資源教育研究センター. Food Resources Education and Research Center, Kobe University

本手法は、遺伝子発現ネットワークの推定 [6] などでも有用性を示している。これらの応用と同様に、放牧牛の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ロバスト性が重要であることから、本研究でも疎構造学習を応用する。

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×140 m の広さの放牧場で実施した。日時は 2022 年 4 月 1 日。首輪型デバイスを装着した放牧牛 1 頭につき、3 軸速度データを 5 秒間隔で取得し、3 軸合成加速度に加工する。サンプル長 $T = 17280$ 、個体数 $M = 13$ の時系列速度データを作成する。

このデータを入力として、Graphical Lasso、および比較手法である Ledoit-wolf を適用し、個体間の速度の関係性ネットワークを推定する。解釈性については、得られた精度行列 Λ を定性的に評価する。ロバスト性については、ペナルティ係数による交差検証スコアの変化を確認する。

さらに、求めた個体間の速度の関係性ネットワークが、アソシエーションを推定する手段として適切であることを確かめるため、各牛から取得される時系列速度以外の種類のデータと照合する。ここでは、各牛から取得された GPS データをもとに、サンプル数 40 の時系列空間位置情報を作成する。また、この時系列空間位置情報を入力として、アソシエーション指標として Sample ratio $A_{i,j,r}(i, j = 1, \dots, M)$ を用いて空間近接性ネットワークを作成する。ここで、閾値 r として 11m から 55m の 5 通りを用意し、ネットワークを 5 通り作成する。3 種類の精度行列 Λ 、すなわち速度の関係性ネットワーク (標本, Ledoit-wolf, Graphical Lasso) に対して、空間近接性ネットワークとの距離を求めることで、推定した速度の関係性と位置の関係性の類似度を評価する。ネットワーク同士の距離 d は、式 (4) で示す $M \times M - 1$ 次元の対数平方ユークリッド距離により算出する。閾値を変動した 5 通りの空間近接性ネットワークそれぞれに対して d_r を求め相加平均をとる。

$$d_r = \log \left(\sum_i \sum_{j:i \neq j}^{M-1} (|A_{i,j}| - A_{i,j,r})^2 \right) \quad (4)$$

5 実験結果

左から標本, Ledoit-wolf, Graphical Lasso (提案手法) により求めた精度行列 Λ を図 1 に示す。色が濃いほど絶対値が大きい、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 より、 $i = 1$ の個体は他のほとんどの個体との速度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。

Graphical Lasso の交差検証スコアを図 2 に示す。横軸はペナルティ係数、縦軸は対数尤度を表す。図 2 より、交差検証スコアを最良にするペナルティ係数を選択しネットワークを推定することで、ノイズに対するロバスト性が確認できた。

さらに、各牛の GPS から取得した空間的位置をタイムラプスでプロットした図を図 3 に示す。図 3 の左下部に注目すると、 $i = 1$ の個体は空間的にも孤立していることがわかる。図 1 で示した結果と照合すると、速度

データから空間的に孤立した個体を強調して記述することができた。

また、精度行列 Λ と空間近接性ネットワークとの距離 d について、標本は 8.112, Ledoit-wolf は 8.101, Graphical Lasso は 7.845 であった。速度から推定したネットワークの中でも、Graphical Lasso が最も空間近接性ネットワークと類似していた。

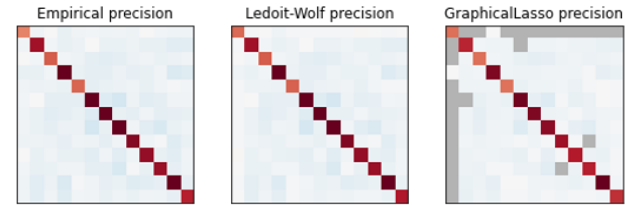


図 1 精度行列の比較

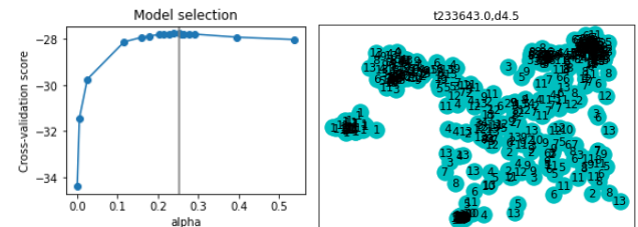


図 2 交差検証スコア

図 3 各牛の位置の分布

6 まとめ

各牛の時系列速度データから、疎構造学習手法である Graphical Lasso を適用することで牛同士の関係を推定した。疎なネットワークを推定することで、解釈性に富みロバスト性に長けたネットワークを得ることができた。個体間の真の関係は観測が困難であるため、推定の精度を評価することは難しい。しかし、速度の関係ネットワークの推定により、空間的に孤立する個体を強調して記述できたことから、真の関係の推定について一定の示唆を得られた。今後は、グラフの変化から闘争やマウンティングなど特定のインタラクションの検出など、より高次元情報の検出に取り組む。

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] Patricia Menéndez et al. “Gene regulatory networks from multifactorial perturbations using Graphical Lasso: application to the DREAM4 challenge.”, PloS one, Vol. 5, No. 12, pp. e14147 (2010)。