

# Avaliação da Escolha de Hiperparâmetros para Modelos Gráficos de Cópula Bayesianos

Renato Rodrigues Silva<sup>a,1,\*</sup>, Márcio Augusto Ferreira Rodrigues<sup>a</sup>, Everton Batista da Rocha<sup>a</sup>, Sandro Rogério Rodrigues Batista<sup>b,c</sup>

<sup>a</sup>Federal University of Goiás, Institute of Mathematics and Statistics, Campus Samambaia, CP, Goiânia, 74001-970

<sup>b</sup>Federal University of Goiás, Faculty of Medicine, 235 c/ 1a. s/n - S. Universitário, Goiânia, 74605-020

<sup>c</sup>University of Brasília, Faculty of Medicine, Campus Universitario Darcy Ribeiro Icc Sul, Brasília, 70910-900

---

## Abstract

This is the abstract. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum augue turpis, dictum non malesuada a, volutpat eget velit. Nam placerat turpis purus, eu tristique ex tincidunt et. Mauris sed augue eget turpis ultrices tincidunt. Sed et mi in leo porta egestas. Aliquam non laoreet velit. Nunc quis ex vitae eros aliquet auctor nec ac libero. Duis laoreet sapien eu mi luctus, in bibendum leo molestie. Sed hendrerit diam diam, ac dapibus nisl volutpat vitae. Aliquam bibendum varius libero, eu efficitur justo rutrum at. Sed at tempus elit.

**Keywords:** keyword1, keyword2

---

## 1. Introdução

Análise de rede tem sido utilizada em diversas áreas do conhecimento, tais como: psicologia ambiental, psicopatologia, psicologia da personalidade, saúde pública entre outras<sup>1,2,3,4</sup>. Com a pressuposição de que o fenômeno a ser estudado seja um sistema complexo, essa análise proporciona que a representação das variáveis que compõem o sistema seja feita através de um grafo não direcionado<sup>5</sup>. Além disso, é possível modelar as relações entre as variáveis e identificar padrões de agrupamentos<sup>6</sup>. Por exemplo, em uma aplicação de saúde pública, o fenômeno de multimorbidade crônica em uma população pode ser considerado um fenômeno complexo, cada morbidade crônica (diabetes, pressão arterial, colesterol, depressão, etc ...) pode ser representada por um nodo e a relação entre essas doenças crônicas por arestas de um grafo direcionado. A análise de rede é um método estatístico baseadas em uma classe de modelos probabilísticos denominados Campo Aleatório de Markov (CAM)<sup>5</sup>. Para variáveis aleatórias contínuas, tem-se usado os modelos Campo Aleatório de Markov Gaussianos<sup>7</sup>. O processo de estimação de parâmetros desse modelo é bastante desafiador, pois o número de parâmetros é dado por:  $\frac{k*(k+1)}{2}$ , em que  $k$  é o número de nodos de um grafo não direcionado. Para lidar com o problema de superajustamento, (author?)<sup>8</sup>, propuseram a estimação da correlação parcial com penalização LASSO. Para a modelagem de dados binários, (author?)<sup>9</sup> propuseram um modelo estatístico para a representação da rede baseado no modelo Ising. Dada a equivalência entre o modelo Ising e o modelo de regressão logística, foi proposto o uso da regressão logística com penalização LASSO<sup>10</sup> para lidar com o problema de superajustamento. (author?)<sup>11</sup> apresentaram o

---

\*Corresponding author

Email addresses: [renato.rrsilva@ufg.br](mailto:renato.rrsilva@ufg.br) (Renato Rodrigues Silva), [marcioaugusto@ufg.br](mailto:marcioaugusto@ufg.br) (Márcio Augusto Ferreira Rodrigues), [evertonbatista@ufg.br](mailto:evertonbatista@ufg.br) (Everton Batista da Rocha), [sandrorbatista@gmail.com](mailto:sandrorbatista@gmail.com) (Sandro Rogério Rodrigues Batista)

<sup>1</sup>This is the first author footnote.

modelo gráfico misto, o qual é um caso especial da distribuição conjunta de um CAM que permite combinar um conjunto arbitrário de distribuições condicionais que sejam membros da família exponencial. Mais recentemente, (author?)<sup>12</sup> apresentou um pacote estatístico com diversos modelos gráficos Bayesianos, dentre eles destaca-se o modelo gráfico de cópula Gaussiano<sup>13</sup>, pois este permite a modelagem tanto de dados contínuos, binários, ordinais e mistos, independente dessa variável pertencer ou não a família exponencial. Segundo (author?)<sup>12</sup>, a vantagem de usar os modelos Bayesianos seria a quantificação da incerteza do espaço de busca da seleção dos modelos candidatos e a quantificação da precisão das estimativas. No entanto, uma questão central ainda a ser respondida é qual o impacto da escolha dos hiperparâmetros na distribuição a posteriori do modelo gráfico de cópula Gaussiano. Dessa forma, o presente estudo objetiva avaliar diferentes escolhas de hiperparâmetros para a distribuição a priori e verificar qual impacto dessas escolhas na precisão das estimativas de medidas de centralidade da rede. Com esta finalidade, serão utilizados os dados de morbididades crônicas obtidos na segunda onda do Estudo Longitudinal da Saúde dos Idosos Brasileiros (ELSI-Brasil)<sup>14</sup>

## 2. Material e Método

O conjunto de dados utilizado foi obtido a partir de registros de autorelatos sobre morbididades crônicas dos participantes da segunda onda do estudo longitudinal da saúde dos idosos brasileiros, ELSI<sup>14</sup>. Na etapa de processamento de dados, retirou-se os dados faltantes, totalizando N observações.

### 2.1. Campo Aleatório de Markov

CAM pode ser definido como um conjunto de variáveis aleatórias

### 2.2. Campo Aleatório de Markov Gaussiano

### 2.3. Modelos Gráficos de Cópula Gaussiano

## 3. Resultados

### 3.1. Avaliação do parâmetro de probabilidade de inclusão

## References

- [1] M. V. Zwicker, H. U. Nohlen, J. Dalege, G.-J. M. Gruter, F. van Harreveld, [Applying an attitude network approach to consumer behaviour towards plastic](#), Journal of Environmental Psychology 69 (2020) 101433. doi:<https://doi.org/10.1016/j.jenvp.2020.101433>. URL <https://www.sciencedirect.com/science/article/pii/S0272494419308783>
- [2] D. Borsboom, [A network theory of mental disorders](#), World Psychiatry 16 (1) (2017) 5–13. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/wps.20375>, doi:<https://doi.org/10.1002/wps.20375>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/wps.20375>
- [3] G. Costantini, S. Epskamp, D. Borsboom, M. Perugini, R. Møttus, L. J. Waldorp, A. O. Cramer, [State of the art personality research: A tutorial on network analysis of personality data in r](#), Journal of Research in Personality 54 (2015) 13–29, r Special Issue. doi:<https://doi.org/10.1016/j.jrp.2014.07.003>. URL <https://www.sciencedirect.com/science/article/pii/S0092656614000701>
- [4] G. H. Soares, F. M. R. Bado, C. F. Tenani, P. H. Ribeiro Santiago, L. M. Jamieson, F. L. Mialhe, [A psychometric network perspective to oral health literacy: Examining the replicability of network properties across the general community and older adults from brazil](#), Journal of Public Health Dentistry 82 (3) (2022) 321–329. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/jphd.12539>, doi:<https://doi.org/10.1111/jphd.12539>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/jphd.12539>
- [5] K. P. Murphy, Machine learning: a probabilistic perspective, MIT Press, Cambridge, 2012.
- [6] S. Epskamp, A.-M. Isvoranu, [New trends in network modeling of psychopathology](#), World Psychiatry 21 (3) (2022) 463–464. doi:<https://doi.org/10.1002/wps.21017>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/wps.21017>
- [7] H. Rue, L. Held, Machine learning: a probabilistic perspective, Chapman and Hall/CRC, New York, 2005. doi:<https://doi.org/10.1201/9780203492024>.
- [8] S. Epskamp, D. Borsboom, E. I. Fried, [A tutorial on regularized partial correlation networks.](#), Psychological methods 23 (4) (2018) 617–634. doi:[10.18637/jss.v048.i04](https://doi.org/10.18637/jss.v048.i04). URL <https://doi.org/10.1037/met0000167>
- [9] C. D. vanBorkulo, D. Borsboom, S. Epskamp, T. F. Blanken, L. Boschloo, R. A. Schoevers, L. J. Waldorp, [A new method for constructing networks from binary data](#), Scientific Reports 4 (2014) 5918. doi:<https://doi.org/10.1038/srep05918>.

- [10] J. Friedman, R. Tibshirani, T. Hastie, Regularization paths for generalized linear models via coordinate descent, *Journal of Statistical Software* 33 (1) (2010) 1–22. doi:[10.18637/jss.v033.i01](https://doi.org/10.18637/jss.v033.i01).
- [11] J. M. B. Haslbeck, L. J. Waldorp, *mgm: Estimating time-varying mixed graphical models in high-dimensional data*, *Journal of Statistical Software* 93 (8) (2020) 1–46. doi:[10.18637/jss.v093.i08](https://doi.org/10.18637/jss.v093.i08).  
URL <https://www.jstatsoft.org/index.php/jss/article/view/v093i08>
- [12] K. B. S. Huth, J. de Ron, A. E. Goudriaan, J. Luigjes, R. Mohammadi, R. J. van Holst, E.-J. Wagenmakers, M. Marsman, *Bayesian analysis of cross-sectional networks: A tutorial in r and jasp*, *Advances in Methods and Practices in Psychological Science* 6 (4) (2023) 25152459231193334. arXiv:<https://doi.org/10.1177/25152459231193334>, doi:[10.1177/25152459231193334](https://doi.org/10.1177/25152459231193334).  
URL <https://doi.org/10.1177/25152459231193334>
- [13] R. Mohammadi, E. C. Wit, *Bdgraph: An r package for bayesian structure learning in graphical models*, *Journal of Statistical Software* 89 (3) (2019) 1–30. doi:[10.18637/jss.v089.i03](https://doi.org/10.18637/jss.v089.i03).  
URL <https://www.jstatsoft.org/index.php/jss/article/view/v089i03>
- [14] M. F. Lima-Costa, F. B. de Andrade, P. R. B. d. Souza, A. L. Neri, Y. A. d. O. Duarte, E. Castro-Costa, C. de Oliveira, *The Brazilian Longitudinal Study of Aging (ELSI-Brazil): Objectives and Design*, *American Journal of Epidemiology* 187 (7) (2018) 1345–1353. arXiv:<https://academic.oup.com/aje/article-pdf/187/7/1345/25076724/kwx387.pdf>, doi:[10.1093/aje/kwx387](https://doi.org/10.1093/aje/kwx387).  
URL <https://doi.org/10.1093/aje/kwx387>