

# Research on Regularization Model in Complex Network Structure

Ruixue Zong, Xiaomin Zhong

School of Mathematics and Information Science, Guangxi University, Nanning Guangxi  
Email: 1531874485@qq.com, 846345642@qq.com

Received: August 12<sup>th</sup>, 2019; accepted: August 27<sup>th</sup>, 2019; published: September 3<sup>rd</sup>, 2019

---

## Abstract

Based on the complex network structure under the Gaussian model, the penalty term is considered as the seamless- $L_0$  penalty, and a new regularization model is proposed. The Minorize-Maximization algorithm was used to construct the Minorization function of the proposed regularization model, and then the Maxdet algorithm was used to solve the model. The model simultaneously performs model selection and parameter estimation. And the result is sparse. Numerical simulation studies and examples of gene regulatory networks show that the model has good parameter estimation and model selection effects.

## Keywords

Complex Network, Regularization Model, Maxdet Algorithm, Model Selection

---

# 正则化模型在复杂网络结构中的研究

宗瑞雪, 钟小敏

广西大学, 数学与信息科学学院, 广西 南宁  
Email: 1531874485@qq.com, 846345642@qq.com

收稿日期: 2019年8月12日; 录用日期: 2019年8月27日; 发布日期: 2019年9月3日

---

## 摘要

研究基于高斯图模型下的复杂网络结构, 考虑惩罚项为The seamless- $L_0$ 惩罚, 提出了新的正则化模型。利用Minorize-Maximization算法, 构造所提正则化模型的Minorization函数, 再对该函数利用Maxdet算法求解模型。该模型同时进行模型选择和参数估计, 并且估计结果具有稀疏性。数值模拟研究和基因调控网络的实例说明该模型有良好的参数估计和模型选择效果。

## 关键词

复杂网络, 正则化模型, Maxdet算法, 模型选择

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

复杂网络是由数据庞大的节点及节点之间错综复杂的关系所构成的网络结构。复杂网络普遍存在于现实生活中, 如社会关系网络[1], 交通运输网络, 基因网络等。因此, 对复杂网络的研究成为近年来的热点问题[2]。

本文基于高斯图[3]模型(Gaussian Graphical Model, GGM)对复杂网络结构进行研究。其数学模型如下, 假设  $\mathbf{X} = (\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p)})$  服从  $p$  维正态分布  $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ,  $\boldsymbol{\mu}$  为均值,  $\boldsymbol{\Sigma}$  为协方差矩阵。利用  $\mathbf{X}$  的样本  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$  估计  $\mathbf{A} = \boldsymbol{\Sigma}^{-1}$  [3]。当  $a_{ij} = 0$  时, 表示  $\mathbf{X}^{(i)}$  与  $\mathbf{X}^{(j)}$  条件独立。高斯图模型中, 随机向量用无向图表示  $G = VE$  [4]。其中,  $V$  代表节点, 每个节点分别对应随机向量的一个分量。  $E$  代表边, 两个节点之间无边, 当且仅当这两个节点对应的随机变量条件独立, 即  $a_{ij} = 0$ 。因此, 判断协方差矩阵的逆  $\mathbf{A}$  中的每个元素是否为 0 即是判断网络的结构。

在高斯图模型中, 标准的模型选择方法是 Greedy stepwise。该方法下参数估计结果是基于被选择的模型[5]。在每一步中, 边的选择或删除通常通过在某个水平  $\alpha$  的假设检验来完成。这种方法的缺点是没有正确地考虑到多重比较。另一个缺点是该方法在计算上具有复杂性[5]。为了解决以上问题, Drton 和 Perlman [6]提出产生保守同步的  $1-\alpha$  置信区间, 并在每一步中利用这些置信区间进行模型选择。该方法的提出是基于对渐近性的考虑。Meinshausen 和 Buhlman [7]为协方差选择提出具有较强计算能力的方法, 可用于更庞大的高斯图。该方法为图中的每个节点进行领域选择, 并结合选择结果来研究高斯浓度图模型的结构。Meinshausen 和 Buhlman 说明了他们的方法对于稀疏高维图模型具有一致性。上述的这些方法, 在模型选择和参数估计上都是分开的。并且, 参数的估计是基于被选择的模型。Breiman [8]也证明了这些方法在估计上具有不稳定性, 即数据发生小的变动时, 会导致估计结果有很大的不同。

本文提出的正则化模型在高斯图中可同时进行模型选择和参数估计, 并使得估计结果具有稀疏性。正则子为 The seamless- $L_0$  (SELO) [9]。该正则化模型对所估计矩阵非对角线上的元素进行 SELO 惩罚。该思想类似于 SELO 惩罚在线性回归模型中的应用。本文使用 MM 算法和 Maxdet 算法求解该模型。最后, 通过数值模拟和实例分析说明本文提出的正则化模型有良好的模型选择和参数估计效果。

## 2. 正则化模型

设  $\mathbf{X} = (\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p)})$  服从  $p$  维正态分布  $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ,  $\boldsymbol{\mu}$  为均值,  $\boldsymbol{\Sigma}$  为协方差矩阵。设  $\mathbf{A} = (a_{ij})_{1 \leq i, j \leq p} = \boldsymbol{\Sigma}^{-1}$  为协方差矩阵的逆。基于随机变量  $\mathbf{X}$  样本观测值  $X_1, X_2, \dots, X_n$  的对数似然函数是

$$\frac{n}{2} \log |\mathbf{A}| - \frac{1}{2} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})' \mathbf{A} (\mathbf{X}_i - \boldsymbol{\mu}) \quad (1)$$

当不依赖于常数  $A$  和  $\mu$  时, 设  $(\mu, \Sigma)$  的最大似然估计为  $(\bar{X}, \bar{C})$ 。其中

$$C = \frac{1}{n} \sum_{i=1}^n (\mathbf{X}_i - \bar{X})' (\mathbf{X}_i - \bar{X})$$

通常利用样本协方差矩阵  $S = n\bar{C}/(n-1)$  进行协方差矩阵的估计。因此,  $A = \Sigma^{-1}$  可利用  $S^{-1}$  或者  $\bar{C}^{-1}$  进行估计。但是, 当变量的维数很大时, 矩阵  $S^{-1}$  稳定性差。并且, 当  $n \geq p$  时, 矩阵  $S^{-1}$  是正定矩阵, 矩阵中的元素不包含零项。所以, 以此方法进行的估计结果是不具有稀疏性的图模型。

因此, 通过引入  $L_1$  正则子进行模型选择和参数估计。本文选择的  $L_1$  正则子为 SEL0 惩罚, 改惩罚形式是对  $L_0$  惩罚的形式的模仿[10]。

但当变量维数很大时, 该方法在计算上具有复杂性, 并且具有不连续性和复杂性[9]。为解决上述问题, Dicker L.和 Huang B.等提出 SEL0 惩罚。该惩罚形式为

$$p_{\text{SEL0}}(\beta_j) = \frac{\lambda}{\log(2)} \left( \frac{|\beta_j|}{|\beta_j| + \tau} + 1 \right) \quad (2)$$

其中,  $\tau > 0$  是调节参数, 当  $\tau$  很小时,  $p_{\text{SEL0}}(\beta_j) \approx \lambda I\{\beta_j \neq 0\}$ , 并且, 当  $\tau = 0.01$  时, 该方法呈现更好的效果。

因此, 为了得到具备稀疏性的图模型结构, 并更加准确的估计  $A = \Sigma^{-1}$ 。本文结合 The seamless- $L_0$  惩罚(SEL0)和  $-\log|A| + \frac{1}{n} \sum_{i=1}^n (\mathbf{X}_i - \mu)' A (\mathbf{X}_i - \mu)$  寻找  $(\hat{\mu}, \hat{A})$  的最大值。即

$$\log|A| - \frac{1}{n} \sum_{i=1}^n (\mathbf{X}_i - \mu)' A (\mathbf{X}_i - \mu) \quad \text{st} \quad \sum_{i \neq j} \frac{1}{\log(2)} \log \left( \frac{|a_{ij}|}{|a_{ij}| + \gamma} + 1 \right) \leq \lambda \quad (3)$$

其中,  $\lambda \geq 0$  是调节参数。

因为对观测值作中心化处理。所以  $\hat{\mu} = 0$ 。因此, (1)式可写为

$$\log|A| - \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i' A \mathbf{X}_i \quad \text{st} \quad \sum_{i \neq j} \frac{1}{\log(2)} \log \left( \frac{|a_{ij}|}{|a_{ij}| + \gamma} + 1 \right) \leq \lambda \quad (4)$$

进一步地, 我们可以将(2)式写为

$$\log|A| - \text{tr}(A\bar{C}) \quad \text{st} \quad \sum_{i \neq j} \frac{1}{\log(2)} \log \left( \frac{|a_{ij}|}{|a_{ij}| + \gamma} + 1 \right) \leq \lambda \quad (5)$$

由于(3)式中的目标函数和可行区域都是凸的, 因此, 可以利用拉赫朗日形式将(3)式等价得写为

$$L(A) = \log|A| - \text{tr}(A\bar{C}) - \lambda \sum_{i \neq j} \frac{1}{\log(2)} \log \left( \frac{|a_{ij}|}{|a_{ij}| + \gamma} + 1 \right) \quad (6)$$

因此, 最大化(4)式中的  $L(A)$ , 即可得到图模型的协方差矩阵的逆的估计。

### 3. Maxdet 算法与参数选择

本文利用 Maxdet 算法[11]求解目标函数(6)。但由于本文中目标函数  $L(A)$  的 SEL0 惩罚项是  $\log$  和  $L_1$  惩罚函数的复合形式, 因此, 目标函数较难直接进行优化, 需要利用 Minorize-Maximization (MM) [12] 算法, 构造  $L(A)$  的 Minorization 函数  $H(a|a^n)$ , 其中,  $a^n$  是 MM 算法的第  $n$  部估计, 也是参数  $a$  的一个定值。该函数逼近于原目标函数  $L(A)$  且易于优化。因此, 本节先构造出函数  $H(a|a^n)$ , 再通过构造函数

$H(a|a^n)$  利用 Maxdet 算法, 得到逼近于  $L(\mathbf{A})$  的估计值。最后使用 BIC 方法进行参数估计。

首先构造  $L(\mathbf{A})$  的 Minorization 函数  $H(a|a^n)$ 。

若满足以下条件, 则目标函数  $L(\mathbf{A})$  可用构造函数  $H(a|a^n)$  代替。

$$c_1. L(\mathbf{A}) \geq H(a|a^n)$$

$$c_2. L(\mathbf{A}^n) = H(a|a^n)$$

由于当  $x > 0$ ,  $-\log(1+x) \geq -x$  故

$$\frac{\lambda}{\log(2)} \sum_{i \neq j} \log \left( \frac{|a_{ij}^n|}{|a_{ij}^n| + \gamma} + 1 \right) - \frac{\lambda}{\log(2)} \sum_{i \neq j} \log \left( \frac{|a_{ij}|}{|a_{ij}| + \gamma} + 1 \right) \geq \frac{\lambda}{\log(2)} \sum_{i \neq j} \left( \frac{\frac{|a_{ij}|}{|a_{ij}| + \gamma} + 1}{\frac{|a_{ij}^n|}{|a_{ij}^n| + \gamma} + 1} \right)$$

不等式两边同时加上

$$\log |\mathbf{A}| - \text{tr}(\mathbf{A}\bar{\mathbf{C}}) - \frac{\lambda}{\log(2)} \sum_{i \neq j} \log \left( \frac{|a_{ij}^n|}{|a_{ij}^n| + \gamma} + 1 \right)$$

得

$$L(\mathbf{A}) \geq \log |\mathbf{A}| - \text{tr}(\mathbf{A}\bar{\mathbf{C}}) - \frac{\lambda}{\log(2)} \left( \frac{|a_{ij}^n|}{|a_{ij}^n| + \gamma} + 1 \right) - \frac{\lambda}{\log(2)} \sum_{i \neq j} \left( \frac{\frac{|a_{ij}|}{|a_{ij}| + \gamma} + 1}{\frac{|a_{ij}^n|}{|a_{ij}^n| + \gamma} + 1} \right)$$

令不等式右边为  $H(a|a^n)$ , 即

$$H(a|a^n) = \log |\mathbf{A}| - \text{tr}(\mathbf{A}\bar{\mathbf{C}}) - \frac{\lambda}{\log(2)} \left( \frac{|a_{ij}^n|}{|a_{ij}^n| + \gamma} + 1 \right) - \frac{\lambda}{\log(2)} \sum_{i \neq j} \left( \frac{\frac{|a_{ij}|}{|a_{ij}| + \gamma} + 1}{\frac{|a_{ij}^n|}{|a_{ij}^n| + \gamma} + 1} \right) \quad (7)$$

故  $L(\mathbf{A}) \geq H(a|a^n)$ , 条件  $c_1$  满足。

又因

$$H(a^n|a^n) = \log |\mathbf{A}| - \text{tr}(\mathbf{A}\bar{\mathbf{C}}) - \frac{\lambda}{\log(2)} \left( \frac{|a_{ij}^n|}{|a_{ij}^n| + \gamma} + 1 \right) - \frac{\lambda}{\log(2)} \sum_{i \neq j} \left( \frac{\frac{|a_{ij}^n|}{|a_{ij}^n| + \gamma} + 1}{\frac{|a_{ij}^n|}{|a_{ij}^n| + \gamma} + 1} \right)$$

即

$$H(a^n|a^n) = \log |\mathbf{A}| - \text{tr}(\mathbf{A}\bar{\mathbf{C}}) - \frac{\lambda}{\log(2)} \left( \frac{|a_{ij}^n|}{|a_{ij}^n| + \gamma} + 1 \right)$$

故  $H(a^n | a^n) = L(A)$ , 条件  $c_2$  满足。

因此,  $L(A) \geq H(a | a^n)$ , 且  $H(a^n | a^n) = L(A)$  故  $H(a | a^n)$  满足  $c_1, c_2$  两个条件。所以  $L(A)$  的 Minorization 函数为  $H(a | a^n)$ 。

下面利用 Maxdet 算法求解式(7)。Maxdet 问题的形式如下:

$$\min_{x \in R^m} b'x - \log |G(x)|$$

故对(7)式取负号, 将其转化为求最小值问题。假设  $a_{ij}$  的符号是已知的, 设为  $s_{ij}$ 。根据 Maxdet 问题的形式, 可将转化后的(5)式写成如下形式:

$$\begin{aligned} \min 2 \sum_{i < j} c_{ij} a_{ij} + \sum_i c_{ii} a_{ii} - \log \left| \sum_i a_{ii} I^{(i)} + \sum_{i < j} c_{ij} I^{(ij)} \right| \text{st} \\ \lambda' - 2 \sum_{i < j} \left( \frac{|a_{ij}|}{|a_{ij}| + \gamma} + 1 \right) s_{ij} \geq 0, s_{ij} a_{ij} \geq 0 \end{aligned} \quad (8)$$

其中,  $I^{(ij)}$  为第  $(i, j)$  项为 1, 其余项为 0 的  $n \times n$  的矩阵,  $I^{(i)}$  为第  $(i, i)$  项为 1, 其余项为 0 的  $n \times n$  的矩阵。

关于参数  $\lambda$  的选择, 常用的方法是多重交叉验证法, 但计算上更有效的替代方案是用于模型选择和估计的 BIC 准则。因此, 本文选择如下 BIC 准则选择参数  $\lambda$  [13]:

$$\text{BIC}(\lambda) = -\log |A(\lambda)| + \text{tr} \{ A(\lambda) \bar{C} \} + \frac{\log n}{n} \sum_{i \neq j} \hat{e}_{ij}(\lambda) \text{ 其中, 如果 } a_{ij} = 0, \text{ 则 } e_{ij} = 0; \text{ 其余情 } \hat{e}_{ij} = 1。$$

下面式算法的具体步骤:

第一步: 令  $\hat{A}_{old} = \bar{C}^{-1}$  并且对任意的  $i \neq j$ ,  $\hat{A}_{old} = \bar{C}^{-1} i \neq j s_{ij} = \text{sign} \left\{ \left( \hat{A}_{old} \right)_{ij} \right\}$ 。

第二步: 令  $\hat{A}_{new}$  在正定矩阵集上解决(6)式。

第三步: 如果  $\hat{A}_{new} = \hat{A}_{old}$ , 停止迭代并令  $\hat{A} = \hat{A}_{new}$ 。否则, 对所有的  $(i, j)$  设  $\hat{A}_{new} = \hat{A}_{old}$ , 并且  $s_{ij} = -s_{ij}$ , 所以  $\hat{c}_{ij} = 0$ 。然后返回第二步。

第四步: 使用 BIC 准则选择  $\lambda$ , 并输出相应的  $A$  的估计值。

#### 4. 数值模拟

本文中的网络图模型由 R 语言 huge 包中的 huge.generator 函数生成, 所生成的图形模式为随机模型。并且在此函数的输出结果中可得到该图形模式下随机生成的样本数据及其协方差矩阵。本文设置样本量为  $n = 200$ , 维数  $p$  分别为 15, 30, 50, 100。分别比较本文提出的方法和 Meinshausen & Buhlmann (mb) 提出的方法在不同维数下参数估计和模型选择的情况。用 KL 损失量以及 FPR (false positive rate), TPR (true positive rate) 作为评价正则化模型在参数估计和模型选择上优良性的标准。其中, KL 损失量(KL loss) 为[14]:

$$\text{KL} = -\log |\hat{A}| + \text{tr}(\hat{A} \hat{\Sigma}) - \left( -\log |\Sigma^{-1}| + p \right)$$

FPR 和 TPR 为:

$\text{FPR} = \text{FP} / (\text{FP} + \text{TN})$ ;  $\text{TPR} = \text{TP} / (\text{TP} + \text{FN})$ 。其中, TPR 越大, FPR 越小说明模拟效果越好。

模拟结果如下(表 1~表 4)。

**Table 1.**  $p = 15$ **表 1.**  $p = 15$ 

model	KL loss	FPR	TPR
SELO	0.22	0.116	0.875
mb	0.35	0.137	0.855

**Table 2.**  $p = 30$ **表 2.**  $p = 30$ 

model	KL loss	FPR	TPR
SELO	0.27	0.170	0.841
mb	0.40	0.193	0.824

**Table 3.**  $p = 50$ **表 3.**  $p = 50$ 

model	KL loss	FPR	TPR
SELO	0.32	0.208	0.852
mb	0.45	0.220	0.807

**Table 4.**  $p = 100$ **表 4.**  $p = 100$ 

model	KL loss	FPR	TPR
SELO	0.39	0.243	0.811
mb	0.53	0.254	0.738

模型选择的目的是在高维数据中挑选出更具代表性的变量。表数据结果表明, 在不同维数下, SELO 方法相比于 mb 方法在模型选择和参数估计方面均是最优的。并且在  $p = 15$  时, SELO 方法下所得到的 KL loss 最小, FPR 达到最小, TPR 达到最大。随着维数的增加, 虽然 KL loss 值和 FPR 值逐渐增大, TPR 值逐渐减小, 但总体来说, gSELO 方法仍有较好的模型选择和参数估计能力。故综合以上结果, 本文提出的 gSELO 方法有良好的模型选择和参数估计能力。

## 5. 实例分析

本文将提出的正则化模型应用到基因调控网络的数据中。数据来源于 Faith 等人的论文数据[15]。该数据包含两个对象。第一个对象是 40 行, 153 列的矩阵, 该矩阵的行和列代表了 *Escherichia coli* (*E. coli*) 细菌在 40 种不同实验下, 153 种基因的基因表达水平, 基因表达的测量结果指示了 153 种基因在每个特定的实验干扰下的活动水平。该 153 种基因是转录因子, 在基因调控关系中有重要作用。第二个对象是一个邻接矩阵, 概括了生物学文献中发现的 *E. coli* 细菌的 153 种基因之间已经实际存在的调控关系。结果如图 1 所示。

下面通过统计软件, 首先根据邻接矩阵统计出生物学文献中已经发现的 153 种基因之间实际存在的调控关系的个数; 然后利用本文提出的正则化模型, 统计出在该模型下得到的生物学文献中已发现的实际调控关系及生物学文献中未被发现的调控关系的总量。步骤和结果如下:

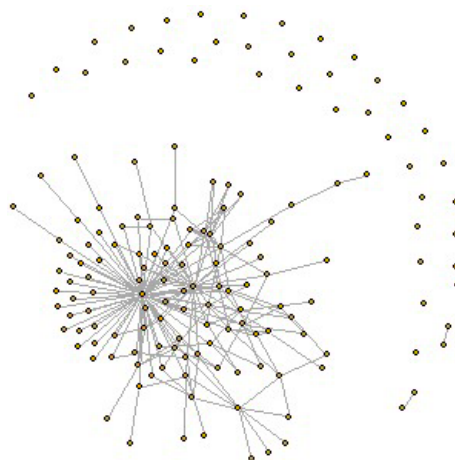
- 1) 对邻接矩阵中基因之间的实际调控关系的个数进行统计, 所得结果如下:

## IGRAPH UN-153 209

## attr: name (v/c)

结果表明: 生物学文献中已经发现的基因之间的实际调控关系共 209 对。

2) 利用第一个数据对象中 40 行, 153 列的矩阵, 根据本文提出的正则化模型, 统计出在该模型下得到的生物学文献中已发现的实际调控关系及生物学文献中未被发现的调控关系的总量。所得结果如下:



**Figure 1.** Actual regulatory relationships among 153 genes

**图 1.** 153 个基因之间实际存在的调控关系

IGRAPH 68d9b0b U-153 554--

结果表明: 根据本文提出的正则化模型共找到 554 对调控关系。

3) 最后统计在步骤 2 中利用本文提出的正则化模型得到的生物学文献中已发现的实际调控关系的个数。所得结果如下:

## IGRAPH UN-153 192--

结果表明: 通过本文提出的正则化模型在已发现的 209 对实际调控关系中共得到了 192 对调控关系。

这一结果进一步说明了该正则化模型在实际应用中的优良性。

## 6. 结论

本文基于高斯图模型进行复杂网络结构的研究, 根据  $SELO$  惩罚在线性回归模型中应用的思想, 将其应用到图模型的模型选择和参数估计中, 提出了新的正则化模型。本文使用  $Minorize-Maximization$  (MM) 算法和  $Madex$  算法求解该模型。数值模拟实验结果显示, 本文提出的正则化模型相比于  $Meinshausen \& Buhlmann$  (mb) 方法,  $KL$  loss 值和  $FPR$  值更小,  $TPR$  值更大。说明本文提出的正则化模型在参数估计和模型选择方面都有良好的效果。最后, 将该模型应用到 *E. coil* 细菌基因调控关系的实际数据当中, 结果表明, 本文提出的正则化模型对基因之间实际存在的调控关系有较好的估计效果, 说明该方法在应用中具有良好的适应性。

## 参考文献

- [1] Kolaczyk, E.D. (2009) *Statistical Analysis of Network Data*. Springer Science + Business Media, Berlin. <https://doi.org/10.1007/978-0-387-88146-1>
- [2] Hand, D.J. (2010) *Statistical Analysis of Network Data: Methods and Models* by Eric D. Kolaczyk. *International Sta-*

- tistical Review*, **78**, 134-159. [https://doi.org/10.1111/j.1751-5823.2010.00109\\_2.x](https://doi.org/10.1111/j.1751-5823.2010.00109_2.x)
- [3] Dempster, A.P. (1972) Covariance Selection. *Biometrika*, **32**, 95-108.
- [4] Koller, D. and Friedman, N. (2012) Probabilistic Graphical Models: Principles and Techniques. The MIT Press, Cambridge.
- [5] 孙建举. 高斯图模型的估计[D]: [硕士学位论文]. 上海: 上海交通大学, 2014.
- [6] Drton, M. and Perlman, M.D. (2004) Model Selection for Gaussian Concentration Graphs. *Biometrika*, **91**, 591-602. <https://doi.org/10.1093/biomet/91.3.591>
- [7] Meinshausen, N. and Bühlmann, P. (2006) High-Dimensional Graphs and Variable Selection with the Lasso. *The Annals of Statistics*, **34**, 1436-1462. <https://doi.org/10.1214/009053606000000281>
- [8] Breiman, L. (1996) Heuristics of Instability and Stabilization in Model Selection. *The Annals of Statistics*, **24**, 2350-2383. <https://doi.org/10.1214/aos/1032181158>
- [9] Dicker, L., Huang, B. and Lin, X. (2013) Variable Selection and Estimation with the Seamless- $L_0$  Penalty. *Statistica Sinica*, **23**, 929-962. <https://doi.org/10.5705/ss.2011.074>
- [10] Foster, D.P. and George, E.I. (1994) The Risk Inflation Criterion for Multiple Regression. *The Annals of Statistics*, **22**, 1947-1975. <https://doi.org/10.1214/aos/1176325766>
- [11] Yuan, M. and Lin, Y. (2007) Model Selection and Estimation in the Gaussian Graphical Model. *Biometrika*, **94**, 19-35. <https://doi.org/10.1093/biomet/asm018>
- [12] Lange, K., Hunter, D.R. and Yang, I. (2000) Optimization Transfer Using Surrogate Objective Functions. *Journal of Computational and Graphical Statistics*, **9**, 60-77. <https://doi.org/10.2307/1390605>
- [13] Friedman, J., Hastie, T. and Tibshirani, R. (2008) Sparse Inverse Covariance Estimation with the Graphical Lasso. *Biostatistics*, **9**, 432-441. <https://doi.org/10.1093/biostatistics/kxm045>
- [14] Ravikumar, P., Raskutti, G., Wainwright, M.J., et al. (2011) High-Dimensional Covariance Estimation by Minimizing L1-Penalized Log-Determinant. *Electronic Journal of Statistics*, **5**, 935-980. <https://doi.org/10.1214/11-EJS631>
- [15] Faith, J.J., Hayete, B., Thaden, J.T., et al. (2007) Large-Scale Mapping and Validation of *Escherichia coli* Transcriptional Regulation from a Compendium of Expression Profiles. *PLoS Biology*, **5**, e8. <https://doi.org/10.1371/journal.pbio.0050008>

### 知网检索的两种方式:

1. 打开知网首页: <http://cnki.net/>, 点击页面中“外文资源总库 CNKI SCHOLAR”, 跳转至: <http://scholar.cnki.net/new>, 搜索框内直接输入文章标题, 即可查询;  
或点击“高级检索”, 下拉列表框选择: [ISSN], 输入期刊 ISSN: 2324-7991, 即可查询。
2. 通过知网首页 <http://cnki.net/>顶部“旧版入口”进入知网旧版: <http://www.cnki.net/old/>, 左侧选择“国际文献总库”进入, 搜索框直接输入文章标题, 即可查询。

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [aam@hanspub.org](mailto:aam@hanspub.org)