Bayesian Variable Selection —— Stochastic Search Variable Selection

Dom Tsang

Email: 395091490@qq.com

School of Mathematical Sciences, South China Normal University

May 11, 2017

CONTENTS

- 1 前情回顾
- 2 Stochastic Search Variable Selection(随机搜索变量选择)
- 3 算法
- 4 模型选择
- 5 参考文献

Dom Tsang SCNU 2017/5/11 1/22

前情回顾

1 前情回顾

- 变量选择
- 3 算法
- 4 模型选择
- 5 参考文献

Dom Tsang SCNU 2017/5/11 2 / 22

前情回顾

00

变量选择

■ 一范数约束

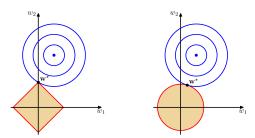


Figure 1: 正则化限制区域

机器学习中的范数规则化之(一)L0、L1 与 L2 范数 http:

//blog.csdn.net/zouxy09/article/details/24971995 逐步回归 hierarchical bayesian

<u>分层贝叶斯模型示例</u>

- 模型: $\mathbf{y}_{(P,1)} = \mathbf{H}_{(P,M)} \mathbf{x}_{(M,1)} + \mathbf{n}_{(P,1)}$ 其中, $\mathbf{n} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_p)$
- 似然函数: $f(\mathbf{y}|\mathbf{x}, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{P/2} \exp\left(-\frac{||\mathbf{y} \mathbf{H}\mathbf{x}||^2}{2\sigma^2}\right)$
- 参数先验: $f(\mathbf{x}|w,a) = [(1-w)^{n_0} \prod \delta(x_i)][w^{n_1} \prod g_a(x_i)]$ $\sigma^2 | \nu, \gamma \sim \mathcal{IG}(\frac{\nu}{2}, \frac{\gamma}{2})$
- 超参数的先验: $a|\alpha \sim \mathcal{IG}(\alpha_0, \alpha_1)$ $w \sim \mathcal{U}([0,1])$ $f(\gamma) \propto \frac{1}{\gamma}$
- 参数后验: $w|\mathbf{x} \sim \mathcal{B}\mathcal{E}(1+n_1,1+n_0)$ $a|\mathbf{x}, \alpha_0, \alpha_1 \sim \mathcal{IG}(||\mathbf{x}||_0 + \alpha_0, ||\mathbf{x}||_1 + \alpha_1)$ $f(x_i|w,a,\sigma^2,\mathbf{x}_{-i},\mathbf{y}) \propto (1-w_i)\delta(x_i) + w_i\phi_+(x_i|\mu_i,\eta_i^2)$ $\sigma^2 | \mathbf{x}, \mathbf{v} \sim \mathcal{IG}(\frac{P}{2}, \frac{||\mathbf{y} - \mathbf{H}\mathbf{x}||^2}{2})$

Dom Tsang SCNU 2017/5/11 4 / 22

算法 O 模型选择 0000 参考文献 00

hierarchical bayesian

前情回顾

0

算法

■ 初始化

← 从先验中

- 1 随机生成 **x**⁽⁰⁾
- **2** 随机生成 σ⁽⁰⁾
- 循环: for t = 1, 2, ..., do:
 - 1 随机生成 $w^{(t)}$
 - 型 随机生成 a^(t)
 - 3 For i = 1, ..., M 随机生成 $x_i^{(t)}$
 - 4 随机生成 $\sigma^{2(t)}$
 - $5 t \leftarrow t+1$

← 从后验中

Dom Tsang SCNU 2017/5/11 5/ 22

CONTENTS

- 1 前情回顾
- 2 Stochastic Search Variable Selec- β 的先验 tion(随机搜索变量选择)
- 3 算法
- 5 参考文献

- 模型
- γ 的先验
- σ^2 的先验
- β 的后验
- γ 的后验
- σ^2 的后验

Dom Tsang **SCNU** 6 / 22 2017/5/11

模型

(广义) 线性模型

$$Y|\beta, \sigma^2 \sim \mathcal{MVN}(X\beta, \sigma^2 I)$$
 (2.1)

其中, $Y \in \mathbb{R}^n$, $X \in \mathbb{R}^{n \times p}$, $\beta = (\beta_1, ..., \beta_n)^{\top}$, σ^2 是一个标量. β π σ² 都是未知参数.

- * $\mathcal{MVN}(X\beta, \sigma^2 I)$ 即是通常的正态分布 $\mathcal{N}(X\beta, \sigma^2 I)$
- * 式 (2.1) 的意思是 $Y = X\beta + n$, 其中 $\sim \mathcal{N}(0, \sigma^2 I)$

Dom Tsang **SCNU** 2017/5/11 7 / 22

β的先验

■ 使用两个不同方差的高斯分布的混合分布对 β 的每个成分 进行建模. 引入隐变量 $\gamma_i = 0$ 或 1, 可将高斯分布混合成,

β_i (混合高斯)

$$\beta_i|\gamma_i \sim (1-\gamma_i)\mathcal{N}(0,\tau_i^2) + \gamma_i\mathcal{N}(0,c_i^2\tau_i^2)$$
 (2.2)

$$P(\gamma_i = 1) = 1 - P(\gamma_i = 0) = \omega_i$$
 (2.3)

* 式 (8) 表达的模型用大家相对熟悉的表达形式即是 $P(\beta_i) = P(\gamma_i = 0)P(\beta_i|\gamma_i = 0) + P(\gamma_i = 1)P(\beta_i|\gamma_i = 1) = (1 - \omega_i)\mathcal{N}(0, \tau_i^2) = \omega_i \mathcal{N}(0, c_i^2 \tau_i^2)$

β 的先验

β (多元高斯先验)

$$\beta | \gamma \sim \mathcal{MVN}(0, \mathbf{D}_{\gamma} \mathbf{R} \mathbf{D}_{\gamma})$$
 (2.4)

其中
$$\gamma = (\gamma_1, ..., \gamma_p)^{\top}$$
, R 是先验协方差矩阵, $D_{\gamma} = diag[\{d_i\}_{i=1}^p]$, 其中 $d_i = (1 - \gamma_i)\tau_i + \gamma_i c_i \tau_i$.

Dom Tsang SCNU 2017/5/11 9/22

 γ 的先验

$$\gamma$$
 (二项分布)

$$\gamma \sim \prod_{i=1}^{p} \omega_i^{\gamma_i} (1 - \omega_i)^{1 - \gamma_i} \tag{2.5}$$

Dom Tsang SCNU 2017/5/11 10 / 22

σ^2 的先验

$$\sigma^2$$
(逆 gamma 分布)

$$\sigma^2 | oldsymbol{\gamma} \sim \mathcal{IG}(rac{
u_{oldsymbol{\gamma}}}{2}, rac{
u_{oldsymbol{\gamma}} \lambda_{oldsymbol{\gamma}}}{2})$$

(2.6)

Dom Tsang SCNU 2017/5/11 11/22

β的后验

β的后验

$$\boldsymbol{\beta}^{(j)} \sim f(\boldsymbol{\beta}^{(j)}|\boldsymbol{Y}, \sigma^{2^{(j-1)}}, \boldsymbol{\gamma}^{j-1})$$

$$= \mathcal{MVN}(A_{\gamma^{j-1}}(\sigma^{j-1})^{-2}\boldsymbol{X}^{\top}\boldsymbol{Y}, A_{\gamma^{j-1}})$$
(2.7)

其中,
$$A_{\gamma^{j-1}} = ((\sigma^{j-1})^{-2} \boldsymbol{X}^{\top} \boldsymbol{X} + \boldsymbol{D}_{\gamma^{j-1}}^{-1} \boldsymbol{R}^{-1} \boldsymbol{D}_{\gamma^{j-1}}^{-1}).$$

Dom Tsang SCNU 2017/5/11 12 / 22

σ^2 的后验

$$\sigma^2$$
 的后验

$$\sigma^{2(j)} \sim f(\sigma^{2(j)} | \mathbf{Y}, \boldsymbol{\beta}^{j}, \boldsymbol{\gamma}^{j-1})$$

$$= \mathcal{IG}(\frac{n + \nu_{\gamma^{j-1}}}{2}, \frac{|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}^{j}|^{2} + \nu_{\gamma^{j-1}}\lambda_{\gamma^{j-1}}}{2})$$
(2.8)

Dom Tsang SCNU 2017/5/11 13 / 22

γ 的后验

$$\gamma_i^j \sim f(\gamma_i^j | \boldsymbol{Y}, \boldsymbol{\beta}^j, \sigma^{2(j)}, \gamma_{(\neq i)}^j) = f(\gamma_i^j | \boldsymbol{\beta}^j, \sigma^{2(j)}, \gamma_{(\neq i)}^j)$$
 (2.9)

每一个伯努利分布的概率为:

$$P(\gamma_i^j = 1 | \boldsymbol{\beta}^{(j)}, \sigma^{2(j)}, \gamma_{(\neq i)}^{(j)}) = \frac{a}{a+b}$$
 (2.10)

其中,

$$a = f(\boldsymbol{\beta}^{(j)}|\boldsymbol{\gamma}_{(\neq i)}^{(j)}), \gamma_{(i)}^{(j)} = 1)f(\sigma^{2(j)}|\boldsymbol{\gamma}_{(\neq i)}^{(j)}), \gamma_{(i)}^{(j)} = 1)f(\boldsymbol{\gamma}_{(\neq i)}^{(j)}), \gamma_{(i)}^{(j)} = 1)$$
(2.11)

$$b = f(\boldsymbol{\beta}^{(j)}|\boldsymbol{\gamma}_{(\neq i)}^{(j)}), \gamma_{(i)}^{(j)} = 0) f(\sigma^{2(j)}|\boldsymbol{\gamma}_{(\neq i)}^{(j)}), \gamma_{(i)}^{(j)} = 0) f(\boldsymbol{\gamma}_{(\neq i)}^{(j)}), \gamma_{(i)}^{(j)} = 0)$$
(2.12)

Dom Tsang SCNU 2017/5/11 14/22

γ 的后验

当 σ 的先验参数为常数时 $(\nu_{\gamma} \equiv \nu, \lambda_{\gamma} \equiv \lambda)$, 式2.11, 2.12可进一步简化为

$$a = f(\boldsymbol{\beta}^{(j)}|\boldsymbol{\gamma}_{(\neq i)}^{(j)}), \boldsymbol{\gamma}_{(i)}^{(j)} = 1)f(\boldsymbol{\gamma}_{(\neq i)}^{(j)}), \boldsymbol{\gamma}_{(i)}^{(j)} = 1)$$

$$= f(\boldsymbol{\beta}^{(j)}|\boldsymbol{\gamma}_{(\neq i)}^{(j)}), \boldsymbol{\gamma}_{(i)}^{(j)} = 1)\omega_{i}$$

$$b = f(\boldsymbol{\beta}^{(j)}|\boldsymbol{\gamma}_{(\neq i)}^{(j)}), \boldsymbol{\gamma}_{(i)}^{(j)} = 0)f(\boldsymbol{\gamma}_{(\neq i)}^{(j)}), \boldsymbol{\gamma}_{(i)}^{(j)} = 0)$$

$$= f(\boldsymbol{\beta}^{(j)}|\boldsymbol{\gamma}_{(\neq i)}^{(j)}), \boldsymbol{\gamma}_{(i)}^{(j)} = 0)(1 - \omega_{i})$$
(2.14)

Dom Tsang SCNU 2017/5/11 15 / 22

算法

SSVS

- 1 初始化: $\beta^0, \sigma^2, \gamma, c, \tau, R, \omega, \nu, \lambda$
- **2** 迭代: for t = 1, 2, ...:
 - 从式2.7中抽取 β;
 - 从式2.8中抽取 σ²;
 - 从式2.10中抽取 γ.

Dom Tsang SCNU 2017/5/11 16 / 22

CONTENTS

- 1 前情回顾
- 2 Stochastic Search Variable Selection(随机搜索变量选择)
- 3 算法
- 4 模型选择
- 5 参考文献

- 问题描述
- 模型空间
- 模型的量化表示
- 模型的后验概率估计

Dom Tsang SCNU 2017/5/11 17 / 22

- 问题: 从模型集合中选择一个最优模型
- 在贝叶斯框架中,模型选择问题被转化为参数估计的形式而不是寻找单一的最优模型,贝叶斯方法将估计所有模型的后验概率.在很多情况下,这个问题从变量的层面被提出:例如,估计模型中变量的边缘后验概率.

Dom Tsang SCNU 2017/5/11 18 / 22

模型空间

在模型选择中,模型空间的均匀分布被定义成:

$$f(m) = \frac{1}{|\mathcal{M}|}, for all m \in \mathcal{M}$$
 (4.1)

Dom Tsang SCNU 2017/5/11 19 / 22

■ γ 的向量空间与模型空间建立起一个一一对应的映射:

$$m(\gamma) = \sum_{j=k}^{p} \gamma_j 2^{j-k} \tag{4.2}$$

Dom Tsang SCNU 2017/5/11 20/22

模型的后验概率

$$\hat{f}(m|\mathbf{Y}) = \frac{1}{T - B} \sum_{t = B + 1} I(m^{(t)} = m)$$
 (4.3)

其中, T 和 B 分别是算法总共的迭代次数和预模拟运行 (burnin) 的迭代次数, $m^{(t)}$ 是模型在迭代次数 t 下的指示值.

Dom Tsang SCNU 2017/5/11 21/22

Reference



O'Hara R B, Sillanpaa M J. A review of Bayesian variable selection methods: what, how and which[J]. Bayesian analysis, 2009, 4(1): 85-117.



George E I, McCulloch R E. Variable selection via Gibbs sampling[J]. Journal of the American Statistical Association, 1993, 88(423): 881-889.

Dom Tsang SCNU 2017/5/11 22 / 22

法

模型选择 0000



参考文献

Thanks!

Dom Tsang SCNU 2017/5/11 22 / 22