Reversible Jump MCMC 与准备金估计模型选择

刘乐平、高磊

天津财经大学统计系

2013 年 3 月 25 日

参考文献

- ⋆ GREEN, P. J. 1995. Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination.Biometrika.
- ⋆ VERRALL, R. J., AND M. V. WUTHRICH. 2012.Reversible Jump Markov Chain Monte Carlo Method For Parameter Reduction In Claims Reserving.North American Actuarial Journal.

Contents

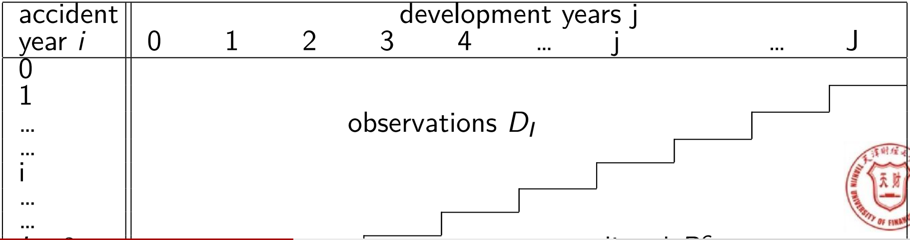
Contents

Contents

内容

损失流量三角形

accyear \ devyear	0	1	2	3	4	5	6	7	8	9
0	357848	766940	610542	482940	527326	574398	146342	139950	227229	67948
1	352118	884021	933894	1183289	445745	320996	527804	266172	425046	NA
2	290507	1001799	926219	1016654	750816	146923	495992	280405	NA	NA
3	310608	1108250	776189	1562400	272482	352053	206286	NA	NA	NA
4	443160	693190	991983	769488	504851	470639	NA	NA	NA	NA
5	396132	937085	847498	805037	705960	NA	NA	NA	NA	NA
6	440832	847631	1131398	1063269	NA	NA	NA	NA	NA	NA
7	359480	1061648	1443370	NA	NA	NA	NA	NA	NA	NA
8	376686	986608	NA	NA	NA	NA	NA	NA	NA	NA
9	344014	NA	NA	NA	NA	NA	NA	NA	NA	NA



乘法模型

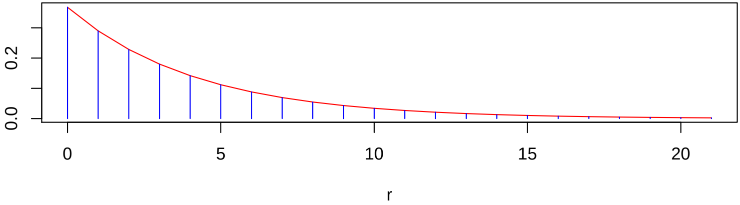
有以下记号：

- ⋆ 事故年水平：μ_i
- ⋆ 进展年水平：r_j
- ⋆ 乘法模型：$E_{\mu_i,\gamma_j}[X_{i,j}] = \mu_i\gamma_j, i, j \in 0, ..., I$

根据乘法模型，为了估计增量下三角形，须估计出 $2(I+1)$ 个参数 $\mu_0, ..., \mu_I; \gamma_0, ..., \gamma I$。而我们仅有 $(I+1)(I+2)/2$ 个数据 (上三角形数据的个数)。统计中，我们把它称为过度参数化问题。

解决过度参数化问题

为解决过度参数化问题，很自然地就是用一条曲线拟合这些参数，例如，用指数递减曲线拟合 $\gamma_0, \ldots, \gamma_I$。



解决过度参数化问题

- 举例，$I = 8, k = 4$

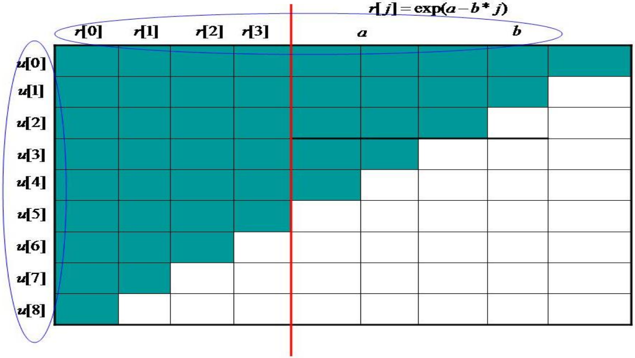


Figure 1: GAM

ODP Model–超散布泊松分布模型

⋆ 给定 $\vartheta = (\mu_0, ..., \mu_I; \gamma_0, ..., \gamma_I, \phi)$,$X_{i,j}$ 为独立随机变量，满足

$$\frac{X_{i,j}}{\phi}|_\vartheta \sim Poisson(\mu_i\gamma_j/\phi)$$

⋆ 与泊松分布的区别：

$$E[X_{i,j}|\vartheta] = \mu_i\gamma_j$$

$$Var(X_{i,j}|\vartheta) = \boxed{\phi\mu_i\gamma_j}$$

ODP Model–超散布泊松分布模型

- ⋆ 给定 $\vartheta = (\mu_0, ..., \mu_I; \gamma_0, ..., \gamma_I, \phi)$,$X_{i,j}$ 为独立随机变量，满足

$$\frac{X_{i,j}}{\phi}|_{\vartheta} \sim Poisson(\mu_i\gamma_j/\phi)$$

- ⋆ 与泊松分布的区别：

$$E[X_{i,j}|\vartheta] = \mu_i\gamma_j$$

$$Var(X_{i,j}|\vartheta) = \boxed{\phi\mu_i\gamma_j}$$

ODP Model–超散布泊松分布模型

- ⋆ 给定 $\vartheta = (\mu_0, ..., \mu_I; \gamma_0, ..., \gamma_I, \phi)$,$X_{i,j}$ 为独立随机变量，满足

$$\frac{X_{i,j}}{\phi}|_{\vartheta} \sim Poisson(\mu_i\gamma_j/\phi)$$

- ⋆ 与泊松分布的区别：

$$E[X_{i,j}|\vartheta] = \mu_i\gamma_j$$

$$Var(X_{i,j}|\vartheta) = \boxed{\phi\mu_i\gamma_j}$$

ODP model + GAM + Bayes

固定 k:

- 模型参数为:

$$\theta_k = (\alpha, \beta, \mu_0, ..., \mu_I; \gamma_0, ..., \gamma_{k-1})$$

- 对于 $j \in k, \ldots, I,$

$$\gamma_j = exp(\alpha - j\beta)$$

- 先验分布为

$$\mu_i \sim \Gamma(s, s/m_i), i = 0, \ldots, I,$$

$$\gamma_j \sim \Gamma(v, v/c_j), j = 0, \ldots, k-1$$

$$\alpha \sim N(\alpha, \sigma^2), \beta \sim N(b, \tau^2)$$

ODP model + GAM + Bayes

固定 k:

- 模型参数为:

$$\theta_k = (\alpha, \beta, \mu_0, ..., \mu_I; \gamma_0, ..., \gamma_{k-1})$$

- 对于 $j \in k, \ldots, I$,

$$\gamma_j = exp(\alpha - j\beta)$$

- 先验分布为

$$\mu_i \sim \Gamma(s, s/m_i), i = 0, \ldots, I,$$

$$\gamma_j \sim \Gamma(v, v/c_j), j = 0, \ldots, k-1$$

$$\alpha \sim N(\alpha, \sigma^2), \beta \sim N(b, \tau^2)$$

ODP model + GAM + Bayes

固定 k:

⋆ 模型参数为：

$$\theta_k = (\alpha, \beta, \mu_0, ..., \mu_I; \gamma_0, ..., \gamma_{k-1})$$

⋆ 对于 $j \in k, \ldots, I$,

$$\gamma_j = exp(\alpha - j\beta)$$

⋆ 先验分布为

$$\mu_i \sim \Gamma(s, s/m_i), i = 0, \ldots, I,$$

$$\gamma_j \sim \Gamma(v, v/c_j), j = 0, \ldots, k-1$$

$$\alpha \sim N(\alpha, \sigma^2), \beta \sim N(b, \tau^2)$$

ODP model + GAM + Bayes

固定 k:

- 模型参数为:

$$\theta_k = (\alpha, \beta, \mu_0, ..., \mu_I; \gamma_0, ..., \gamma_{k-1})$$

- 对于 $j \in k, \ldots, I$,

$$\gamma_j = exp(\alpha - j\beta)$$

- 先验分布为

$$\mu_i \sim \Gamma(s, s/m_i), i = 0, \ldots, I,$$

$$\gamma_j \sim \Gamma(v, v/c_j), j = 0, \ldots, k-1$$

$$\alpha \sim N(\alpha, \sigma^2), \beta \sim N(b, \tau^2)$$

后验分布的推导

⋆ 后验分布：

$$p_k(\theta_k|(X_{i,j})_{(i,j)\in\Omega}) \propto f_k((X_{i,j})_{(i,j)\in\Omega}, \theta_k) \tag{1}$$

⋆ 利用MCMC可以从该分布中采样，进而获得相关参数的估计值，以此可以估计得到准备金。

⋆ k?????,Let the data speak.How?Answer:Bayes idea again! + Tool:RJMCMC!

后验分布的推导

⋆ 后验分布：

$$p_k(\theta_k|(X_{i,j})_{(i,j)\in\Omega}) \propto f_k((X_{i,j})_{(i,j)\in\Omega}, \theta_k) \quad (1)$$

⋆ 利用MCMC可以从该分布中采样，进而获得相关参数的估计值，以此可以估计得到准备金。

⋆ k?????,Let the data speak.How?Answer:Bayes idea again! + Tool:RJMCMC!

后验分布的推导

- ⋆ 后验分布：

$$p_k(\theta_k|(X_{i,j})_{(i,j)\in\Omega}) \propto f_k((X_{i,j})_{(i,j)\in\Omega}, \theta_k) \quad (1)$$

- ⋆ 利用MCMC可以从该分布中采样，进而获得相关参数的估计值，以此可以估计得到准备金。
- ⋆ k?????,Let the data speak.How?Answer:Bayes idea again! + Tool:RJMCMC!

后验分布的推导

- ⋆ 后验分布：

$$p_k(\theta_k|(X_{i,j})_{(i,j)\in\Omega}) \propto f_k((X_{i,j})_{(i,j)\in\Omega}, \theta_k) \quad (1)$$

- ⋆ 利用MCMC可以从该分布中采样，进而获得相关参数的估计值，以此可以估计得到准备金。
- ⋆ k?????,Let the data speak.How?Answer:Bayes idea again! + Tool:RJMCMC!

内容

采样样本概况

利用 RJMCMC 抽样得到的样本概况（M1 有一个参数，M2 有两个参数）。

T	模型	参数	
t	1	θ_{11}^{t}	
t+1	2	θ_{21}^{t+1}	θ_{22}^{t+1}
t+2	2	θ_{21}^{t+2}	θ_{22}^{t+2}
t+3	1	θ_{11}^{t+3}	
t+4	1	θ_{11}^{t+4}	
t+5	1	θ_{11}^{t+5}	
t+6	2	θ_{21}^{t+2}	θ_{22}^{t+2}
...	...	...	...

Table 1: RJMCMC 样本示例

内容

关于数据的设计

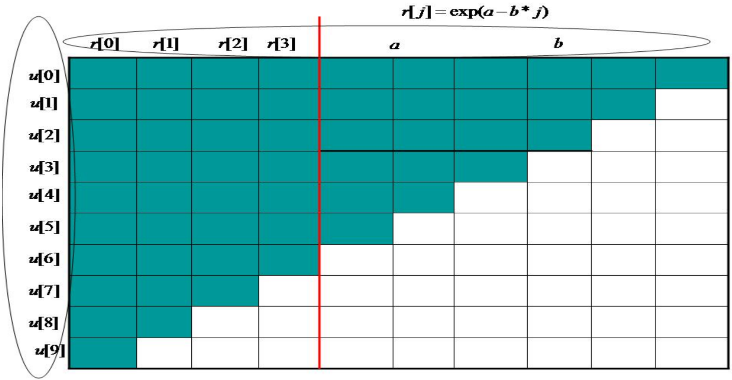


Figure 2: 模拟数据设计

关于数据的设计

- ⋆ 我们定义事故年数为 10，进展年也为 10.

```
I <- 9
```

- ⋆ 事故年水平的定义为

```
u = 1.02^(0:I) * 10^7
u

## [1] 10000000 10200000 10404000 10612080 10824322 11040808 11261624 114
```

- ⋆ 我们设计的数据在 k=4 处划分为两段，$\gamma_0, \ldots, \gamma_3$ 有具体的数值定义，而 $\gamma_4, \ldots, \gamma_9$ 则服从一指数递减曲线，相关参数也已经设置。

```
r0_3 <- c(0.159, 0.179, 0.179, 0.139)
alpha = -1.6159
beta = 0.2
r4_9 <- exp(alpha - beta * (4:9))
```

关于数据的设计

	0	1	2	3	4	5	6	7	8	9
0	1619686	1605365	1789134	1204245	701947	644759	741318	606242	419492	433979
1	2096654	1828792	1912791	1283858	1172218	714845	707710	525486	397655	0
2	2013759	1736024	1684836	1885917	852724	612746	734329	566127	0	0
3	1508176	1748213	1714909	1453170	890544	669430	480932	0	0	0
4	1565750	1640238	2160561	1985603	1127682	871973	0	0	0	0
5	2199588	2158210	1854062	1792833	852008	0	0	0	0	0
6	2065669	2217071	1818990	1270410	0	0	0	0	0	0
7	1858507	2625637	1863646	0	0	0	0	0	0	0
8	1783441	1900148	0	0	0	0	0	0	0	0
9	1848961	0	0	0	0	0	0	0	0	0

RJMCMC 采样结果

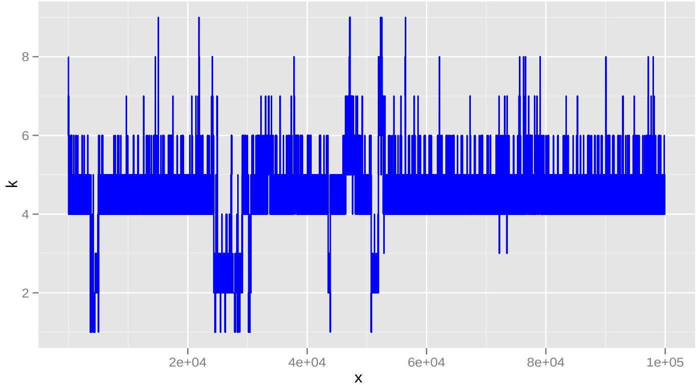


Figure 3: RJMCMC 采样 K 值

RJMCMC 采样结果

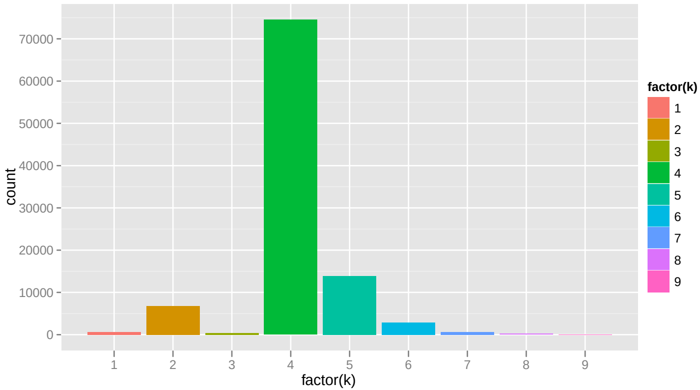


Figure 4: RJMCMC 采样 K 值

RJMCMC 采样结果

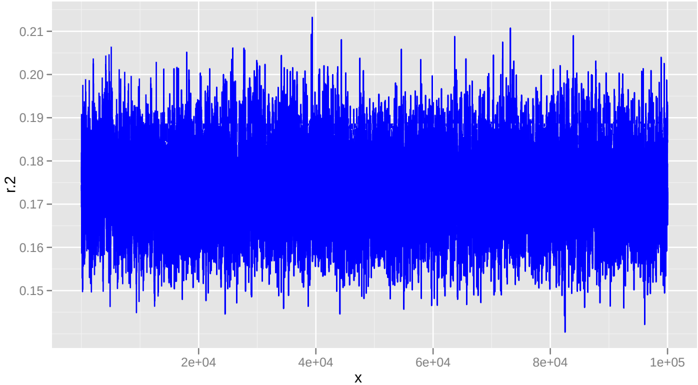


Figure 5: RJMCMC 采样 r.2 样本

RJMCMC 采样结果

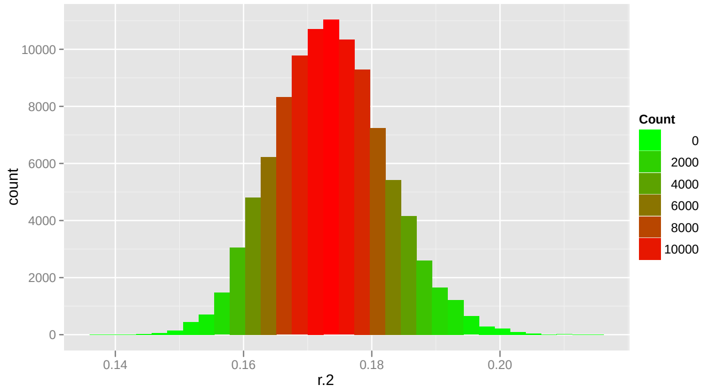


Figure 6: RJMCMC 采样 r.2 分布

RJMCMC 采样结果

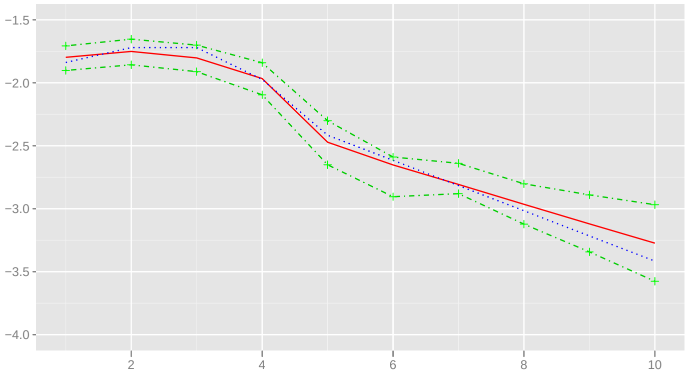


Figure 7: RJMCMC 采样 r 均值