*The Korean Journal of Applied Statistics* (2015) DOI: http://dx.doi.org/10.5351/KJAS.2015.28.2.281
**28**(2), 281–288

# Sparse Matrix Computation in Mixed Effects Model

Won Son[a] · Yong-Tae Park[b] · Yu Kyeong Kim[c] · Johan Lim[a,1]

[a]Department of Statistics, Seoul National University
[b]Department of Industrial Engineering, Seoul National University
[c]Department of Diagnostic Radiology, Seoul National University Hospital

(Received March 18, 2015; Revised April 1, 2015; Accepted April 1, 2015)

**Abstract**

In this paper, we study an approximate procedure to evaluate a penalized maximum likelihood estimator (MLE) for a mixed effects model. The procedure approximates the Hessian matrix of the penalized MLE with a structured sparse matrix or an arrowhead type matrix to speed its computation. In this paper, we numerically investigate the gain in computation time as well as approximation error from the considered approximation procedure.

Keywords: Arrow head type matrix, mixed effects model, penalized maximum likelihood estimator, sparse matrix.

## 1. 서론

혼합모형(mixed effects model)은 설명변수들의 반응변수에 대한 상관도를 나타내는 고정효과(fixed effect)와 궁극적으로는 반응변수들 간의 상호의존성을 결정짓는 랜덤효과(random effect)를 동시에 포함한 모형으로 유전학, 공학, 뇌과학 등 다양한 의학 자료에 널리 사용되는 모형이다 (Beckmann 등, 2003; Sohn 등, 2007; Yoon과 Sohn, 2007). 혼합모형을 조금 더 수리적으로 살펴보면 다음과 같다. 고정효과 $\boldsymbol{\beta}$와 랜덤효과 $\boldsymbol{\nu}$에 대응하는 설명변수 벡터를 각각 $\mathbf{x}$와 $\mathbf{z}$라 하고 이에 대응하는 반응변수 값을 $Y$라 하면 혼합모형은 $\boldsymbol{\nu}$가 주어진 상태에서 $Y$의 확률분포 $f(Y|\boldsymbol{\nu};\boldsymbol{\beta},\boldsymbol{\theta})$를 가정한다. 여기에서 $\boldsymbol{\theta}$는 장애모수(nuisance parameter)로 주로 분산성분들(variance components)에 대한 벡터이다.

혼합모형의 추론에 널리 쓰이는 방법 중 하나는 최대우도추정량(maximum likelihood estimator; MLE)에 기반 한 것으로 관측값 $\{(y_i, \mathbf{x}_i, \mathbf{z}_i), i = 1, 2, \ldots, \}$이 혼합모형을 따른다고 할 때 모수벡터 $(\boldsymbol{\beta}, \boldsymbol{\theta})$에 대한 추정은 로그-우도함수(log-likelihood function)

$$\ell(\boldsymbol{\beta}, \boldsymbol{\theta}) = \sum_{i=1}^{n} \int_{\boldsymbol{\nu}} f(y_i|\boldsymbol{\nu}; (\mathbf{x}_i, \mathbf{z}_i), \boldsymbol{\beta}, \boldsymbol{\theta})\, f(\boldsymbol{\nu}; \boldsymbol{\theta})\, d\boldsymbol{\nu} \tag{1.1}$$

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST)(No. 2011-0030810).
[1]Corresponding author: Department of Statistics, Seoul National University, 1, Gwanak-ro, Gwanak-gu, Seoul 151-742, Korea. E-mail: johanlim@snu.ac.kr

를 최대로 하는 값으로 정의된다. 최대우도추정량은 통계적 효율성(statistical efficiency)을 포함한 여러 좋은 성질들을 지니고 있으나 설명변수 벡터 $\mathbf{z}_i$와 $\boldsymbol{\nu}$의 특성에 따라 식 (1.1)의 다중적분계산에 어려움이 있고 이를 해결하기 위한 많은 연구가 수행되어왔다.

식 (1.1)의 계산과 관련한 여러 대안들 중 가장 널리 이용되는 것은 벌점우도방법(penalized lielihood method)이고 여기서 벌점우도는 완전로그우도함수(complete log-likeliood function)

$$\ell(\boldsymbol{\beta}, \boldsymbol{\nu}) = \sum_{i=1}^{n} \log f(Y|\boldsymbol{\nu}; \boldsymbol{\beta}, \boldsymbol{\theta})$$

와 벌점함수 $p(\boldsymbol{\nu}; \boldsymbol{\theta})$의 합인

$$p\ell(\boldsymbol{\beta}, \boldsymbol{\theta}) = \ell(\boldsymbol{\beta}, \boldsymbol{\nu}) + p(\boldsymbol{\nu}; \boldsymbol{\theta}) \tag{1.2}$$

로 정의된다. 또한 벌점함수로는 $p(\boldsymbol{\nu}; \boldsymbol{\theta}) = \log f\boldsymbol{\nu}; \boldsymbol{\theta})$가 보편적으로 사용된다. $\theta$가 주어졌을 때 식 (1.2)을 최대화하는 $(\boldsymbol{\beta}, \boldsymbol{\nu})$의 벌점최대우도추정량(penalized MLE)은

$$\begin{pmatrix} \mathbf{X}^{\mathrm{T}} W \mathbf{X} & \mathbf{X}^{\mathrm{T}} W \mathbf{Z} \\ \mathbf{Z}^{\mathrm{T}} W \mathbf{X} & \mathbf{Z}^{\mathrm{T}} W \mathbf{Z} + \mathrm{U} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\nu} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^{\mathrm{T}} W \mathbf{y} \\ \mathbf{Z}^{\mathrm{T}} W \mathbf{y} \end{pmatrix} \tag{1.3}$$

의 해로 표시된다. 따라서 추정방정식 (1.3)에 대한 효율적 계산은 벌점우도추정량의 계산에 있어서 매우 중요한 부분이다. 여기에서 $\mathbf{X}$는 $i$번째 열벡터가 $\mathbf{x}_i$인 $n \times p$행렬이고, $\mathbf{Z}$는 $n \times q$행렬로 $\mathbf{z}_i$를 $i$번째 열벡터로 가지며, $n \times n$행렬 $W$는 $\mu = \mathrm{E}(Y)$의 함수로 형태는 $Y$의 분포가정으로부터 결정된다. 마지막으로 $q \times q$행렬 U는 $\mathrm{U} = -\{\partial^2 p(\boldsymbol{\nu}; \boldsymbol{\theta})/(\partial \boldsymbol{\nu} \partial \boldsymbol{\nu}^{\mathrm{T}})\}$로 정의된다.

추정방정식 (1.3)은 본 메모의 주제인 혼합모형 뿐만 아니라 다른 여러 모형들에 대한 벌점우도방법에서도 등장하고 이들 문제들에서 벌점최대우도추정량의 계산 또한 추정방정식 (1.3)을 반복적으로 풀 것을 요구하게 된다. 따라서 이들 문제에서도 추정방정식 (1.3)에 대한 빠른 계산은 벌점우도방법에 기반한 다양한 문제들을 해결하는 데 중요한 부분이 된다. 특히 최근 컴퓨터 기술의 발달과 더불어 고차원 대용량 문제들이 등장하고 있으며 이에 따라 종종 랜덤효과 $\boldsymbol{\nu}$의 차원이 상당히 커지는 상황이 발생하고 있어 추정방정식 (1.3)의 계산을 더욱 어렵게 한다 (Zhu와 Hastie, 2004; Lee와 Oh, 2014).

본 논문에서는 $\boldsymbol{\nu}$의 차원이 큰 경우(또는 동등하게 $q$가 $p$에 비하여 상당히 큰 경우) 희소행렬을 이용한 추정방정식 (1.3)의 근사계산에 대하여 살펴본다. 해당 근사계산에서 사용되는 희소행렬은 그 모양에 기인하여 화살촉행렬(arrow-head type matrix)이라 불리고 해당 행렬의 역행렬 계산 등 여러 성질에 관련하여 본 논문의 2절에서 공부하여 본다. 본 논문의 3절에서는 두 개의 예제를 통하여 희소행렬 계산으로 우리가 얻게 되는 계산 시간의 절약과 이에 대한 비용으로 지불하는 추정량에서 발생하는 근사오차에 대하여 살펴본다. 두 예제 중 하나는 가상의 자료이고 다른 하나는 Fleming과 Harrington (2005)의 만성육아종병(chronic granulomtous disease; CGD) 자료에 기반하여 생성된 자료이다. 마지막으로 4절에서는 본문에서 다루지 못한 몇 가지 문제들에 대한 논의를 추가하며 본 논문를 마친다.

## 2. 벌점우도의 헷시안(Hessian)행렬에 대한 희소행렬근사

본절에서는 식 (1.3)의 좌변에 위치한 헷시안행렬에 대한 희소행렬근사에 사용되는 화살촉행렬과 이의 역행렬 계산식들을 살펴본다.

먼저 본 메모에서 다루는 화살촉행렬은 $p \times p$ 대칭행렬 $A_{11}$, $p \times q$행렬 $A_{12}$와 $q \times q$ 대각(또는 블록대각)행렬 D에 대하여

$$\mathbf{K} = \begin{pmatrix} \mathrm{A}_{11} & \mathrm{A}_{12} \\ \mathrm{A}_{12}^{\mathrm{T}} & \mathrm{D} \end{pmatrix} \tag{2.1}$$

의 형태를 지니는 행렬을 의미한다. 통상 $(p+q) \times (p+q)$ 차원의 역행렬은 $O((p+q)^3)$의 작업이 필요하나 희소-가우스소거법(sparse Gauss elimination)을 이용하면 위 행렬의 역행렬 계산에 $O(\max(pq, p^3))$의 작업이 필요함이 알려져 있다 (Demmel, 1997).

본 메모에서는 $\mathbf{K}$의 역행렬에 대한 계산 효율성을 지닌 수리적 식을 제시하고자 한다. 제시한 식은 우리가 무심코 지나친 여러 회귀분석 교재의 분할행렬의 역행렬식에 기반하고 있고 희소-가우스소거법과 같이 $O(\max(pq, p^3))$의 계산복잡성(computational complexity)을 지닌다. 제시하는 역행렬 공식은 다음과 같다.

**정리 2.1** 목적행렬 $\mathbf{K}$의 우상행렬인 $\mathrm{A}_{11}$의 역행렬이 존재한다는 가정하에

$$\mathbf{K}^{-1} = \begin{pmatrix} \mathrm{A}_{11}^{-1} + \mathrm{A}_{11}^{-1}\mathrm{A}_{12}\mathrm{Q}\mathrm{A}_{21}\mathrm{A}_{11}^{-1} & -\mathrm{A}_{11}^{-1}\mathrm{A}_{12}\mathrm{Q} \\ -\mathrm{Q}\mathrm{A}_{21}\mathrm{A}_{11}^{-1} & \mathrm{Q} \end{pmatrix} \tag{2.2}$$

이고 여기서 $\mathrm{Q} = \mathrm{D}^{-1} + \mathrm{D}^{-1}\mathrm{A}_{21}(\mathrm{A}_{11} - \mathrm{A}_{12}\mathrm{D}^{-1}\mathrm{A}_{21})^{-1}\mathrm{A}_{12}\mathrm{D}^{-1}$이다.

증명: 식 (2.1)의 목적행렬 $\mathbf{K}$는 분할행렬의 역행렬공식 (Park, 2007)에 의하여 행렬

$$Q = \left(\mathrm{D} - A_{21}A_{11}^{-1}A_{12}\right)^{-1}$$

에 대하여

$$\begin{pmatrix} \mathrm{A}_{11}^{-1} + \mathrm{A}_{11}^{-1}\mathrm{A}_{12}\mathrm{Q}\mathrm{A}_{21}\mathrm{A}_{11}^{-1} & -\mathrm{A}_{11}^{-1}\mathrm{A}_{12}\mathrm{Q} \\ -\mathrm{Q}\mathrm{A}_{21}\mathrm{A}_{11}^{-1} & \mathrm{Q} \end{pmatrix}$$

로 표현되고 역행렬에 관한 Sherman-Morrison-Woodbery 공식 (Hager, 1989)에 의하여

$$\begin{aligned} Q &= \left(\mathrm{D} - \mathrm{A}_{21}\mathrm{A}_{11}^{-1}\mathrm{A}_{12}\right)^{-1} \\ &= \mathrm{D}^{-1} + \mathrm{D}^{-1}\mathrm{A}_{21}\left(\mathrm{A}_{11} - \mathrm{A}_{12}\mathrm{D}^{-1}\mathrm{A}_{21}\right)^{-1}\mathrm{A}_{12}\mathrm{D}^{-1} \end{aligned}$$

로 표현된다. □

위의 공식에서 $\mathrm{D}^{-1}$와 $(\mathrm{A}_{11} - \mathrm{A}_{12}\mathrm{D}^{-1}\mathrm{A}_{21})^{-1}$는 각각 $O(q)$와 $O(p^3)$의 계산-복잡성을 지난다.

본 연구에서 제시하는 헷시안행렬의 역행렬에 대한 희소행렬근사 절차는 다음과 같다.

$$\begin{aligned} \mathbf{H} &= \begin{pmatrix} \mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{X} & \mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{Z} \\ \mathbf{Z}^{\mathrm{T}}\mathrm{W}\mathbf{X} & \mathrm{diag}\left(\mathbf{Z}^{\mathrm{T}}\mathrm{W}\mathbf{Z}\right) + \mathrm{U} \end{pmatrix} + \begin{pmatrix} \mathbf{0}_{p\times p} & \mathbf{0}_{p\times q} \\ \mathbf{0}_{q\times p} & \mathbf{Z}^{\mathrm{T}}\mathrm{W}\mathbf{Z} - \mathrm{diag}\left(\mathbf{Z}^{\mathrm{T}}\mathrm{W}\mathbf{Z}\right) \end{pmatrix} \\ &\approx \begin{pmatrix} \mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{X} & \mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{Z} \\ \mathbf{Z}^{\mathrm{T}}\mathrm{W}\mathbf{X} & \mathrm{D} = \left\{\mathrm{diag}\left(\mathbf{Z}^{\mathrm{T}}\mathrm{W}\mathbf{Z}\right) + \mathrm{U}\right\} \end{pmatrix} \equiv \mathbf{K} \end{aligned} \tag{2.3}$$

로 근사하고 따라서 $\mathbf{H}^{-1} \approx \mathbf{K}^{-1}$을 사용한다.

위에서 제안된 희소행렬을 이용한 $\mathbf{H}^{-1}$의 근사오차에 대하여 살펴보면 $\mathbf{M} \equiv \mathbf{Z}^{\mathrm{T}}\mathrm{W}\mathbf{Z} - \mathrm{diag}(\mathbf{Z}^{\mathrm{T}}\mathrm{W}\mathbf{Z})$이 크지 않다는 전제하에

$$\begin{aligned}\mathbf{H}^{-1} &= \left\{\mathbf{K} + \begin{pmatrix} \mathbf{0}_{p\times p} & \mathbf{0}_{p\times q} \\ \mathbf{0}_{q\times p} & \mathbf{M} \end{pmatrix}\right\}^{-1} \\ &= \mathbf{K}^{-1} - \mathbf{K}^{-1}\begin{pmatrix} \mathbf{0}_{p\times p} & \mathbf{0}_{p\times q} \\ \mathbf{0}_{q\times p} & \mathbf{M} \end{pmatrix}\mathbf{K}^{-1}\end{aligned}$$

이고 대략적인 근사오차는 위의 정리 2.1에서 정의된 행렬

$$\mathrm{Q} = \mathrm{D}^{-1} + \mathrm{D}^{-1}\mathbf{Z}^{\mathrm{T}}\mathrm{W}\mathbf{X}\left\{\mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{X} - \mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{Z}\mathrm{D}^{-1}\mathbf{Z}^{\mathrm{T}}\mathrm{W}\mathbf{X}\right\}^{-1}\mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{Z}\mathrm{D}^{-1}$$

에 대하여

$$\begin{aligned}&\mathbf{K}^{-1} - \mathbf{H}^{-1} \\ &= \mathbf{K}^{-1}\begin{pmatrix} \mathbf{0}_{p\times p} & \mathbf{0}_{p\times q} \\ \mathbf{0}_{q\times p} & \mathbf{M} \end{pmatrix}\mathbf{K}^{-1} \\ &= \begin{pmatrix} \left(\mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{Z}\mathrm{Q}\mathbf{M}\mathrm{Q}\mathbf{Z}^{\mathrm{T}}\mathrm{W}\mathbf{X}\left(\mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{X}\right)^{-1} & -\left(\mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{Z}\mathrm{Q}\mathbf{M}\mathrm{Q} \\ -\mathrm{Q}\mathbf{M}\mathrm{Q}\mathbf{Z}^{\mathrm{T}}\mathrm{W}\mathbf{X}\left(\mathbf{X}^{\mathrm{T}}\mathrm{W}\mathbf{X}\right)^{-1} & \mathrm{Q}\mathbf{M}\mathrm{Q} \end{pmatrix}\end{aligned} \tag{2.4}$$

으로 표현된다. 따라서 근사오차는 QMQ에 의하여 결정된다 할 수 있다. 실제 근사식을 계산함에 있어 Q와 **M**을 알고 있으므로 약간의 부수적인 계산을 통하여 대략적인 근사오차를 이해할 수 있다.

논의한 화살촉행렬을 이용한 희소행렬계산에 대한 생각은 Therneau와 Grambsch (2000)의 교재에서 언급이 되었듯이 이미 연구자들에 의하여 실제로 종종 쓰이는 계산기법이고 본 연구에서는 여러 가상실험을 통하여 해당 근사를 통하여 얻을 수 있는 계산시간의 절약과 이에 대한 비용으로 지불하여야 하는 근사오차에 대하여 이해하여 보고자 한다.

## 3. 예제들

### 3.1. 가상자료

본 절에서는 본 논문에서 제안하는 희소행렬계산 자체를 통하여 절약되는 시간을 이해하기 위하여 아래와 같은 행렬 **H**의 역행렬을 직접 구하는 경우와 제안된 희소행렬의 역행렬 공식을 이용하여 계산하는 경우의 계산시간을 비교하여 보았다. 실험에서 고려된 행렬은

$$\mathbf{H} = \begin{pmatrix} \mathrm{A}_{11} & \mathrm{A}_{12} \\ \mathrm{A}_{21} & \mathrm{D} \end{pmatrix}$$

으로 여기서 $\mathrm{A}_{11} = 10\,\mathrm{I}_{10} + \mathrm{J}_{10}$, $\mathrm{A}_{12} = \mathrm{J}_{10,n-10}$, $\mathrm{D} = 10\,\mathrm{I}_{n-10}$로 설정되었다. 행렬 $\mathrm{I}_k$는 $k \times k$ 단위행렬이고, $\mathrm{J}_{l,m}$은 $l \times m$차원 행렬로 모든 원소가 1인 행렬이다. 실험은 Intel Core i7 (16GB RAM) 컴퓨터에서 R3.1.2를 이용하여 수행되었다. 행렬의 크기 $n$을 변화시켜가면서 일반적인 역행렬 계산 방법과 주어진 희소행렬 공식에 의한 방법의 계산소요시간을 기록하였으며 이를 Figure 3.1에 표현하였다. 각 방법들의 계산복잡성을 통하여 예상하였던 바와 같이 두 방법 간 계산소요시간에 큰 차이가 있으며 그 차이는 차원이 커질수록 급격히 커지는 것을 확인할 수 있다.

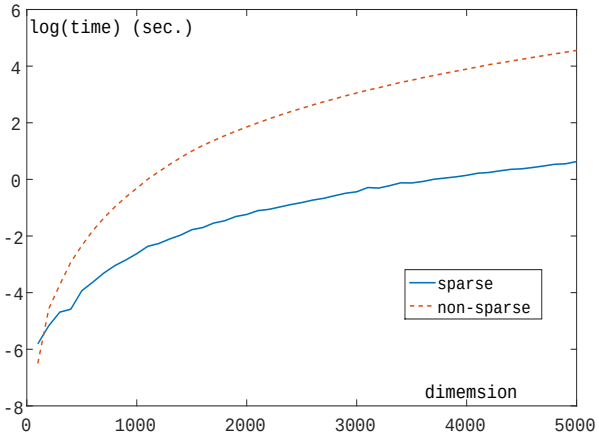


**Figure 3.1.** Log of computation time

### 3.2. 만성육아종병(CGD) 자료

본 실험에서는 Fleming과 Harrington (2005)에서 보고된 만성육아종병 환자 자료를 10배로 복제한 자료에 대하여 논문에서 제안된 희소행렬근사 방법을 적용하여 보고 이를 통하여 계산시간의 변화와 함께 실제 모수 추정량 값들에의 영향에 대하여도 살펴본다. 만성육아종병 자료는 총 203개의 관측치로 구성되어 있는데 본 원고에서는 차원이 큰 행렬을 얻기 위해 이 자료를 10번 복제하여 총 2,030개의 관측치를 가진 자료를 만들었다. 감염까지 걸린 기간을 종속변수로 하였으며 설명변수는 모두 11개로 각각 처치 종류(treatment, $\beta_1$), 성별($\beta_2$), 나이($\beta_3$), 키($\beta_4$), 몸무게($\beta_5$), 유전적 특질($\beta_6$), 스테로이드제 및 예방 목적 항생제(prophylactic antibiotics) 사용 여부($\beta_7, \beta_8$), 병원 구분 변수($\beta_9, \beta_{10}, \beta_{11}$) 등을 포함한다. 모수 추정을 위하여 로그정규 프레일티 모형(lognoraml frailty model)을 가정하고 Ha 등 (2001)에서 제안한 h-likelhood를 이용하여 프레일티 분산 $\theta$를 포함한 다음의 추정량들을 얻었다. 추정량의 계산에서 (1.3)를 반복하여 계산하게 되고 반복수에 따라 얻어진 추정량들을 Table 3.1에 보고하였다. 표에서 알 수 있듯이 일반적인 역행렬을 이용한 방법과 희소행렬을 이용한 근사적 추정방법의 모수 추정치에는 큰 차이가 없었으나 계산 시간 측면에서는 희소행렬을 이용한 근사적 추정방법이 역행렬을 이용한 직접 계산방법에 비해 이점을 가지고 있음을 알 수 있다.

## 4. 결론

혼합모형에서 최대우도추정량의 계산은 Pinheiro와 Bates (2000)의 책을 통하여 알 수 있듯이 통계학에서 오랜 기간 연구되어 왔으나 현재까지도 어려움을 겪고 있는 문제 중 하나이다. 본 연구에서는 최대우도추정량을 근사적으로 계산하는 절차를 제시하고 이의 계산효율성과 근사계산의 정확성에 대하여 살펴보았다. 본 연구에서 제안한 희소행렬을 이용한 근사절차는 (본 논문의 제한된 실험 하에서) 높은 계산효율성과 정확도를 보여주어 실제 많은 랜덤효과를 지니고 있는 모형들의 추정에 있어 유용하게 사용될 수 있으리라 사료된다.

마지막으로 본 논문의 수정 원고를 준비함에 있어 근사오차와 관련하여 두 분 심사위원들의 의미 있는 의견이 있었고 또한 본 논문에서는 보고되지 않은 근사오차와 관련한 가상실험의 일부 결과를 독자들과 공유하며 논문을 마치고자 한다. 본 연구에서 제안한 희소행렬을 이용한 근사의 근사오차는 2절의

**Table 3.1.** Parameter estimates and computation time for the CGD data

| | 추정 | 반복 횟수 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 방법 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 | (s.e.) |
| $\beta_1$ | Inv. | −1.151 | −1.153 | −1.155 | −1.156 | −1.157 | −1.158 | −1.158 | −1.158 | −1.159 | −1.159 | (0.101) |
| | Sps. | −1.147 | −1.152 | −1.154 | −1.156 | −1.158 | −1.159 | −1.159 | −1.160 | −1.160 | −1.160 | (0.100) |
| $\beta_2$ | Inv. | 0.793 | 0.792 | 0.792 | 0.792 | 0.792 | 0.791 | 0.791 | 0.791 | 0.791 | 0.791 | (0.153) |
| | Sps. | 0.788 | 0.791 | 0.791 | 0.791 | 0.791 | 0.791 | 0.791 | 0.791 | 0.791 | 0.791 | (0.151) |
| $\beta_3$ | Inv. | −0.088 | −0.088 | −0.088 | −0.088 | −0.087 | −0.087 | −0.087 | −0.087 | −0.087 | −0.087 | (0.013) |
| | Sps. | −0.088 | −0.088 | −0.088 | −0.088 | −0.087 | −0.087 | −0.087 | −0.087 | −0.087 | −0.087 | (0.013) |
| $\beta_4$ | Inv. | 0.008 | 0.008 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | (0.004) |
| | Sps. | 0.008 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | (0.004) |
| $\beta_5$ | Inv. | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.011 | 0.011 | 0.011 | 0.011 | 0.011 | (0.006) |
| | Sps. | 0.010 | 0.010 | 0.010 | 0.010 | 0.011 | 0.011 | 0.011 | 0.011 | 0.011 | 0.011 | (0.006) |
| $\beta_6$ | Inv. | −0.673 | −0.674 | −0.674 | −0.675 | −0.675 | −0.676 | −0.676 | −0.676 | −0.676 | −0.676 | (0.110) |
| | Sps. | −0.668 | −0.672 | −0.674 | −0.675 | −0.675 | −0.676 | −0.676 | −0.676 | −0.677 | −0.677 | (0.109) |
| $\beta_7$ | Inv. | 2.074 | 2.057 | 2.049 | 2.043 | 2.040 | 2.038 | 2.036 | 2.035 | 2.035 | 2.034 | (0.248) |
| | Sps. | 2.068 | 2.055 | 2.046 | 2.041 | 2.037 | 2.034 | 2.033 | 2.032 | 2.031 | 2.030 | (0.244) |
| $\beta_8$ | Inv. | −0.770 | −0.755 | −0.747 | −0.742 | −0.739 | −0.737 | −0.736 | −0.735 | −0.735 | −0.734 | (0.134) |
| | Sps. | −0.769 | −0.754 | −0.746 | −0.740 | −0.737 | −0.734 | −0.733 | −0.732 | −0.731 | −0.731 | (0.132) |
| $\beta_9$ | Inv. | −0.170 | −0.158 | −0.151 | −0.147 | −0.145 | −0.143 | −0.142 | −0.141 | −0.141 | −0.141 | (0.115) |
| | Sps. | −0.174 | −0.159 | −0.151 | −0.146 | −0.143 | −0.141 | −0.140 | −0.139 | −0.138 | −0.138 | (0.114) |
| $\beta_{10}$ | Inv. | 0.111 | 0.116 | 0.120 | 0.122 | 0.123 | 0.124 | 0.125 | 0.125 | 0.125 | 0.125 | (0.197) |
| | Sps. | 0.108 | 0.116 | 0.120 | 0.122 | 0.124 | 0.125 | 0.126 | 0.126 | 0.127 | 0.127 | (0.196) |
| $\beta_{11}$ | Inv. | 0.941 | 0.937 | 0.935 | 0.934 | 0.933 | 0.932 | 0.932 | 0.932 | 0.932 | 0.932 | (0.159) |
| | Sps. | 0.938 | 0.936 | 0.934 | 0.933 | 0.932 | 0.932 | 0.931 | 0.931 | 0.931 | 0.931 | (0.157) |
| $\theta$ | Inv. | 0.704 | 0.587 | 0.529 | 0.497 | 0.478 | 0.465 | 0.457 | 0.452 | 0.448 | 0.446 | |
| | Sps. | 0.721 | 0.587 | 0.521 | 0.484 | 0.461 | 0.446 | 0.437 | 0.430 | 0.426 | 0.423 | |
| 계산 | Inv. | 26.360 | 26.380 | 26.380 | 26.470 | 26.550 | 29.570 | 26.560 | 26.600 | 27.940 | 27.760 | |
| 시간 | Sps. | 19.130 | 18.860 | 17.830 | 17.800 | 17.820 | 17.830 | 17.830 | 19.000 | 17.820 | 19.220 | |

표에서 "Inv."은 실제 역행렬을 이용한 계산을 "Sps."는 희소행렬에 기반한 계산을 의미하며 단위는 모두 sec.이다. 또, $\theta$는 프레일티 분산을, "s.e."는 추정량의 표준오차를 의미한다.

후반부에서 살펴보았듯이 헷시안 행렬의 오차인 **M** 자체보다는 헷시안의 역행렬에서의 발생하는 오차를 결정하게 되는 **QMQ** 행렬을 통하여 살펴볼 수 있고 여기서 행렬 **Q**는 랜덤효과의 분산에 정비례하는 항으로 랜덤효과의 분산이 커지면 근사오차 또한 커짐을 예측하여 볼 수 있다. 또한 Therneau와 Grambsch (2000)의 9.7절에 언급된 바와 같이 이 경우 모수추정을 위한 Newton-Raphson절차의 반복횟수가 늘어나는 경향이 있어 계산효율성 또한 떨어지게 된다. 다음으로 일견 행렬 **M**의 비 대각 원소들의 0이 아닌 비율이나 크기가 (헷시안 역행렬에 관한) 근사오차와 정 비례관계가 있을 것이라 생각되나 실제에 있어서는 이런 경향이 잘 나타나지 않음을 관측하였다.

## References

Beckmann, C. F., Jenkinson, M. and Smith, S. M. (2003). General multilevel linear modeling for group analysis in FMRI, *NeuroImage*, **20**, 1052–1063.

Demmel, J. W. (1997). *Applied Numerical Linear Algebra*, SIAM, Philadelphia, PA.

Fleming, T. R. and Harrington, D. P. (2005). *Counting Processes and Survival Analysis*, John Wiley & Sons,

Inc., Hoboken, NJ.

Ha, I. D., Lee, Y. J. and Song, J.-K. (2001). Hierarchical likelihood approach for frailty models, *Biometrika*, **88**, 233–243.

Hager, W. W. (1989). Updating the inverse of a matrix, *SIAM Review*, **31**, 221–239.

Lee, Y. and Oh, H.-S. (2014). A new sparse variable selection via random-effect model, *Journal of Multivariate Analysis*, **125**, 89–99.

Park, S. (2007). *Regression Analysis*, 3/e, Minyoungsa, Seoul.

Pinheiro, J. C. and Bates, D. M. (2000). *Mixed Effects Model in S and S-PLUS*, Springer, New York.

Sohn, S., Chang, I. and Moon, H. (2007). Random effects Weibull regression model for occupational lifetime, *European Journal of Operational Research*, **179**, 124–131.

Therneau, T. M. and Grambsch, P. M. (2000). *Modeling Survival Data: Extending the Cox Model*, Springer, New York.

Yoon, K. and Sohn, S. Y. (2007). Finding the optimal CSP inventory level for multi-echelon system in Air Force using random effects regression model, *European Journal of Operational Research*, **180**, 1076–1085.

Zhu, J. and Hastie, T. (2004). Classification of gene microarrays by penalized logistic regression, *Biostatistics*, **5**, 427–443.

# 희소행렬 계산과 혼합모형의 추론

손원[a] · 박용태[b] · 김유경[c] · 임요한[a,1]

[a]서울대학교 통계학과, [b]서울대학교 산업공학과, [c]서울대학교 의과대학 핵의학교실

(2015년 3월 18일 접수, 2015년 4월 1일 수정, 2015년 4월 1일 채택)

**요 약**

본 연구에서는 혼합모형의 추론을 위한 벌점-최대우도추정량의 빠른 계산절차를 제안하다. 제안된 절차는 벌점-최대우도추정량을 위한 추정방정식에서 헷시안 행렬을 화살촉형태를 지닌 희소행렬을 통하여 근사 시킴으로써 계산속도의 향상을 가져왔다. 두 가지 가상실험을 통하여 제안된 근사식을 사용함으로써 얻게되는 계산시간의 감소와 동시에 이를 위하여 지불하여야 하는 근사오차에 대하여 살펴보았다.

주요용어: 벌점-최대우도추정량, 화살촉행렬, 혼합모형, 희소행렬.

이 논문은 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. 2011-0030811).

[1]교신저자: (151-742) 서울특별시 관악구 관악로 1, 서울대학교 통계학과. E-mail: johanlim@snu.ac.kr