한국통계학회논문집 제3권 1호
1996년 4월 pp. 219-227

# 역추정에서 추가된 독립변수의 효과[1)]

박 래 현[2)], 이 석 훈[2)], 이 상 호[3)]

## 요 약

역추정 문제에서 독립변수가 하나 더 추가되었을때 나타나는 효과가 각 변수간의 상관계수들 및 이들의 함수로 계산되는 결정계수에 따라 어떻게 달라지는지를 살펴보았다. 연구방법으로는 베이지안 방법을 택하였고, 계산상 어려움을 극복하기 위해 Gibbs sampling 방법 및 Mean Field Annealing 방법을 도입하였다.

## 1. 서 론

인간은 항상 미래를 예측하면서 현재를 살고 있으며 통계학에 있어서도 미래 예측이 최종 목표중의 하나라고도 할 수 있을 것이다. 예측의 한 분야에 해당되는 역추정(Calibration, 교정, 보정)은 종속 변수 $y$와 독립 변수 $x$ 사이의 관계를 설명해 주는 회귀 방정식을 바탕으로 하는데, 보통 회귀분석에서 실시하는 주어진 독립 변수 $x_0$에서 종속 변수 $y_0$의 값을 예측하는 것과는 반대로 확률 변수인 $y_0$가 주어졌을 때 확률 변수가 아닌 $x_0$을 예측하는 분야로 통계학자를 포함한 많은 학자들의 연구 대상이 되어 왔으며 의학, 화학, 식품학, 공학 등 다수의 응용 분야를 갖고 있다.

역추정 문제에서는 여러 개의 종속 변수(확률 변수임)들을 바탕으로 여러 개의 확률 변수가 아닌 독립 변수들 $\underline{x}_0$을 예측(예를 들어 밀의 NIR(Near Infrared Reflectance) 측정치들을 가지고 밀의 수분과 단백질의 비율을 예측)하는 다변량 역추정에 대한 연구의 비중이 매우 크다고 생각되는데, 그 중에서 종속 변수를 1개에서 2개로 늘렸을 때 예측하고자 하는 한 개의 독립 변수에 대한 효과는 박래현과 이석훈(1990)에 의해 연구된 바 있다. 본 연구에서 다루려고 하는 것은 위와는 반대로 하나의 종속 변수 $y$와 하나의 독립 변수 $x_1$에 관한 모형으로부터 주어진 $y_0$에서 $x_{10}$을 예측하는 것과 하나의 종속 변수 $y$와 두 개의 독립 변수 $x_1$, $x_2$에 관한 모형에서 $x_{10}$을 예측하는 것 과의 차이점이다. 즉 $x_2$가 모형에 더 추가되었을 경우 $x_{10}$의 예측에 미치는 효과라 할 수 있고 본 효과만 알면 $x_{10}$의 예측을 $x_1$, $y_1$만을 바탕으로 해야 할 지 또는 $x_1$, $x_2$, $y$를 가지

1) 이 논문은 1994년도 학술진흥재단의 공모과제 연구비에 의하여 연구되었음.
2) (305-764) 대전광역시 유성구 궁동 220 충남대학교 통계학과 교수.
3) (301-010) 대전광역시 중구 대흥동 344 대전고등학교.

고 예측함이 좋은지를 판단하는데 큰 도움이 되리라 생각된다.

이에 대한 연구로서 비베이지안(Non-Bayesian) 방법은 효과를 수리적으로 파악할 수 없기 때문에 베이지안 방법을 도입하였고, 베이지안 방법도 뒤에서 논의한 바와 같이 $\underline{x}_0=(x_{10}, x_{20})'$의 사전 분포를 비정보적으로 줄 수 없다. 물론 $x_1$과 $x_2$를 종속 변수로, $y$를 독립변수로 보고 회귀 분석으로 접근하면 $\underline{x}_0$의 predictive distribution을 통해서 해결될 수 있지만 이는 잘못된 모형을 바탕으로 하는 점이 문제이기 때문에 연구 방법에서 배제하였다.

결론적으로 우리가 사용한 방법은 보통의 다변량 선형 회귀 모형하에서 $\underline{x}_0$의 사전 분포를 정보적인 정규분포로 주고 수행한 베이지안 접근인데, $x_2$가 더 추가되었을 때 $x_{10}$의 예측에 영향을 줄 수 있는 요소들로는 $y$와 $x_1$의 상관계수 $r_1$, $y$와 $x_2$의 상관계수 $r_2$, $x_1$과 $x_2$의 상관계수 $r_3$ 및 이로 부터 계산되어 나오는 결정계수( $R_2^2$)로 보고 이들이 복합적으로 어떻게 작용하는가를 살펴 보았다.

## 2. 베이지안 다변량 역추정(Gibbs sampling 접근)

(1) 다변량 Calibration 문제

본 연구를 위해 도입하려고 하는 것은 다음과 같이 흔히 사용하는 다변량 Calibration 모형이다.

$$\begin{aligned} \underline{y}_i &= \underline{\alpha}+B'\,\underline{x}_i+\underline{\varepsilon}_i \quad , \quad i=1,\cdots\cdots,n \\ \underline{y}_{0j} &= \underline{\alpha}+B'\,\underline{x}_0+\underline{\varepsilon}_{0j} \quad , \quad j=1,\cdots\cdots,k \end{aligned} \qquad \cdots\cdots\cdots(2.1)$$

위에서 $\underline{y}_i$, $\underline{y}_{0j}$는 $p$x1 확률벡터, $\underline{\alpha}$와 $B$는 각각 px1, qxp 회귀계수 행렬, $\underline{x}_i$와 $\underline{x}_0$은 수학적 변수인 qx1벡터이고, 오차항 $\underline{\varepsilon}_i$, $\underline{\varepsilon}_{0j}$는 $N_p(\underline{0}, \Sigma)$($\Sigma$는 미지)를 따르는 IID인 확률변수라 가정한다.

모형 (2.1)을 다음과 같이 다시 쓸 수도 있다.

$$Y = ZF + E$$

$$Y_0 = Z_0F + E_0$$

여기서

$$Y=\begin{pmatrix} \underline{y}_1' \\ \underline{y}_2' \\ \cdot \\ \cdot \\ \cdot \\ \underline{y}_n' \end{pmatrix}, \qquad Z=\begin{pmatrix} 1 & \underline{x}_1' \\ 1 & \underline{x}_2' \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & \underline{x}_n' \end{pmatrix}, \qquad F=\begin{pmatrix} \underline{\alpha}' \\ B \end{pmatrix}$$

$$Y_0 = \begin{pmatrix} y_{01}' \\ y_{02}' \\ \cdot \\ \cdot \\ \cdot \\ y_{0k}' \end{pmatrix}, \qquad Z_0 = \begin{pmatrix} 1 & x_0' \\ 1 & x_0' \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & x_0' \end{pmatrix}, \quad E = \begin{pmatrix} \varepsilon_1' \\ \varepsilon_2' \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_n' \end{pmatrix}, \quad E = \begin{pmatrix} \varepsilon_{01}' \\ \varepsilon_{02}' \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_{0k}' \end{pmatrix}$$

이다.

다변량 역추정 문제란 데이타인 ( $x_i$, $y_i$), $i=1$, $\cdots$, $n$과 $y_{0j}$, $j=1$, $\cdots$, $k$를 바탕으로 미지의 $x_0$을 추론하는 것으로 제안된 통계적 방법은 고전적인 비베이지안 방법과 베이지안 방법으로 크게 나눌 수가 있다. 주 논의 대상인 구간 추정시 비베이지안 방법은 예측하여야할 모수 $x_0$이 여러 통계량에 혼합되어 간접적으로 다루어져서 정확한 신뢰 수준을 줄 수 없고 폐쇄형 신뢰 영역도 특정한 조건하에서만 얻어지는 반면에 베이지안 방법은 모수 $x_0$의 사후 분포를 사용하므로 모수를 직접적으로 다룬 것이 되어 비베이지안 방법의 단점을 극복한 것이라 할 수 있지만 베이지안 방법의 핵심인 사후 분포의 계산에 있어 어려움이 뒤따른다.

독립변수를 하나 더 추가시켰을 때 $x_{10}$의 예측에 미치는 효과를 연구시 어려웠던 점은 비베이지안 방법은 위에서 언급한 바와 같은 한계 때문에 제외시켰고 베이지안 방법을 택하였는데 이 방법 역시 ($x_{10}$, $x_{20}$)의 결합 사분포에서 $x_{10}$의 주변 사후분포를 계산하는 어려움의 문제는 갖고 있었으나 이를 Gibbs sampling 방법을 도입하여 극복하였다.

베이지안 적분 계산시 도입한 알고리즘인 Gibbs sampling 방법은 Geman과 Geman (1984)에 의해 소개되었고 통계학에서는 Gelfand와 Smith(1990)가 처음 도입한 계산 방법으로 박래현과 et al(1994)에 소개되어 있다.

(2) 사후 분포의 계산(Gibbs sampling 방법)

먼저 사전 분포를 어떻게 주느냐에 대해 살펴 보기로 하자. 사전 분포를 보통의 비정보적인 $p(x_0, F, \Sigma) \propto |\Sigma|^{-\frac{1}{2}(p+1)}$로 주면 뒤에서 알 수 있듯이 $p<q$인 경우 (우리가 다루려고 하는 것은 $p=1$, $q=2$로 $p<q$인 경우임) 사후 분포의 계산이 불가능하게 된다. 그래서 우리는 사전 분포를 $x_0$에 대해서는 정보적인 $N_q(\tau, V)$($\tau$, $V$는 기지)로 $F$와 $\Sigma$는 비정보적인 $p(F, \Sigma) \propto |\Sigma|^{-\frac{1}{2}(p+1)}$로 주고 $x_0$와 ($F$, $\Sigma$)는 사전적으로 서로 독립이라 가정하겠다.

위와 같은 사전 분포하에서 $x_0$의 사후 분포를 Gibbs sampling 방법으로 계산하려면 $x_0$, $F$, $\Sigma^{-1}$($\Sigma$보다 $\Sigma^{-1}$가 random variates 생성에 더 적합하므로 $\Sigma^{-1}$를 먼저 생성한 후 그것의 역행렬로 $\Sigma$를 생성하려함) 각각에 대해 FCD(Full Conditional Distribution)가 필요한데 이는 아래와 같다.(증명은 부록에 있음.)

$\underline{x}_0$, $F$, $\Sigma^{-1}$의 FCD

$p(\underline{x}_0|F, Y, Y_0, \Sigma)$는 $N_q(\hat{\underline{x}}_0, (V^{-1}+kB\Sigma^{-1}B')^{-1})$의 밀도함수,

$p(F \mid \underline{x}_0, Y, Y_0, \Sigma)$는 $N(\hat{F}, (Z'Z)^{-1}\otimes\Sigma)$의 밀도함수이고,

$p(\Sigma^{-1}|F, Y, Y_0, \underline{x}_0)$는 $W_p(n+k, D^{-1})$의 밀도함수이다.

여기서

$$\hat{\underline{x}}_0 = (V^{-1}+kB\Sigma^{-1}B')^{-1}(V^{-1}\underline{\tau}+kB\Sigma^{-1}(\bar{\underline{y}}_0-\underline{a})),$$

$$\bar{\underline{y}}_0 = k^{-1}\sum_{j=1}^{k}\underline{y}_{0j}, \quad \hat{F}=(Z'Z)^{-1}Z'Y,$$

$$D = (Y-Z\hat{F})'(Y-Z\hat{F})+(Y_0-Z_0\hat{F})'(Y_0-Z_0\hat{F}),$$

$W_p(\cdot, \cdot)$는 Wishart분포를 나타내고, $H\otimes G$는 두 행렬 $H$, $G$의 직적(direct product)이다.

위에서 한가지 언급할 사항은 $\underline{x}_0$의 사전 분포를 비정보적으로 잡으면 $(V^{-1}+kB\Sigma^{-1}B')^{-1}$ 대신 $(kB\Sigma^{-1}B')^{-1}$가 필요한데 $p<q$인 경우에는 존재하지 않으므로 우리가 다루려는 $p=1$, $q=2$인 경우는 $\underline{x}_0$의 사전 분포를 비정보적으로 잡을 수 없다는 것이다.

$\underline{x}_0$의 사후 분포 계산을 위한 모든 FCD가 구해졌으므로 다음은 Gibbs sampling방법으로 $\underline{x}_0$의 사후 분포를 추정하기로 하자. 초기치로 ( $\underline{x}_0^{(0)}$, $F^{(0)}$, $\Sigma^{(0)}$ )를 주어 i번째 까지의 iteration을 m번씩 반복하여 얻어진 생성값들을

$$(\underline{x}_{0j}^{(i)}, F_j^{(i)}, \Sigma_j^{(i)}), \quad j = 1, 2, \cdots, m$$

으로 표현하면 $p(\underline{x}_0|data)$는 아래와 같이 추정된다.

$$\hat{p}(\underline{x}_0|data)=\frac{1}{m}\sum_{j=1}^{m}p(\underline{x}_0|F_j^{(i)}, \Sigma_j^{(i)}, Y, Y_0) \qquad \cdots\cdots (2.2)$$

위에서 $p(\underline{x}_0|F_j^{(i)}, \Sigma_j^{(i)}, Y, Y_0)$은 $N_q(\hat{\underline{x}}_{0j}^{(i)}, (V^{-1}+k B_j^{(i)}\Sigma_j^{-1(i)}B_j^{(i)\prime})^{-1})$의 밀도함수, $\hat{\underline{x}}_{0j}^{(i)} = (V^{-1}+k B_j^{(i)}\Sigma_j^{-1(i)}B_j^{(i)\prime})^{-1}(V^{-1}\underline{\tau}+kB_j^{(i)}\Sigma_j^{-1(i)}(\bar{\underline{y}}_0-\underline{a}_j^{(i)}))$이다.

이와같은 Gibbs sampling방법의 장점중 하나는 기존의 방법으로는 힘들었던 $\underline{x}_0$의 요소중 관심있는 r번째 요소 $x_{r0}$의 사후분포를 추정하는 것으로 이는 아래와 같다.

$$\hat{p}(x_{r0}|data)=\frac{1}{m}\sum_{j=1}^{m}p(x_{r0}|F_j^{(i)}, \Sigma_j^{(i)}, Y, Y_0) \qquad \cdots\cdots (2.3)$$

여기서 $p(x_{r0}|F_j^{(i)}, \Sigma_j^{(i)}, Y, Y_0)$은 $N(\hat{x}_{r0j}^{(i)}, A_{rr})$의 밀도 함수, $\hat{x}_{r0j}^{(i)}$는 $\hat{\underline{x}}_{0j}^{(i)}$의 r번째 원소이고, $A_{rr}$은 $(V^{-1}+k\,B_j^{(i)}\Sigma_j^{-1(i)}B_j^{(i)\prime})^{-1}$의 r번째 대각원소를 나타낸다.

## 3. 효과를 알아보기 위한 절차

(1) 모형 및 사전 분포

우리가 비교하려는 모형은 2절에서 $p=q=1$과 $p=1,\ q=2$의 경우로 다음과 같다.

$$\begin{cases} y_i = w_0+w_1x_{1i}+v_i, & i=1,\ \cdots,\ n \\ y_{0j} = w_0+w_1x_{10}+v_{0j}, & j=1,\ \cdots,\ k \end{cases} \quad \cdots\cdots\cdots(3.1)$$

$$\begin{cases} y_i = \beta_0+\beta_1x_{1i}+\beta_2x_{2i}+\varepsilon_i, & i=1,\ \cdots,\ n \\ y_{0j} = \beta_0+\beta_1x_{10}+\beta_2x_{20}+\varepsilon_{0j}, & j=1,\ \cdots,\ k \end{cases} \quad \cdots\cdots\cdots(3.2)$$

위에서는 연구에 영향을 미치지 않기 때문에 $y,\ x_1,\ x_2$는 전부 표준화시킨 것으로 하였으며 오차항 $v$와 $\varepsilon$의 분산을 각각 $\varphi^2,\ \sigma^2$으로 표기하기로 하자.

(3.1)에서 $x_{10}$의 사전분포를 $N(\tau_1,\ \sigma_1^2)$로 줄때 $\tau_1$은 $x_{10}$의 사후분산(추가된 독립변수의 효과를 사후분산의 크기를 기준으로 하였음)에 영향을 미치지 않으므로 $\tau_1$=0로 하였고 $\sigma_1$은 1과 100의 두가지로 주어졌으며, (3.2)에서 $\underline{x}_0=(x_{10},\ x_{20})'$의 사전 분포로는

$$N_2\left(\begin{pmatrix}0\\0\end{pmatrix},\ \begin{pmatrix}\sigma_1^2 & r_3\sigma_1\sigma_2\\ r_3\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\right)$$

로 하고 $(\sigma_1^2,\ \sigma_2^2)$을 (1, 1)과 (100, 100)의 두 종류로 주어 추가된 독립변수의 효과에 대해 알아보았다. 사전 분산을 1로 잡은 것은 $x_1,\ x_2$를 표준화시켰기 때문이고, 100으로 한 것은 비정보적인 경우로 가깝게 잡아 비정보적인 경우를 유추해 보기 위함이다.

우리가 목적하는 것은 (3.1)에서 (2.2)로 $x_{10}$의 사후분포 $\hat{p}_1(x_{10}|data)$와 (3.2)에서 (2.3)으로 주변사후분포 $\hat{p}_2(x_{10}|data)$를 추정하고, 이 두 개의 추정된 사후분포에서 계산된 두 개의 사후분산 $V_1=Var_1(x_{10}|data),\ V_2=Var_2(x_{10}|data)$의 크기를 비교하여 추가된 $x_2$가 $x_{10}$의 예측에 미치는 효과를 알아보는 것이다.

(2) $x_2$의 생성(MFA) 및 비교과정

두 사후분산 $V_1$, $V_2$를 앞에서 언급한 세 종류의 상관계수 $r_1$, $r_2$, $r_3$와 $R_2^2$의 함수로 표시하여 크기를 대수적으로 비교하는 것은 불가능하므로 생각한 방법이 다음과 같다.

먼저 모형(3.1)에서 $|r_1|$이 큰 것(0.98정도), 중간 것(0.82정도), 낮은 것(0.57정도)의 세 종류의 데이타를 택하고(이렇게 하면 모형 (3.1)의 결정계수 $R_1^2 = r_1^2$가 높은 것, 중간 것, 낮은 것의 세 종류를 택한 것이 되는데, 데이타는 Brown(1982)의 Wheat quality data와 Chatterjee와 Price(1977)의 television data에서 가져왔음), $r_2$, $r_3$는 각각 -0.9에서 0.9까지 0.1의 간격으로 변화시켜 가면서 주어진 $r_2$, $r_3$를 만족시키는 표준화된 독립변수 $x_2$의 생성을 아래와 같이 하였다.

$x_{2i}$, $i=1, \cdots, n$는 $\frac{1}{n}\sum_{i=1}^{n} y_i x_{2i} = r_2$, $\frac{1}{n}\sum_{i=1}^{n} x_{1i} x_{2i} = r_3$, $\sum_{i=1}^{n} x_{2i}^2 = n$, $\sum_{i=1}^{n} x_{2i} = 0$ 을 만족하도록 만들어 내면 되는데 이는 결국

$\psi = (\frac{1}{n}\sum_{i=1}^{n} y_i x_{2i} - r_2)^2 + (\frac{1}{n}\sum_{i=1}^{n} x_{1i} x_{2i} - r_3)^2 + (\sum_{i=1}^{n} x_{2i}^2 - n)^2 + (\sum_{i=1}^{n} x_{2i})^2$ 의 값이 0이 되게 생성하게 될 것이다. $\psi$의 최소값인 0을 만족하는 $x_{2i}$, $i=1, \cdots, n$을 찾기 위한 알고리즘으로는 컴퓨터 분야의 인공지능에서 사용하는 MFA(Mean Field Annealing) 방법을 택하였다.

이는 $r_2$, $r_3$에 따른 수많은 종류의 데이타 생성을 위해 한 것인데($x_1$, $x_2$는 모두 수학적 변수(controlled variable)이어서 보통하는 Normal random variates로 생성하면 안됨), $(y, x_1, x_2)$의 상관계수 행렬

$$\rho = \begin{pmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_3 \\ r_2 & r_3 & 1 \end{pmatrix}$$

이 양정치이어야 하므로 양정치 행렬의 조건인 $|\rho| = 1 + 2r_1r_2r_3 - (r_1^2 + r_2^2 + r_3^2)$이 0보다 크게 되는 $(r_2, r_3)$들만 택해서 $x_{2i}$를 생성하였으며 이 조건하에서 $r_2$, $r_3$에 따른 모형(3.2)의 결정계수$(R_2^2)$ 및 $r_2$, $r_3$의 변화에 따라 $x_{10}$의 예측에 미치는 효과가 어떻게 변화하는지도 살펴보았다.

## 4. 시뮬레이션 결과 및 토의

대수적으로 추가된 독립변수$(x_2)$의 효과를 알아보기 어려워 앞에서 언급한 $r_1$에 따른세가지 예를 바탕으로 $x_2$의 자료를 MFA로 생성한 후 iteration수를 i=50, 반복수 m을 100 ( i, m의 선택

에 대해서는 현재 여러 학자들에 의해 연구가 진행되고 있는 것으로 알고 있음)으로 한 Gibbs sampling 방법으로 $V_1$, $V_2$를 추정해 효과를 분석한 결과를 다음처럼 요약할 수 있다. 효과는 $E=\frac{V_1-V_2}{V_1}*100(\%)$로 살펴보았는데 이것의 의미는 $x_2$를 도입함으로써 감소된 양의 $V_1$에 대한 비율을 뜻한다고 볼 수 있고 $y_0$의 값은 $V_1$, $V_2$에 전혀 영향을 주지 않으므로 어느 값으로 주어도 상관없다.

먼저 $x_{01}$, $\underline{x}_0$의 사전 분포를 각각 $N(0,\ 1)$, $N_2\left(\begin{pmatrix}0\\0\end{pmatrix},\ \begin{pmatrix}1 & r_3\\ r_3 & 1\end{pmatrix}\right)$로 한 경우 세가지 예에서 모두 $r_2$, $r_3$의 값으로는 효과를 설명하기 어렵고 대신 $R_2^2$ 의 값이 클수록 효과가 큰 것으로 나타났는데 기타 중요한 사항을 요약한 표가 아래와 같다.

| | $V_1$ | $V_2$의 범위 | E의 최대값 | E의 최대값일때 ( $r_2$, $r_3$) |
|---|---|---|---|---|
| $R_1^2=0.94456$ ( $r_1=-0.98$ )의 예 | 0.070629 | 0.061-0.05156 | 27.0% | (-0.8, 0.7)<br>(0.8, -0.7) |
| $R_1^2=0.6778$ ( $r_1=-0.82$ ) 〃 | 0.35649 | 0.405-0.323 | 9.4% | (-0.9, 0.5)<br>(0.9, -0.5) |
| $R_1^2=0.325$ ( $r_1=0.57$ ) 〃 | 0.70729 | 0.7420-0.6801 | 3.84% | (-0.6, 0.3)<br>(0.6, -0.3) |

위 표 및 표에 나타나지 않은 결과를 살펴보면 $x_2$의 효과는 $|r_1|$이 클수록, $R_2^2$(독립변수를 추가시켰을 때 결정계수)가 클수록 효과가 크게 나타난 것으로 되어 있어서 어느 경우에 $x_2$를 추가시키면 좋은지를 쉽게 판단할 수 있다.

결론적으로 $x_2$의 효과는 앞의 사전분포하에서 $r_2$, $r_3$에는 관계없이 $R_2^2$의 증가와 $r_1$에 전적으로 달려 있는 것으로 나타났고, $r_1$이 작으면 $x_2$를 하나 더 첨가해도 큰 효과를 볼 수 없는 것으로 되어 있다.

$x_{10}$, $\underline{x}_0$의 사전분포를 각각 비정보적인 것에 가까운 $N(0,\ 100)$, $N_2\left(\begin{pmatrix}0\\0\end{pmatrix},\ \begin{pmatrix}100 & 100r_3\\ 100r_3 & 100\end{pmatrix}\right)$으로 잡으면 $x_2$를 추가할 경우는 항상 $V_2$가 $V_1$보다 크게 나타나

효과가 전혀 없는 것으로 나타나 사전 분포에 따라 효과가 큰 차이를 보여 일반화에 어려운 점이 아쉬운 면으로 남아 있다.

## 참 고 문 헌

[1] 박래현, 이석훈, 이낙영, 박영옥, 이상호(1994). 중도절단 자료에서의 역추정의 문제, 「응용통계연구」 제7권 제1호, 1-17.

[2] Brown, P. J.(1982). Multivariate Calibration, *Journal of the Royal Statistical Society*, B, 44, 287-321.

[3] Chatterjee, S., and Price, B.(1977). Regression Analysis by Example, John Wiley & Sons, New York.

[4] Gelfand, A. E., and Smith, A. F. M.(1990). Sampling Based Approaches to Calculating Marginal Densities, *Journal of the American Statistical Association*, 85, 398-409.

[5] Geman, S., and Geman, D.(1984). Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration Images, *IEEE Transactions on Pattern Analysis and Machince Intelligence*, 6, 721-741.

[6] Park, N. H., and Lee, S. H.(1990). A Baysian Analysis in Multivariate Bioassay and Multivariate Calibration, *Journal of the Korean Statistical Society*, 19, 71-79.

## 부 록

$\underline{x}_0$과 Y는 독립이고, $\underline{x}_0$의 충분통계량은 $\bar{\underline{y}}_0 = \frac{1}{k}\sum_{j=1}^{k} \underline{y}_{oj}$이며 $\bar{\underline{y}}_0$은 $N_q(\underline{\alpha}+B'\underline{x}_0,\ \frac{1}{k}\Sigma)$의 분포이므로

$$
\begin{aligned}
p(\underline{x}_0|F,\ \Sigma,\ Y,\ Y_0) &= p(\underline{x}_0|F,\ \Sigma,\ Y_0) \\
&\propto p(\underline{x}_0)p(\bar{\underline{y}}_0|F,\ \Sigma,\ x_0) \\
&\propto \exp\left[-\frac{1}{2}(\underline{x}_0-\underline{\tau})'V^{-1}(\underline{x}_0-\underline{\tau})\right]\cdot\exp\left[-\frac{k}{2}(\bar{\underline{y}}_0-\underline{\alpha}-B'\underline{x}_0)'\Sigma^{-1}(\bar{\underline{y}}_0-\underline{\alpha}-B'\underline{x}_0)\right] \\
&\propto \exp\left[-\frac{1}{2}\{(\underline{x}_0-\widehat{\underline{x}_0})'(V^{-1}+kB\Sigma^{-1}B')(\underline{x}_0-\widehat{\underline{x}_0})\}\right]
\end{aligned}
$$

따라서 $p(\underline{x}_0|F,\ \Sigma,\ Y,\ Y_0)$은 $N_q(\widehat{\underline{x}_0},\ (V^{-1}+kB\Sigma^{-1}B')^{-1})$의 밀도함수이다.

F는 ( $x_0$, $Y_0$)와 독립이므로

$$p(F \mid x_0, \Sigma, Y, Y_0) = p(F \mid \Sigma, Y) \propto p(F)p(Y \mid F, \Sigma)$$

$$\propto p(Y \mid F, \Sigma) \propto \exp\left[ -\frac{1}{2} tr\{(Y-ZF)'(Y-ZF)\Sigma^{-1}\}\right]$$

위에서

$(Y-ZF)'(Y-ZF) = (Y-Z\hat{F})'(Y-Z\hat{F}) + (F-\hat{F})'Z'Z(F-\hat{F})$ 이므로

$p(F \mid \Sigma, Y) \propto \exp\left[ -\frac{1}{2} tr\{(F-\hat{F})'Z'Z(F-\hat{F})\Sigma^{-1}\}\right]$ 이다.

그러므로 $p(F \mid x_0, \Sigma, Y, Y_0)$은 $N(\hat{F}, (Z'Z)^{-1} \otimes \Sigma)$의 밀도함수이다.

$x_0$와 F가 주어지면 $\Sigma$의 충분통계량 D는 Wishart분포 $W_p(n+k, \Sigma)$를 따르므로

$$p(\Sigma \mid F, x_0, Y, Y_0) = p(\Sigma \mid D) \propto p(\Sigma)p(D \mid \Sigma)$$

$$\propto |\Sigma|^{-\frac{1}{2}(n+k+p+1)} \exp\left[ -\frac{1}{2} tr(\Sigma^{-1}D)\right]$$

$H = \Sigma^{-1}$로 변수변환하면 Jacobian이 $|J| = |H|^{-(p+1)}$ 이므로

$$p(H \mid D) \propto |H|^{\frac{1}{2}(n+k+p+1)} \exp\left[ -\frac{1}{2} tr(HD)\right] |H|^{-(p+1)}$$

$$\propto |H|^{\frac{1}{2}(n+k-p-1)} \exp\left[ -\frac{1}{2} tr(HD)\right]$$

따라서 $p(\Sigma^{-1} \mid x_0, F, Y, Y_0)$은 $W_p(n+k, D^{-1})$의 밀도함수이다.