김일성종합대학학보 주체109(2020)년
수 학 제66권 제3호

인자사영에 관한 D－유효초포화계획의 구성방법

김철호, 김정만

경애하는 최고령도자 **김정은**동지께서는 다음과 같이 말씀하시였다.

《현시대는 과학기술의 시대이며 과학기술의 발전수준은 나라의 종합적국력과 지위를 규정하는 징표로 됩니다.》

최근에 실험회수를 줄이기 위한 실험계획적연구가 심화되고있으며 균형성과 D－최량성을 보장하는 초포화계획에 대한 연구가 많이 진행되고있다.

선행연구[1－3]에서는 1차회귀모형에 대한 $E(s^2)$－최량초포화계획의 개념과 여러가지 수법에 의한 몇가지 구성방법들을, 선행연구[4]에서는 $UE(s^2)$－최량초포화계획의 개념과 구성방법들을 연구하였다.

론문에서는 균형성과 최량성을 가지는 인자사영에 관한 D－유효초포화계획의 개념을 제기하고 $m=4t-1$인 경우 인자사영에 관한 D－유효초포화계획의 한가지 구성방법에 대하여 연구하였다.

먼저 $UE(s^2)$－최량초포화계획과 인자사영에 관한 D－유효초포화계획에 대하여 보자.

이제 X_d를 $n\times m$형 2수준 +1, －1을 가지는 초포화계획 d의 계획행렬이라고 하고 $Z_d=[1\ X_d]$를 계획 d에 의한 독립변수행렬이라고 하자.

정의 1 [1] Z_d의 두 렬벡토르 u, v가 $u=v$ 또는 $u=-v$를 만족시킬 때 계획 d를 가상렬을 가진 계획이라고 말한다.

정의 2 [4] 계획 d에 의한 계획행렬 X_d의 매 렬에서 +1과 －1의 개수가 동일할 때 이 d를 수준균형인 계획이라고 부르며 수준균형인 렬에 배당된 인자를 수준균형인 인자라고 부른다. 이 경우는 실험점개수 n이 짝수일 때만 성립된다.

정의 3 [4] 계획 d에 의한 계획행렬 X_d의 매 렬에서 +1과 －1의 개수의 차가 1일 때 d를 거의수준균형인 계획이라고 부르며 거의수준균형인 렬에 배당된 인자를 거의수준균형인 인자라고 부른다.

일반적으로 거의수준균형인 렬들에서는 +1이 $(n-1)/2$번, -1이 $(n+1)/2$번 나타난다.

초포화계획에 대하여 렬들이 모두 서로 직교되는것은 불가능하다.

정의 4 [4] 초포화계획 d에 대하여 X_d의 렬들중에서 비직교성에 대한 단순한 척도 $E(s^2)=\frac{1}{C_m^2}\sum_{i<j}(x_i^{\mathrm{T}}x_j)^2$이 최소인 계획을 $E(s^2)$－최량초포화계획이라고 부른다.

선행연구[4]에서는 수준균형성을 포함하는 다음과 같은 새로운 기준을 제기하였다.

$$UE_d(s^2)=\frac{1}{C_{m+1}^2}\left(\sum_{i=1}^{m}(1^{\mathrm{T}}x_i)^2+\sum_{i<j}(x_i^{\mathrm{T}}x_j)^2\right)$$

이제 주어진 $n,\ m$에 대하여 $D_U(m,\ n)$을 수준균형 또는 거의수준균형제한이 없는 모든 초포화계획들의 모임이라고 하고 $D_R(m,\ n)$은 수준균형 또는 거의수준균형인 모든

초포화계획들의 모임이라고 하자.

정의 5[4] 임의의 $d \in D_U(m,\ n)$에 대하여 $UE_{d^*}(s^2) \le UE_d(s^2)$이 성립될 때 $d^* \in D_U(m,\ n)$을 $UE(s^2)$－최량계획이라고 한다.

한편 m개의 인자들중에서 $f(f \le n-1)$개의 활용인자들의 모임을 F라고 표시하자.

정의 6 임의의 m개의 인자들로 구성된 계획 d에 대하여 활용인자 $f \in F$개의 렬로 구성된 계획을 m개의 인자들에 의한 d의 F에로의 사영계획이라고 부르고 d^F로 표시한다.

X_d^F를 X_d에 대응하는 사영계획 d^F의 계획행렬이라고 하자.

이때 독립변수행렬을 $Z_d^F = [1 X_d^F]$라고 하고 계획 d^F의 정보행렬을 M_{d^F}로 놓으면

$$M_{d^F} = Z_{d^F}^{\mathrm{T}} Z_{d^F} = \begin{bmatrix} n, & 1^{\mathrm{T}} X_{d^F} \\ X_{d^F}^{\mathrm{T}} 1, & X_{d^E}^{\mathrm{T}} X_{d^F} \end{bmatrix}$$

이다.

정의 7 초포화계획 d의 f인자사영에 관한 값

$$D_f(d) = \frac{1}{C_m^f} \sum_{F:|F|=f} D(X_{d^F}) \tag{1}$$

를 D－유효성척도라고 부르며 모든 $f = 1, 2, \cdots, n-1$에 대하여 $D_f(d)$가 최대인 초포화계획을 모임 F우에로 인자사영에 관한 D－유효초포화계획이라고 부른다.

식 (1)에서 합은 f개 인자로 이루어지는 부분모임 F우에서 진행한것이다.

보조정리 1 $D(X_{d^F})$를 최대로 하는 계획을 찾는 문제는 $\mathrm{tr}[M_{d^F}]$를 최대로 하거나 $\mathrm{tr}[M_{d^F}]$가 최대일 때 $\mathrm{tr}[M_{d^F}]^2$을 최소화하는 계획을 찾는 문제와 동등하다.

보조정리 2 어떤 계획에 대한 정보행렬의 모든 고유값들이 1이면 이 계획은 D－최량계획이다.

보조정리 3[2, 3] 임의의 $n \times m$형초포화계획 d가 $E(S^2)$－최량계획이기 위하여서는 $X_d X_d^{\mathrm{T}} = \begin{pmatrix} m & -t & \cdots & -t \\ -t & m & \cdots & -t \\ \vdots & \vdots & \ddots & \vdots \\ -t & -t & \cdots & m \end{pmatrix}$일것이 필요하고 충분하다. 여기서 X_d는 d의 계획행렬이다.

보조정리 4 초포화계획 d^*이 수준균형 또는 거의수준균형계획이고 $E(s^2)$－최량계획이면 계획 d^*은 모임 F우에로 인자사영에 관한 D－유효초포화계획이다.

증명 D－최량계획 d^*에 대하여 $\mathrm{tr}[M_{d^F}] = n(f+1)$은 상수이며 $D_f(d)$를 최대로 하는것은 $\frac{1}{C_m^f} \sum_{F:|F|=1} \mathrm{tr}[M_{d^F}]^2 = \frac{1}{C_m^f} \sum_{F:|F|=1} \mathrm{tr}[Z_{d^F}^{\mathrm{T}} Z_{d^F}]^2$을 최소로 하는 문제에 귀착된다.

한편 크기 m인 모집단에서 반환없이 표본 f개를 단순표본뽑기하는 경우 1차표본이 뽑힐 확률은 f/m이고 2차표본이 뽑힐 확률은 $f(f-1)/m(m-1)$이므로

$$\frac{1}{C_m^f} \sum_{F:|F|=1} \mathrm{tr}[Z_{d^F}^{\mathrm{T}} Z_{d^F}]^2 = \frac{2f}{m}[1^{\mathrm{T}} X_d X_d^{\mathrm{T}} 1] + \frac{f(f-1)}{m(m-1)} \mathrm{tr}[(X_d^{\mathrm{T}} X_d)^2] + A \tag{2}$$

이다. 여기서 A는 상수이고 f/m는 작으며 $f(f-1)/m(m-1)$은 f/m보다 더 작다.

식 (2)를 최소화하는것은 $1^{\mathrm{T}}X_dX_d^{\mathrm{T}}1=ss$와 $\mathrm{tr}[(X_d^{\mathrm{T}}X_d)^2]$을 최소화하는것이다. 여기서 $1^{\mathrm{T}}X_dX_d^{\mathrm{T}}1=ss$를 최소로 하는것은 수준균형 또는 거의수준균형계획이라는것이고 $\mathrm{tr}[(X_d^{\mathrm{T}}X_d)^2]$을 최소로 하는것은 $E_d(s^2)$을 최소로 한다는것이다.(증명끝)

다음 인자사영에 관한 $D-$유효초포화계획의 구성방법을 $m=4t-1$인 경우에 대하여서만 론의하자.

먼저 아다마르행렬에 의한 초포화계획 $d^*(n,\ 4t-1)$의 구성방법에 대하여 론의하자.

이제 H를 첫렬과 첫행이 모두 1인 $4t$차아다마르행렬이 주어졌다고 하자.

이때 계획을 다음과 같이 구성한다.

① H에서 $4t$개의 행들중에서 서로 직교하는 n개 행들을 고르고 나머지 $4t-n$개의 행들을 떼버린다.

② 얻어진 $n\times 4t$형행렬에서 첫렬을 빼버린다.

정리 1 초포화계획 $d^*(n,\ 4t-1)$은 인자사영에 관한 $D-$유효계획이다.

증명 먼저 계획 $d^*(n,\ 4t-1)$이 가상렬을 포함하지 않는다는것을 보기로 하자.

일반성을 잃지 않고 $H^{\mathrm{T}}=[d^{\mathrm{T}}G^{\mathrm{T}}]$, $d^{\mathrm{T}}d=[u_{ij}]$, $G^{\mathrm{T}}G=[w_{ij}]$로 놓자.

이때 $H^{\mathrm{T}}H=(4t)I_{4t}$이므로 모든 $i,\ j$에 대하여 다음의 식이 성립된다.

$$|u_{ij}|=|w_{ij}|,\ |u_{ij}|\le n,\ |w_{ij}|\le 4t-n \tag{3}$$

만일 계획 d가 2개의 가상렬로 i_0째 렬과 j_0째 렬을 가졌다면 $|u_{i_0j_0}|=n$이므로 식 (3)으로부터 $n\le 4t-n$이다. 즉 $n\le 2t$이다. 이것은 $n>2t$인 n에 대하여 H에서 $4t-n$개 행을 떼버린것에 의해 얻어진 계획행렬 d는 가상렬을 포함하지 않는다는것이다.

다음으로 인자사영에 관한 $D-$유효성을 보기로 하자.

계획 d^*에 의한 계획행렬을 X_{d^*}이라고 하면 독립변수행렬은 $Z_d=[1\ X_d]$이며 구성방법에 의하여 이 행렬의 임의의 두 행들은 직교한다. 이때 $Z_{d^*}Z_{d^*}^{\mathrm{T}}=(4t)I_n$이고 $\mathrm{tr}[(Z_{d^*}^{\mathrm{T}}Z_{d^*})^2]=\mathrm{tr}[(Z_{d^*}Z_{d^*}^{\mathrm{T}})^2]=(4t)^2n$이다. 이로부터 $X_{d^*}X_{d^*}^{\mathrm{T}}$는 모든 대각선원소가 $4t-1$과 같으며 모든 비대각선원소는 -1이다. 따라서 보조정리 3에 의하여 계획 $d^*(n,\ 4t-1)$은 $E(s^2)-$최량계획이다.

한편 주어진 $n,\ m=4t-1$에 대하여 $\mathrm{tr}[(X_{d^*}^{\mathrm{T}}X_{d^*})^2]=\mathrm{tr}[(X_{d^*}X_{d^*}^{\mathrm{T}})^2]=n(4t-1)^2+n(n-1)$로서 상수이라는것을 알수 있다. 그리고 $\mathrm{tr}[(Z_dZ_d^{\mathrm{T}})^2]=n^2+2\cdot 1^{\mathrm{T}}X_dX_d^{\mathrm{T}}1+\mathrm{tr}[(X_dX_d^{\mathrm{T}})^2]$이므로 $ss=1^{\mathrm{T}}X_dX_d^{\mathrm{T}}1=\{\mathrm{tr}[(Z_dZ_d^{\mathrm{T}})^2]-\mathrm{tr}[(X_dX_d^{\mathrm{T}})^2]-n^2\}/2$이며 따라서 $ss=1^{\mathrm{T}}X_dX_d^{\mathrm{T}}1$도 일정한 값으로 된다.

결국 보조정리 4와 정의 6에 의하여 결과가 얻어진다.(증명끝)

다음으로 크로네커적에 의한 초포화계획 $d_*(n,\ 4t-1)$의 구성방법에 대하여 론의하자.

$4t$가 2의 제곱 즉 $4t=2^w$이라고 할 때 임의의 $n=w+1$에 대하여 행렬 $D(n,\ 4t)$의 1행은 $\bigotimes_{i=1}^{w}(1,\ 1)$로, 2행은 $(1,\ -1)\bigotimes_{i=2}^{w}(1,\ 1)$로, 3행부터 w행까지는 $\bigotimes_{i=1}^{j}(1,\ 1)\otimes(1,\ -1)\bigotimes_{i=j+2}^{w}(1,\ 1)$로, $w+1$행은 $\bigotimes_{i=1}^{w-1}(1,\ 1)\otimes(1,\ -1)$로 구성한다.

마지막에 얻어진 행렬 $D(n, 4t)$ 에서 첫렬을 제거하면 초포화계획 $d(n, 4t-1)$ 을 구성하게 된다.

정리 2 계획 $d_*(n, 4t-1)$ 은 가상렬을 가지지 않는 인자사영에 관한 D－유효초포화계획이다.

증명 먼저 행렬 $D(n, 4t)$ 가 가상렬을 포함하지 않는다는것을 보자.

선행연구[3]에 의하여 행의 크기가 n 인 1과 -1 을 원소로 가지는 $n\times 4t$ 형행렬에서의 가상렬이 없는 렬의 최대개수는 2^{n-1} 이다.

따라서 $n\times 4t$ 형행렬이 가상렬을 가지지 않자면 $2^{n-1}\geq 4t$ 또는 $n\geq[\log_2 t]+3$ 을 만족시켜야 한다. 여기서 $[z]$는 z 와 같거나 작지 않은 옹근수이다.

$4t=2^w$ 이므로 $w+1=[\log_2 t]+3$ 이며 우의 조건을 만족시키므로 행렬 $D(n, 4t)$ 는 가상렬을 가지지 않는다. 따라서 $d_*(n, 4t-1)$ 은 가상렬을 포함하지 않는다.

한편 구성된 행렬 $D(n, 4t)$ 는 행들이 서로 직교하며 따라서 정리 1에서와 마찬가지로 계획 $d_*(n, 4t-1)$ 의 인자사영에 관한 D －유효성이 얻어진다.(증명끝)

참 고 문 헌

[1] A. Das et al.; J. Statist. Plann. Inference, 138, 3749, 2008.

[2] J. Christopher et al.; Computational Statistics, 54, 3158, 2010.

[3] A. B. Durun et al.; The Annals of Statistics, 32, 4, 1662, 2004.

[4] B. Jones et al.; J. Amer. Statist. Assoc., 109, 1592, 2014.

주체109(2020)년 3월 15일 원고접수

Construction Methods of *D*-Efficiency Supersaturated Designs on Factor Projections

Kim Chol Ho, Kim Jong Man

We introduce concept of *D*-efficiency supersaturated designs on factors projections with balance and optimality and in the case $m=4t-1$ study a construction methods of *D*-efficiency supersaturated designs on factor projections.

Keyword: factor projection