Proceedings of the Autumn Conference in 2001
and the 30th Anniversary of the Korean Statistical Society

그룹분류가능계획를 이용한 최적 블록 CDC의 설계

김 진[1], 배 종 성[2]

요 약

$m=2$ 또는 $n=2$이고, $\lambda_1 < \lambda_2$인 그룹분류가능계획을 매개디자인으로 사용한 완전이면교배가 A-최적, D-최적임을 보였다. 또한, $\lambda_2=\lambda_1+1$이면 일반화된 최적계획이 됨을 보였다.

주요용어 : 그룹분류가능계획, 블록완전이면교배, A-, D-최적계획, 일반화된 최적계획

1. 서론

이면교배(diallel crosses)는 동물 또는 식물 육종학에서 자식대의 근교계통(inbred line)들간의 유전적인 성질을 분석하여 어미대의 유전적인 성질을 연구하는데 사용되는 짝짓기 계획이다.
서로 다른 특성을 갖는 p개의 근교계통이 있을 때, i번째 근교계통과 j번째 근교계통간의 교배를 $(i \times j)$, $i < j = 1, \cdots, p$ 로 나타내고, 실험에 사용되는 교배수를 n_c라 하자.
교배수가 $n_c = p(p-1)/2$인 실험을 완전이면교배(Complete Diallel Cross : CDC)실험이라 하고, CDC에서 일부분의 교배($n=ps/2$, $s < p-1$)만 사용하는 실험을 부분이면교배(Partial Diallel Cross : PDC)실험이라 한다. 근교계통 p가 증가하면 교배수의 증가로 인하여 n_c개의 교배를 동일한 환경에서 동시에 실험하기가 곤란해진다. 이러한 경우는 동일한 조건하에서 실험 가능한 교배를 블록에 나누어 배치함으로써 실험오차를 줄이는 블록화 방법을 사용한다.
완전이면교배의 블록화 방법에 대한 연구로 Agarwal 과 Das(1990)는 균형된 불완비 블록계획(Balanced Incomplete Block Design : BIBD)를 이용하였으며, Divecha 와 Gosh(1994)는 삼각형 PBIBD(Triangular Partially Balanced Incomplete Block Design)을 이용하여 블록 완전이면교배(Blocked CDC)를 설계하였다. Dey 와 Midha(1996)는 삼각형 PBIBD를 이용하여 설계한 블록완전이면교배가 총체적 최적(universal optimality)임을 보였다. Das, Dey 와 Dean(1998)는 특별한 모수 조건을 만족하는 삼각형 PBIBD를 이용하여 설계한 블록완전이면교배는 총체적 최적이 됨을 보였다. 본 논문에서는 $m=2$ 또는 $n=2$이고, $\lambda_1 < \lambda_2$인 그룹분류가능계획을 매개디자인으로 사용하여 CDC를 블록화한 경우 블록 CDC가 A-최적, D-최적계획이 됨을 보였다. 또한, $\lambda_2=\lambda_1+1$이면 일반화된 최적계획(generalized optimal design)이 됨을 보였다.

2. 블록완전이면교배의 모형과 최적화

근교계통의 수가 p인 이면교배를 블록 크기가 K인 B개의 블록에 R번 반복하여 배치하는 블록 이면교배의 모형은 다음과 같다.

---

1) 박사과정, 광주광역시 북구 용봉동 300 전남대학교 통계학과
2) 교수, 광주광역시 북구 용봉동 300 전남대학교 통계학과

$$Y = \mu 1_n + \Delta_1 g + \Delta_2 \beta + \varepsilon$$

여기서, 크기 $n\times 1$인 Y는 관측치 벡터, μ는 전체 평균, 1_n은 모든 원소가 1인 $n\times 1$벡터, 크기 $p\times 1$과 $B\times 1$인 g , β는 p개의 일반조합능력 효과와 B개의 블록 효과를 나타내는 모수벡터이다. n은 실험횟수이다. 크기 Δ_1, Δ_2 는 g 와 β 에 대응하는 빈도행렬(incidence matrix)이고, ε은 평균이 0, 분산이 σ^2인 정규분포를 따른다. 크기 $n\times p$인 행렬 Δ_1의 (u, w)원소는 u번째 교배가 w번째 근교계통을 포함하면 1, 아니면 0이고, $u=1,2,\cdots,n$, $w=1,2,\cdots,p$이다. 크기 $n\times B$ 인 행렬 Δ_2의 (u, w) 원소는 u번째 교배가 w번째 블록에서 나타나면 1, 아니면 0이다. 그리고 $G=\Delta_1'\Delta_1=(G_{ij})$, $\Gamma=\Delta_1'\Delta_2$ 라면 $G_{ij}=R$, $G_{ii}=R(p-1)$ 이다. $\Gamma=(n_{il})$의 n_{il} 은 i번째 근교계통이 l 번째 블록에 나타나는 횟수이다. 일반조합능력, g를 추정하기 위한 $C-$행렬은

$$C= G-\frac{1}{K}\Gamma\Gamma' = (c_{ij}), \qquad i,j=1,2,\cdots,p$$

이고, $C1_p=0$, $rank(C)=p-1$이다.

Cheng(1978)은 최적함수를 다음과 같이 정의하였다.

$$\phi_f(C)=\sum_i^{v-1} f(x_i)$$

여기서, f는 $(0, M_0)$에서 정의되고, x_i값들은 $C-$행렬의 양의 고유치들이다. 또한, M_0는 같은 모수를 갖는 디자인들 중에서 최대인 $tr(C)$값이다. 또한 A-최적과 D-최적계획을 다음과 같이 정의하였다.

A-최적계획은 $f(x)=\frac{1}{x}$이며 같은 모수를 갖는 전체 디자인 중에서 $\sum 1/x_i$를 최소화하는 블록계획을 말한다.

D-최적계획은 $f(x)=-\log x$이며 같은 모수를 갖는 전체 디자인 중에서 $\sum -\log x_i$ 또는 $\prod x_i^{-1}$를 최소화하는 블록계획을 말한다.

(보조정리2.1) $A^2>B>A^2/(v-1)$를 만족하는 고정된 $A=\sum_1^{v-1} x_i$ 와 $B=\sum_1^{v-1} x_i^2$에 대해서, 임의의 함수 f에 대해서, $\phi_f=\sum_i^{v-1} f(x_i)$는 x_i의 값들 중 정확히 하나가 $\{A+\delta(v-2)P\}/(v-1)$이고, 나머지가 $\{A-\delta P\}/(v-1)$인 조건을 만족하면 최소가 된다. 여기서, $P^2=B-A^2/(v-1)$이고 $\delta=\sqrt{(v-1)/(v-2)}$이다.

(보조정리2.2) $D(b,v,k)$내에 $C-$행렬이 다음과 같은 근을 갖는 디자인 d^*가 존재한다고 하자(Cheng, 1978).

$\mu^*=\{A^*+\delta(v-2)P^*\}/(v-1)$: 1개의 중근

$\mu'^*=\{A^*-\delta P^*\}/(v-1)$: $(v-2)$개의 중근

여기서 A^* 와 P^*는 d^*에 대응되는 A와 P이고, $\delta=\sqrt{(v-1)/(v-2)}$이다.

게다가 $D(b,v,k)$중에서 (A,P)의 모든 쌍에 대해

(i) $A^*\geq A$ (ii) $A^*-\delta P^*\geq A-\delta P$

을 만족하면 d^*는 일반화된 최적이다.

3. 그룹분류가능계획을 매개디자인으로 사용한 블록 CDC의 설계

3.1 $m=2$인 그룹분류가능계획을 매개디자인으로 이용한 최적 블록 CDC

블록 CDC를 $m=2$인 그룹분류가능계획을 매개디자인으로 이용하여 설계하는 방법에 대해서 살펴보자. $v=p(p-1)/2$, $b=B$, $k=K$, $r=R$인 그룹분류가능계획을 찾아 그룹분류가능계획의 각 처리를 교배로 바꾸어주면 된다. $v, b, k, r, \lambda_1, \lambda_2, m=2$인 그룹분류가능계획을 매개디자인으로 사용한 블록 CDC의 $C-$행렬은 다음과 같다.

(i). $p=4$인 경우

$$C=\begin{pmatrix} x & y & y & t \\ y & x & y & t \\ y & y & x & t \\ t & t & t & u \end{pmatrix}$$

여기서, $x=3r-\frac{1}{k}(3r+2\lambda_1+4\lambda_2)$, $y=r-\frac{1}{k}(r+4\lambda_1+4\lambda_2)$,

$t=r-\frac{1}{k}(r+2\lambda_1+6\lambda_2)$, $u=3r-\frac{1}{k}(3r+6\lambda_1)$ 이다.

평균효율 E를 구하기 위해 $m=2$인 그룹분류가능계획을 매개디자인으로 사용하여 설계된 블록 CDC의 $C-$행렬의 고유치를 구할수 있다.

$C=\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$로 분할되었다 하자.

$$A_1=\begin{pmatrix} x & y & y \\ y & x & y \\ y & y & x \end{pmatrix} \quad A_2=\begin{pmatrix} t \\ t \\ t \end{pmatrix} \quad A_3=(t \ t \ t) \quad A_4=(u) \text{이다.}$$

$$\begin{aligned}|C| = |A_1||A_4-A_3A_1^{-1}A_2| &= (x-y)^2(x+2y)(u-\frac{3t^2}{x+2y}) \\ &= (x-y)^2(x+2y)\{u(x+2y)-3t^2\}\end{aligned}$$

$C-$행렬은 한 행의 합이 모두 0이므로 $m=2, p=4$인 경우 $C-$행렬은 $x+2y+t=3t+u=0$이 된다. 따라서, $x+2y=u+2t$이다. 따라서,

$|C| = (x-y)^2\{u(u+2t)-3t^2\} = (x-y)^2(u-t)(u+3t)$

이고 $u+3t=0$이므로 고유치는 1개의 $u-t$ 와 2개의 $x-y$ 이다.

$p=4$인 경우의 블록 CDC의 고유치는 $u-t$, $x-y$, $x-y$이므로 $A=(u-t)+2(x-y)$, $B=(u-t)^2+2(x-y)^2$ 이고

$$\begin{aligned}(\delta P)^2 &= \frac{p-1}{p-2}\times\frac{B(p-1)-A^2}{p-1} = \frac{3B-A^2}{2} \\ &= [3\{(u-t)^2+2(x-y)^2\}-\{(u-t)+2(x-y)\}^2]/2 = \{(u-t)-(x-y)\}^2\end{aligned}$$

이다.

(정리3.1) $m=2$이고, $\lambda_1 < \lambda_2$인 그룹분류가능계획을 매개디자인으로 사용한 $p=4$인 블록 CDC은 A-, D- 최적이다.

(증명) 양수인 δP는 $\pm\{(u-t)-(x-y)\}$을 갖는다.

$\delta P=(x-y)-(u-t)$이면 (보조정리2.1)이 성립되지 않고, $\delta P=(u-t)-(x-y)$이면 (보조정리2.1)이 성립된다. $\delta P=(u-t)-(x-y)$를 (보조정리2.1)에 적용하면

$$\frac{A+\delta(p-2)P}{p-1} = \frac{(u-t)+(p-2)(x-y)+(p-2)\{(u-t)-(x-y)\}}{p-1} = u-t$$

$$\frac{A-\delta P}{p-1} = \frac{(u-t)+(p-2)(x-y)-\{(u-t)-(x-y)\}}{p-1} = x-y$$

이다. 따라서, $\delta P=(u-t)-(x-y)>0$인 경우에만 A-, D-최적이 된다.

$\delta P=(u-t)-(x-y)>0$

$$(u-t)-(x-y) = 2r-\frac{1}{k}(2r+4\lambda_1-6\lambda_2)-2r+\frac{1}{k}(2r-2\lambda_1)=6(\lambda_2-\lambda_1)>0$$

이므로 $\lambda_1<\lambda_2$ 이다.

따라서, $m=2$이고, $\lambda_1<\lambda_2$를 만족하는 그룹분류가능계획을 매개디자인으로 사용한 $p=4$인 블록 CDC는 A-, D-최적이다. □

(ii). $p=5$인 경우

$$C=\begin{pmatrix} x & y & y & t & t \\ y & x & y & t & t \\ y & y & x & t & t \\ t & t & t & u & w \\ t & t & t & w & u \end{pmatrix}$$

여기서, $x=4r-\frac{1}{k}(4r+4\lambda_1+8\lambda_2)$, $y=t=r-\frac{1}{k}(r+7\lambda_1+8\lambda_2)$,

$u=4r-\frac{1}{k}(4r+6\lambda_1+6\lambda_2)$, $w=r-\frac{1}{k}(r+5\lambda_1+10\lambda_2)$ 이다.

$$A_1=\begin{pmatrix} x & y & y \\ y & x & y \\ y & y & x \end{pmatrix} \quad A_2=\begin{pmatrix} t & t \\ t & t \\ t & t \end{pmatrix} \quad A_3=\begin{pmatrix} t & t & t \\ t & t & t \end{pmatrix} \quad A_4=\begin{pmatrix} u & w \\ w & u \end{pmatrix}$$이다.

$$\begin{aligned} |C| = |A_1||A_4-A_3A_1^{-1}A_2| &= (x-y)^2(u-w)(x+2y)(u+w-\frac{6t^2}{x+2y}) \\ &= (x-y)^2(u-w)\{(u+w)(x+2y)-6t^2)\} \end{aligned}$$

이다. $C-$행렬의 한 행 또는 열의 합은 모두 0이므로, $m=2, p=5$인 경우 $C-$행렬은 $x+2y+2t=3t+u+w=0$이고 $y=t$이므로, $x+y=u+w$이다. 따라서,

$|C| = (x-y)^2(u-w)\{(x+y)(x+2y)-6y^2)\} = (x-y)^3(u-w)(x+4y)$

이고, $x+4y=0$이므로 $C-$행렬의 고유치는 1개의 $u-w$와 3개 $x-y$이다.

$p=5$인 경우의 블록 CDC의 고유치는 $u-w$, $x-y$, $x-y$, $x-y$이므로

$A=(u-w)+3(x-y)$, $B=(u-w)^2+3(x-y)^2$이고

$$\begin{aligned} \{\delta P\}^2 &= \frac{p-1}{p-2}\times\frac{B(p-1)-A^2}{p-1} = (4B-A^2)/3 \\ &= [4\{(u-w)^2+3(x-y)^2\}-\{(u-w)+3(x-y)\}^2]/3 = [(u-w)-(x-y)]^2 \end{aligned}$$

이다.

(정리3.2) $m=2$이고, $\lambda_1<\lambda_2$인 그룹분류가능계획을 매개디자인으로 사용한 $p=5$인 블록 CDC은 A-, D- 최적이다.

(증명) δP는 양수이고 $\pm\{(u-w)-(x-y)\}$을 갖는다.

$\delta P=(x-y)-(u-w)$이면 (보조정리2.1)을 만족하지 않고, $\delta P=(u-w)-(x-y)$이면 (보조정리2.1)을 만족한다. $\delta P=(u-w)-(x-y)$를 (보조정리2.1)에 적용하면

$$\frac{A+\delta(p-2)P}{p-1} = \frac{[(u-w)+(p-2)(x-y)+(p-2)\{(u-w)+(x-y)\}]}{p-1} = u-w$$

$$\frac{A-\delta P}{p-1} = \frac{[(u-w)+(p-2)(x-y)-\{(u-w)+(x-y)\}]}{p-1} = x-y$$

이다. 따라서, $\delta P=(u-w)-(x-y)>0$인 경우에만 A-, D-최적이 된다.

$$\delta P=(u-w)-(x-y)>0$$

$$(u-w)-(x-y) = 3r-\frac{1}{k}(3r+\lambda_1-4\lambda_2)-3r+\frac{1}{k}(3r-3\lambda_1) = 4(\lambda_2-\lambda_1)>0$$

로 $\lambda_1<\lambda_2$ 이다. 따라서 $m=2$이고, $\lambda_1<\lambda_2$를 만족하는 그룹분류가능계획을 매개디자인으로 사용한 $p=5$인 블록 CDC는 A-, D-최적이다. □

3.2 $n=2$인 그룹분류가능계획을 매개디자인으로 이용한 최적 블록 CDC

$n=2$인 그룹분류가능계획을 매개디자인으로 사용한 블록 CDC도 $m=2$인 경우와 마찬가지로 $v=mn=2m$이므로 v가 짝수여야 한다. 따라서, $p=4$ 또는 $p=5$ 즉, $v=6$, 또는 $v=10$인 그룹분류가능계획만을 매개디자인으로 사용할 수 있다. $p=5$인 경우의 $C-$행렬의 형태가 일정하지 못하므로 $p=4$인 경우만 고려한다. 평균효율을 위한 고유치 θ_i를 구하기 위해 $C-$행렬의 구조적 특성을 살펴보자.

$n=2$이고 $v=6$인 그룹분류가능계획을 매개디자인으로 사용하여 $p=4$인 CDC를 블록화 한 경우 블록 CDC의 $C-$행렬은 다음과 같다.

$$C=\begin{pmatrix} x & y & t & t \\ y & x & t & t \\ t & t & u & w \\ t & t & w & u \end{pmatrix}$$

여기서, $x=3r-\frac{1}{k}(3r+6\lambda_2)$, $y=t=r-\frac{1}{k}(r+2\lambda_1+6\lambda_2)$,

$u=3r-\frac{1}{k}(3r+2\lambda_1+4\lambda_2)$, $w=r-\frac{1}{k}(r+8\lambda_2)$ 이다.

$C-$행렬의 고유치는 $\theta_1=2\{r-(r+\lambda_1-2\lambda_2)/k\}$, $\theta_2=2\{r-(r-\lambda_1)/k\}$이고, 각각 1개와 2개의 중근을 갖는다.

$A_1=\begin{pmatrix} x & y \\ y & x \end{pmatrix}$ $A_2=\begin{pmatrix} t & t \\ t & t \end{pmatrix}$ $A_3=\begin{pmatrix} t & t \\ t & t \end{pmatrix}$ $A_4=\begin{pmatrix} u & w \\ w & u \end{pmatrix}$이다. 그러면

$$\begin{aligned} |C| = |A_1||A_4-A_3A_1^{-1}A_2| &= (x-y)(x+y)(u-w)(u+w-\frac{4t^2}{x+y}) \\ &= (x-y)(u-w)\{(u+w)(x+y)-4t^2)\} \end{aligned}$$

이다.

$C-$행렬의 한 행 또는 열의 합이 모두 0이므로 $n=2$, $p=4$인 경우 $C-$행렬은 $x+y+2t=2t+u+w=0$이고 $y=t$이므로, $x+y=u+w$이다, 즉,

$$|C| = (x-y)(u-w)\{(x+y)^2-4y^2)\} = (x-y)^2(u-w)(x+3y)$$

이고, $x+3y=0$이므로 $C-$행렬의 고유치는 1개의 $u-w$와 2개의 $x-y$이다.

$n=2$이고 $p=4$인 경우 역시 $A^2>B>A^2/(p-1)$이다. 이 경우 블록 CDC의 고유치는 $u-w$, $x-y$, $x-y$이므로 $A=(u-w)+2(x-y)$, $B=(u-w)^2+2(x-y)^2$이고,

$$\{\delta P\}^2 = \frac{p-1}{p-2} \times \frac{B(p-1)-A^2}{p-1} = \frac{3B-A^2}{2}$$
$$= [3\{(u-w)^2+2(x-y)^2\} - \{(u-w)+2(x-y)\}^2]/2 = [(u-w)-(x-y)]^2$$

이다.

(정리3.3) $n=2$, $\lambda_1 < \lambda_2$인 그룹분류가능계획을 매개디자인으로 사용한 $p=4$인 CDC는 A-, D-최적이다.

(증명) δP는 양수이고 $\pm\{(u-w)-(x-y)\}$을 갖는다.

$\delta P=(x-y)-(u-w)$이면 (보조정리2.1)을 만족하지 않고, $\delta P=(u-w)-(x-y)$이면 (보조정리2.1)을 만족한다. $\delta P=(u-w)-(x-y)$를 (보조정리2.1)에 적용시키면

$$\frac{A+\delta(p-2)P}{p-1} = \frac{[(u-w)+(p-2)(x-y)+(p-2)\{(u-w)+(x-y)\}]}{p-1} = u-w$$

$$\frac{A-\delta P}{p-1} = \frac{[(u-w)+(p-2)(x-y)-\{(u-w)+(x-y)\}]}{p-1} = x-y$$

이다. 따라서, $\delta P=(u-w)-(x-y)>0$인 경우에만 A-, D-최적이 된다.

$$\delta P=(u-w)-(x-y)>0$$

$$(u-w)-(x-y) = 2r-\frac{1}{k}(2r+2\lambda_1-4\lambda_2)-2r+\frac{1}{k}(2r-2\lambda_1) = 4(\lambda_2-\lambda_1)>0$$

로 $\lambda_1 < \lambda_2$ 이다. 따라서, $n=2$이고, $\lambda_2 > \lambda_1$인 그룹분류가능계획을 매개디자인으로 사용한 $p=4$인 블록 CDC는 A-, D-최적이다. □

(정리3.4) $m=2$ 또는 $n=2$이고 $\lambda_2=\lambda_1+1$를 만족하는 그룹분류가능계획을 매개디자인으로 사용한 블록 CDC는 일반화된 최적이다.

(증명) $n-ary$ designs 중에서 (보조정리2.2)의 (i), (ii)는 $|\lambda_2-\lambda_1|=1$일 때 성립되므로, $\lambda_2=\lambda_1+1$을 만족하는 그룹분류가능계획은 모든 (i) $A^* \geq A$, (ii) $A^*-\delta P^* \geq A-\delta P$이 성립된다. 따라서, $p=4$ 또는 $p=5$인 CDC를 $m=2$ 또는 $n=2$이고, $\lambda_2=\lambda_1+1$인 그룹분류가능계획로 블록화한 경우 일반화된 최적이다. □

참고문헌

Agarwal, S.D and Das, M.N. (1990). Use of $n-ary$ block designs in diallel crosses evalation, *Journal of Applied Statistics*, 17, 125-131.

Cheng, C.S. (1978). Optimality of certain asymmetrical experimental designs. Annals of Statistics, 6, 1239-1261.

Das.A., Dey.A. and Dean.A.M. (1998). Optimal designs for diallel cross experiments. *Statistics & Probability Letters*, 36,427-436.

Dey, A. and Midha, C.K. (1996). Optimal block designs for diallel crosses. *Biometrika*, 83, 484-489.

Divecha, J. and Gosh, D.K. (1994). Incomplete block designs for complete diallel crosses and their analysis, *Journal of Applied Statistics*, 21, 395-408.

Griffing, B. (1956). Concept of general and specific combining ability in relation to diallel crossing systems. *Australian Journal of Biological Sciences* 9, 463-493.