

실험계획의 시각화

장대홍¹

¹부경대학교 통계학과

(2011년 5월 접수, 2011년 8월 채택)

요약

실험계획법의 강의내용은 크게 두 개의 파트인 실험계획과 모형분석으로 대별되는데 시각화 작업은 주로 모형분석 중심으로 이루어져 왔다. 실험계획법의 강의내용에 대한 시각화 작업의 일환으로 우리는 실험계획의 시각화를 라틴방격법의 시각화, 초포화계획법의 시각화, 불완비블록계획법의 시각화로 나누어 고려하여 볼 수 있다. 실험계획을 시각화하는 작업을 위하여 우리는 계획그림을 제안 할 수 있고 기존의 산점도나 산점도행렬을 사용할 수 있다. 이러한 실험계획의 시각화를 통하여 우리는 이론 중심의 실험계획법 강의에 그림들을 삽입함으로써 실험계획법 수업에서의 시너지효과를 얻을 수 있다.

주요용어: 실험계획, 시각화, 계획그림, 산점도행렬.

1. 서론

실험계획법의 강의내용은 크게 실험계획과 모형분석으로 대별된다. 실험계획법의 강의내용에서 시각화 작업은 주로 모형분석 중심으로 이루어져 왔다. 우리는 모형적합성 검토(오차항의 검토)에서 오차항의 독립성, 등분산성, 정규성을 검토하기 위한 그림으로서 잔차도나 정규확률플롯 등을 많이 이용한다. 또한, 우리는 주효과 및 교호작용 그림(평균반응프로파일), 효과 측정을 위한 (반)정규확률플롯, 반응표면그림, 등고선그림, 빗금-블록 행렬그림(shaded-block matrix plot, Barton과 Schruben (1994)) 등을 이용할 수 있다. 반응표면실험계획을 평가하기 위한 그래픽방법으로서는 분산산포그래프(variance dispersion graph, Giovannitti-Jensen과 Myers (1989)), 분위수그림(quantile plot, Khuri 등 (1996)), 분위수산포그래프(quantile dispersion graph, Khuri (1997)), FDS그림(fraction of design space plot, Zahran 등 (2003)), 동적분위수그림(dynamic quantile plot, 장대홍 (2010)) 등이 있다.

실험계획을 시각화하는 작업으로는 요인실험계획이나 부분요인실험계획을 나타내는 계획-그림(design-plot, Barton (1998)), 반응-척도화 계획-그림(response-scaled design-plot) 등이 있다. 실험계획을 시각화하는 또 다른 작업으로서는 라틴방격법을 시각화하는 작업(인터넷 상에서 다수의 작업을 확인할 수 있음)이나 서로 직교하는 라틴방격법(mutually orthogonal Latin squares)의 시각화 작업 (Hyndman, 2009)이 있다. 실험계획법의 강의내용상 수식 전개 중심의 강의에 첨가하여 실험계획의 시각화를 통한 교육을 실시한다면 우리는 실험계획법의 강의에 시너지효과를 얻을 수 있다. 자칫 지루하기 쉬운 수업 내용을 보완할 수 있는 좋은 그래픽도구가 될 수 있다.

2절에서는 실험계획의 시각화에 대하여 세 가지 주제(라틴방격법의 시각화, 초포화계획법의 시각화, 불완비블록계획법의 시각화)로 나누어 언급하였고 3절에서는 결론을 내렸다. 2절에서 언급되는 모든 그림들은 R 패키지를 이용하여 그린 그림들이다.

¹(608-737) 부산광역시 남구 대연3동 599-1, 부경대학교 통계학과, 교수. E-mail: dhjang@pknu.ac.kr

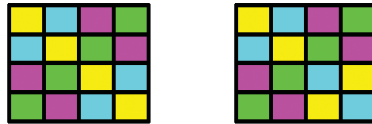
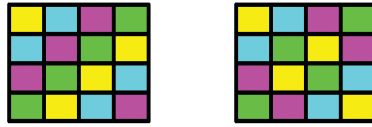


그림 2.1. 4개의 4×4 표준방격을 나타내는 계획그림



그림 2.2. 서로 직교하는 세 개의 7×7 라틴방격법들을 나타내는 계획그림

2. 실험계획의 시각화

2.1. 라틴방격법의 시각화

k 개의 숫자 또는 글자를 어느 행, 어느 열에도 하나씩 나열하여 가로, 세로 k 개의 숫자 또는 글자가 정사각형이 되도록 한 실험계획이 라틴방격법(Latin square design)이다. 우리는 Hyndman (2009)의 아이디어를 사용하여 실험계획법의 강의에 실험계획의 시각화를 통한 교육을 실시할 수 있다. 다음 그림 2.1은 4개의 표준 4×4 라틴방격법을 시각화한 그림들이다. 이러한 그림들을 계획그림(design plot)이라 부르도록 하자. 숫자 또는 글자(처리수준)에 해당하는 색깔은 임의로 선택할 수 있다.

다음 그림 2.2는 서로 직교하는 세 개의 7×7 라틴방격법들을 나타내는 계획그림이다. 이 세 개의 7×7 라틴방격법들을 조합하면 초그레코라틴방격법이 된다. 그림 2.3은 그림 2.2에 나타난 세 개의 7×7 라틴방격법들을 조합하여 만든 초그레코라틴방격법을 나타내는 계획그림이다.

스도쿠(Sudoku)는 라틴방격법을 기반으로 하는 유명한 숫자퍼즐로서, 가로 9칸, 세로 9칸으로 이루어져 있는 9×9 정사각형 표(네모 표)에 1부터 9까지의 숫자를 채워넣는 퍼즐이다. 이 9×9 정사각형 표는 다시 작은 9개의 3×3 칸으로 구성된 정사각형 블록들로 구성되는데 9개의 가로줄에 1부터 9까지의 숫자가 한번씩, 9개의 세로줄에 1부터 9까지의 숫자가 한번씩, 9개의 블록에 각각 1부터 9까지의 숫자가 한번씩만 들어가야 한다. 이 스도쿠는 실험계획법 책에도 언급되고 있다 (Montgomery, 2009). 그림

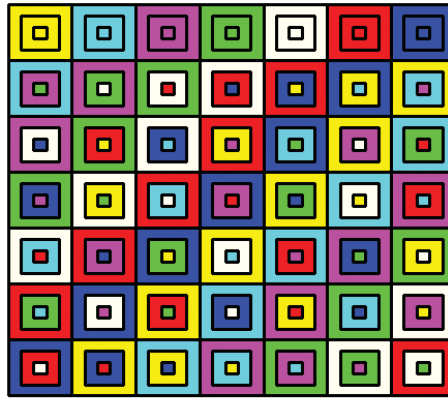


그림 2.3. 7×7 초그레코라틴방격법을 나타내는 계획그림

1	2	3	4	5	6	7	8	9
7	8	9	1	2	3	4	5	6
4	5	6	7	8	9	1	2	3
6	4	5	9	7	8	3	1	2
3	1	2	6	4	5	9	7	8
9	7	8	3	1	2	6	4	5
8	9	7	2	3	1	5	6	4
5	6	4	8	9	7	2	3	1
2	3	1	5	6	4	8	9	7

1	2	3	4	5	6	7	8	9
4	5	6	7	8	9	1	2	3
7	8	9	1	2	3	4	5	6
8	9	7	2	3	1	5	6	4
2	3	1	5	6	4	8	9	7
5	6	4	8	9	7	2	3	1
6	4	5	9	7	8	3	1	2
9	7	8	3	1	2	6	4	5
3	1	2	6	4	5	9	7	8

그림 2.4. 서로 직교하는 두 개의 9×9 스도쿠방격법

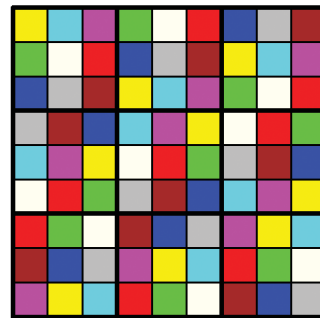
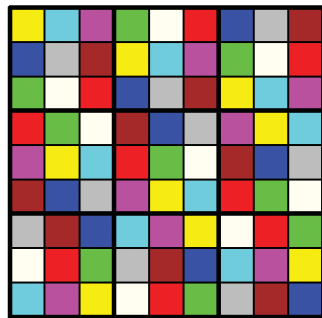


그림 2.5. 서로 직교하는 두 개의 9×9 스도쿠방격법을 나타내는 계획그림

2.4에 나타나는 서로 직교하는 두 개의 스도쿠에 대하여 계획그림을 그리면 그림 2.5와 같고 서로 직교하는 두 개의 스도쿠를 조합한 계획그림은 그림 2.6과 같다. 이 그림을 통하여 두 개의 스도쿠를 조합한 계획그림은 9×9 그레코라틴방격법이 됨을 알 수 있다.

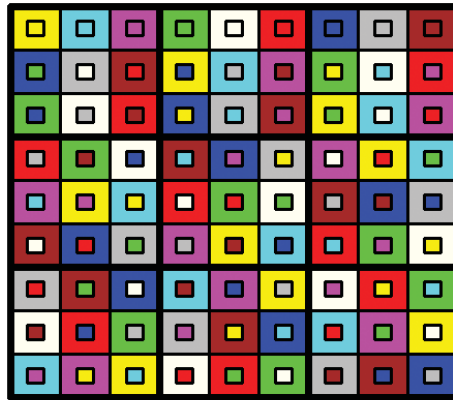


그림 2.6. 서로 직교하는 두 개의 9×9 스도쿠방격법을 조합한 계획그림

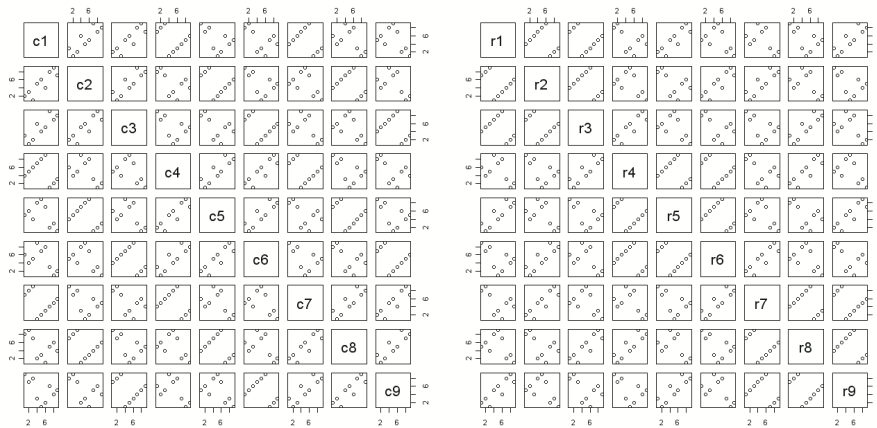


그림 2.7. 첫 번째 9×9 스도쿠방격법에 대한 산점도행렬

각 행, 열, 블록에서 1부터 9까지의 숫자가 한 번씩 나타나는 지를 확인하기 위하여 스도쿠방격법을 숫자로 나타낸 그림 2.4와 스도쿠방격법을 색깔로 나타낸 그림 2.5를 비교하여 보면 대부분의 사람들이 색깔보다는 숫자에 훨씬 익숙하기 때문에 그림 2.4처럼 스도쿠방격법을 숫자로 나타내는 것이 더 효과적일 수 있다. 그러나 전통적으로 실험계획법 수업은 반응표면분석 파트를 제외하면 수식 위주의 교육 중심으로 이루어져 있으므로 학생들이 실험계획법 수업에 대하여 지루해하거나 어렵게 느낄 공산이 크다. 수식 위주의 교육 내용에 이러한 시각화 작업을 첨가한다면 실험계획법 수업이 훨씬 효과적일 수 있다.

우리는 실험계획을 대상으로 산점도행렬을 적용하면 실험계획의 행과 열에 대한 정보를 얻을 수 있다. 그림 2.4에 나타나는 서로 직교하는 두 개의 스도쿠 각각에 대하여 행과 열에 대한 산점도행렬을 그리면 다음 그림 2.7과 그림 2.8과 같다. 그림 2.7에서 왼쪽 산점도행렬은 첫 번째 9×9 스도쿠방격법을 나타내는 행렬을 이용하여 그렸다. 각 산점도는 해당 열들 사이의 산점도를 나타낸다. c1-c9은 스도쿠방격법에서의 9개의 열을 각각 나타낸다. 그림 2.7에서 오른쪽 산점도행렬은 첫 번째 9×9 스도쿠방격법을 나타내는 행렬의 전치행렬을 이용하여 그렸다. 각 산점도는 해당 행들 사이의 산점도를 나타낸다. r1-r9은 스도쿠방격법에서의 9개의 행을 각각 나타낸다. 그림 2.7과 그림 2.8을 통하여 우리는 두 개의

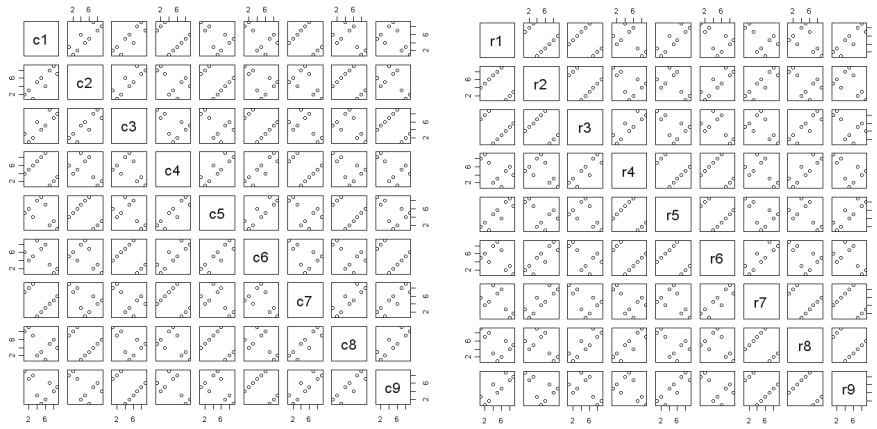


그림 2.8. 두 번째 9×9 스토쿠방격법에 대한 산점도행렬

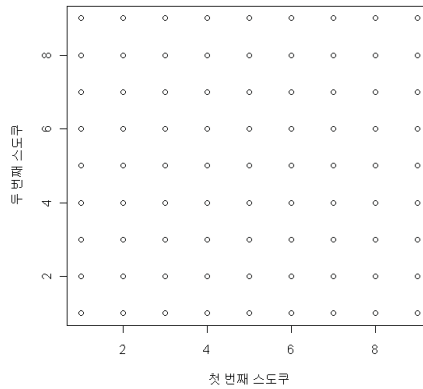


그림 2.9. 서로 직교하는 두 개의 9×9 스토쿠방격법을 나타내는 산점도

스토쿠 모두 9개의 행에 1부터 9까지의 숫자가 한 번씩, 9개의 열에 1부터 9까지의 숫자가 한 번씩 들어감을 알 수 있다. 더불어 두 열 사이나 두 행 사이의 숫자 배열의 구조도 확인할 수 있다. 그림 2.4에 나타나는 서로 직교하는 두 개의 스토쿠를 조합한 후 산점도를 그리면 그림 2.9와 같다. 91개의 모든 조합이 나타나므로 이 두 개의 스토쿠가 직교함을 알 수 있다.

Bailey 등 (2008)은 서로 직교하는 9×9 스토쿠방격법 여섯 개를 보였는데 이러한 여섯 개의 서로 직교하는 9×9 스토쿠방격법을 조합한 계획그림은 다음 그림 2.10과 같다. 여섯 개의 서로 직교하는 9×9 스토쿠방격법에서 각 두 개씩의 스토쿠방격법들을 조합한 후 이들을 이용하여 산점도행렬을 그리면 그림 2.11과 같다. 각 산점도는 해당되는 두 개의 스토쿠를 조합한 후 구한 산점도이다. 모든 산점도가 그림 2.9처럼 나타나 91개의 모든 조합이 나타나므로 여섯 개의 스토쿠가 서로 직교함을 알 수 있다.

2.2. 초포화계획법의 시각화

고려하여야 할 인자의 수는 많으나 실험횟수가 적은 경우 우리는 통상 초포화계획(supersaturated design)을 사용한다. 최근에 실험계획 연구자들 사이에서 이 초포화계획에 대한 관심이 지속적으로 증가하고 있다. 이러한 초포화계획을 계획그림과 산점도행렬로 표시하여 보면 우리는 계획의 구조를 확인할

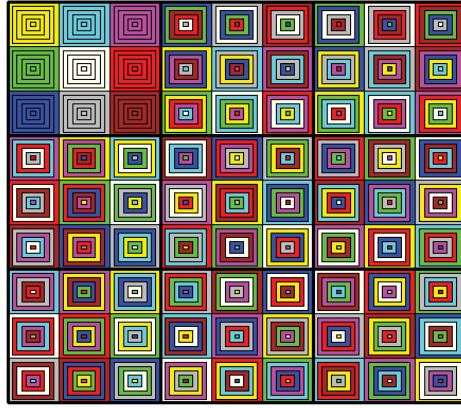


그림 2.10. 서로 직교하는 여섯 개의 9×9 스도쿠방격법을 조합한 계획그림

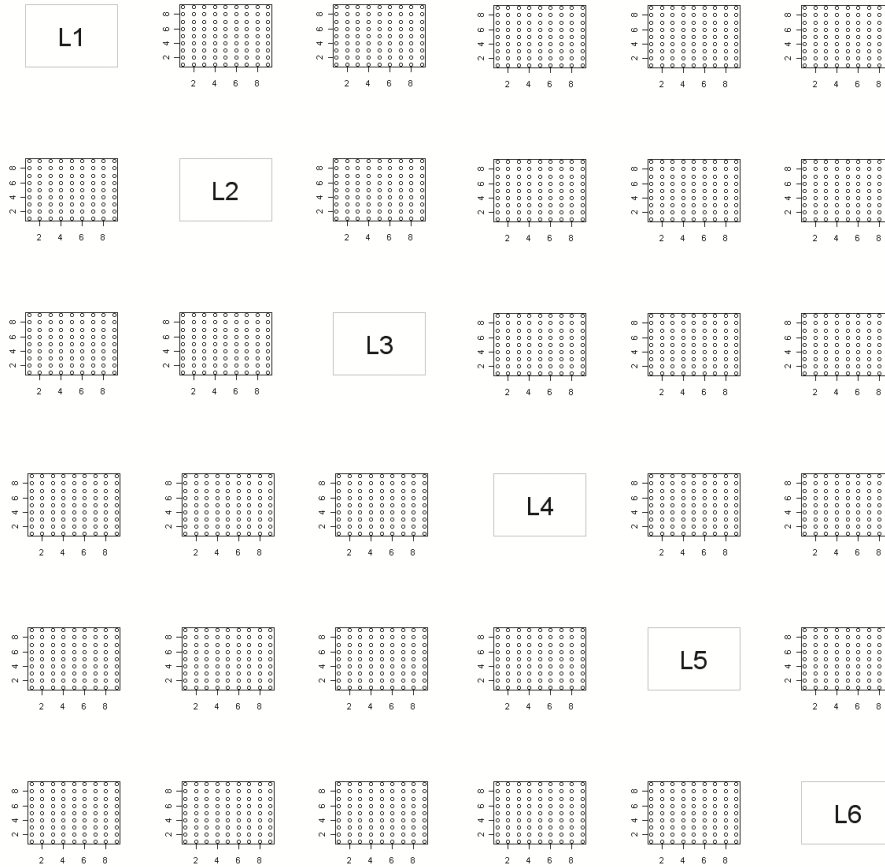


그림 2.11. 서로 직교하는 여섯 개의 9×9 스도쿠방격법을 나타내는 산점도행렬

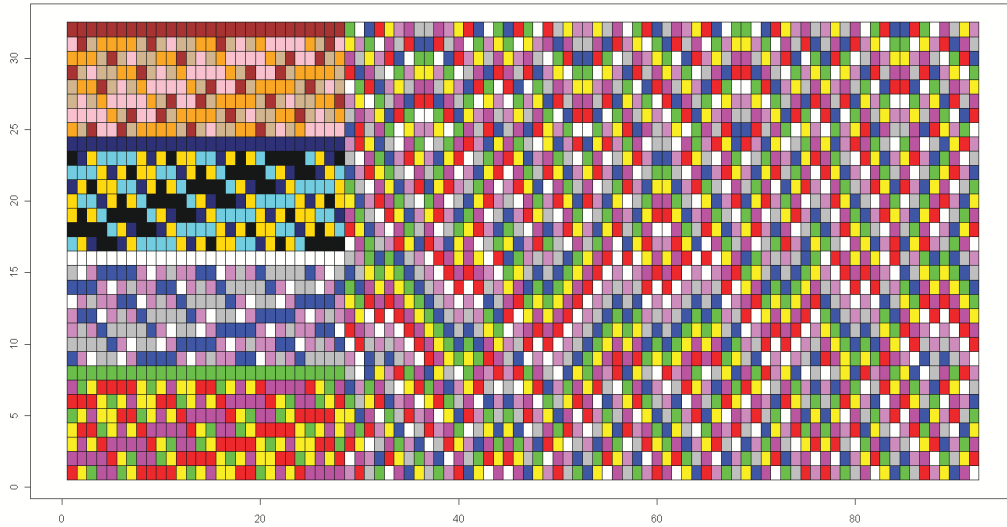


그림 2.12. 초포화계획 $L'_{32}(16^{28} \times 8^{64})$ 에 대한 계획그림

수 있다. 특히 혼합수준계이거나 인자의 수가 과다하게 많은 경우에 우리는 계획그림과 산점도행렬을 이용하여 이러한 초포화계획의 특징을 쉽게 알 수 있게 된다.

$$L'_N (q_1^{m_1} \times q_2^{m_2} \times \dots \times q_s^{m_s}), \quad \sum_{i=1}^s m_i = m$$

를 q_i 개의 수준을 갖는 열들이 m_i 개인 초포화계획이라 하고 직교배열은

$$L_N (q_1^{m_1} \times q_2^{m_2} \times \dots \times q_s^{m_s}), \quad \sum_{i=1}^s m_i = m$$

로 표시하자. 다음 그림 2.12는 Liu와 Liu (2011)가 제시한 초포화계획 $L'_{32}(16^{28} \times 8^{64})$ 에 대한 계획그림이다. 이 계획은 $L'_8(4^{28})$ 과 직교배열 $L_{64}(8^4)$ 을 이용하여 만들어졌다. 열에서는 28번째 열과 29번째 열 사이에서 패턴이 바뀔 수 있다. 행에서는 8개씩 묶어 패턴이 형성됨을 알 수 있다. 그림 2.13은 초포화계획 $L'_{32}(16^{28} \times 8^{64})$ 을 나타내는 행렬을 이용하여 그린 산점도행렬이다. 포화계획의 25번째 열부터 40번째 열까지만 표시하였다. 각 산점도는 해당 열들 사이의 산점도를 나타낸다. c25-c40은 포화계획의 25번째 열부터 40번째 열을 각각 나타낸다. 28번째 열과 29번째 열 사이에서 패턴이 바뀔 수 있다. 첫 번째 그룹(첫 번째 열에서 28번째 열까지)에서는 서로 비직교성의 정도가 아주 강함을 알 수 있고 두 번째 그룹(29번째 열부터 92번째 열까지)에서는 서로 비직교성이 나타나나 다양한 패턴이 나타남을 알 수 있다. 첫 번째 그룹과 두 번째 그룹 사이에는 직교성이 어느 정도 강함을 알 수 있다. 각 산점도에서의 패턴을 통하여 행에서는 8개씩 묶어 패턴이 형성됨을 알 수 있다.

다음 그림 2.14는 Liu와 Liu (2011)가 제시한 초포화계획 $L'_{27}(9^{117})$ 에 대한 계획그림이다. 이 계획은 $L'_9(3^{36})$ 과 직교배열 $L_{81}(9^3)$ 을 이용하여 만들어졌다. 열에서는 36번째 열과 37번째 열 사이에서 패턴이 바뀔 수 있다. 행에서는 9개씩 묶어 패턴이 형성됨을 알 수 있다. 그림 2.15는 초포화계획 $L'_{27}(9^{117})$ 을 나타내는 행렬을 이용하여 그린 산점도행렬이다. 포화계획의 30번째 열부터 45번째 열까

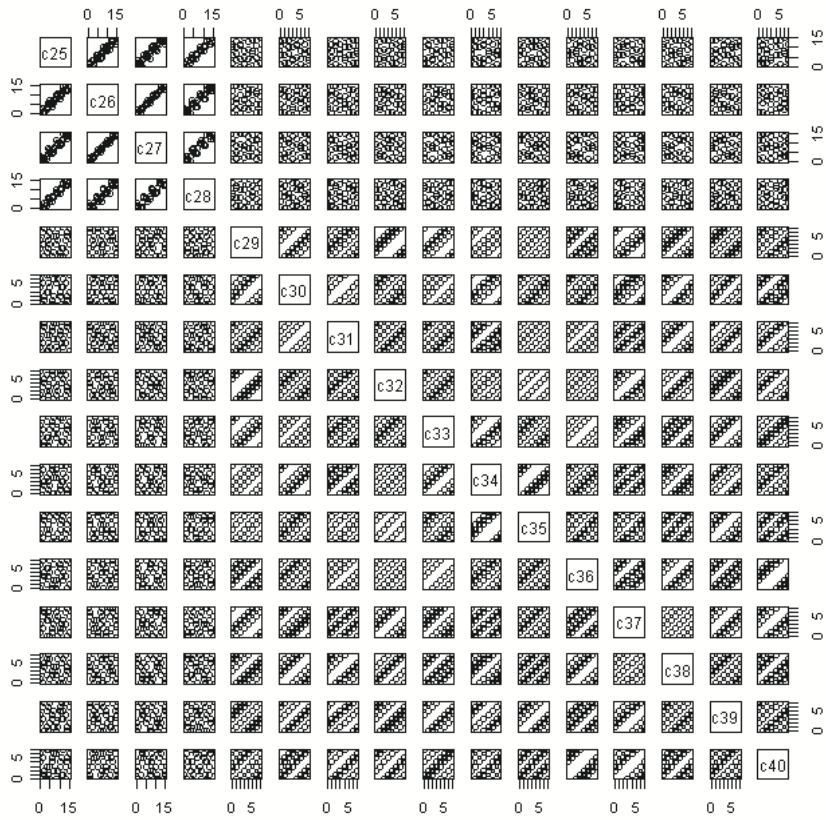


그림 2.13. 초포화계획 $L'_{32}(16^{28} \times 8^{64})$ 에 대한 산점도행렬(25번째 열부터 40번째 열까지만 표시함)

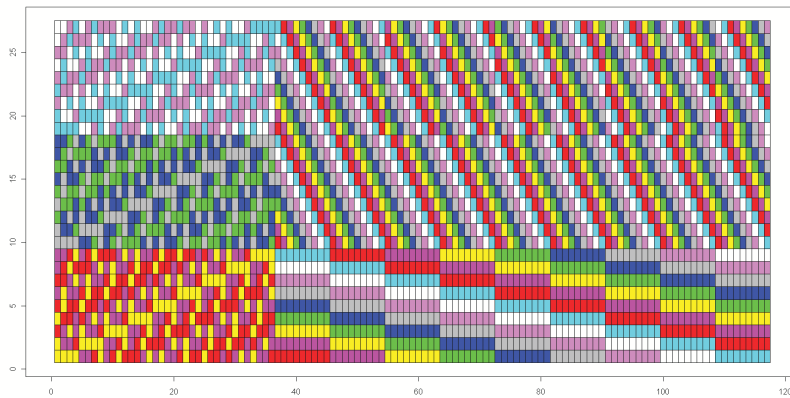


그림 2.14. 초포화계획 $L'_{27}(9^{117})$ 에 대한 계획그림

지만 표시하였다. 각 산점도는 해당 열들 사이의 산점도를 나타낸다. c30-c45는 포화계획의 30번째 열부터 45번째 열을 각각 나타낸다. 36번째 열과 37번째 열 사이에서 패턴이 바뀔 수 있다. 첫 번째

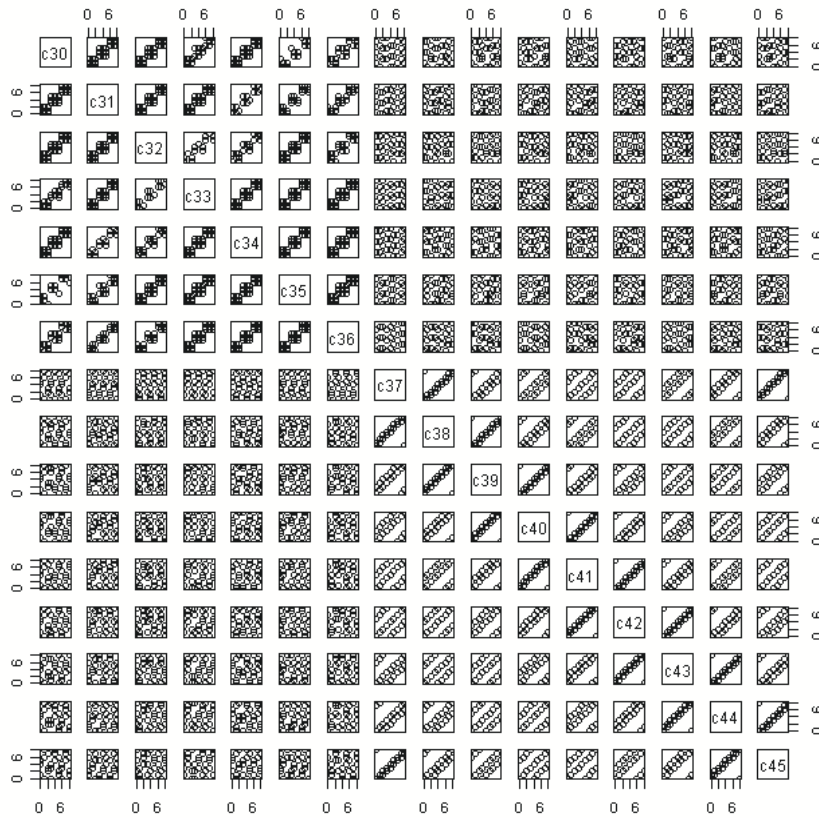


그림 2.15. 초포화계획 $L_{27}^I(9^{117})$ 에 대한 산점도행렬(30번째 열부터 45번째 열까지만 표시함)

그룹(첫 번째 열에서 36번째 열까지)에서는 서로 비직교성의 정도가 아주 강함을 알 수 있고 두 번째 그룹(37번째 열부터 117번째 열까지)에서는 서로 비직교성이 나타나나 다양한 패턴이 나타남을 알 수 있다. 첫 번째 그룹과 두 번째 그룹 사이에는 직교성이 어느 정도 강함을 알 수 있다. 각 산점도에서의 패턴을 통하여 행에서는 9개씩 묶어 패턴이 형성됨을 알 수 있다.

김영일과 장대홍 (2010)은 초포화계획을 평가하기 위한 그래픽방법에 대하여 언급하였다. 실험계획법 강의 시 이러한 그래픽방법과 더불어 본 논문에서 제안하는 계획그림과 산점도행렬을 병행해서 사용하면 우리는 시너지효과를 기대할 수 있다.

2.3. 불완비블록계획법의 시각화

처리의 개수가 t 이고 블록의 개수가 b 인 블록계획법에서 한 블록에 모든 처리를 포함하지 못 하면 불완비블록계획이 된다. 이러한 불완비블록계획이 균형불완비블록계획이 되려면 다음과 같은 세 가지 조건들이 만족이 되어야 한다.

1. 모든 블록에서 k 개의 처리가 이루어진다. 즉, 모든 블록의 크기는 k 개로 같다.
2. 각 처리는 r 개의 블록에 나타난다. 즉, 모든 처리의 반복수는 r 로 같다.

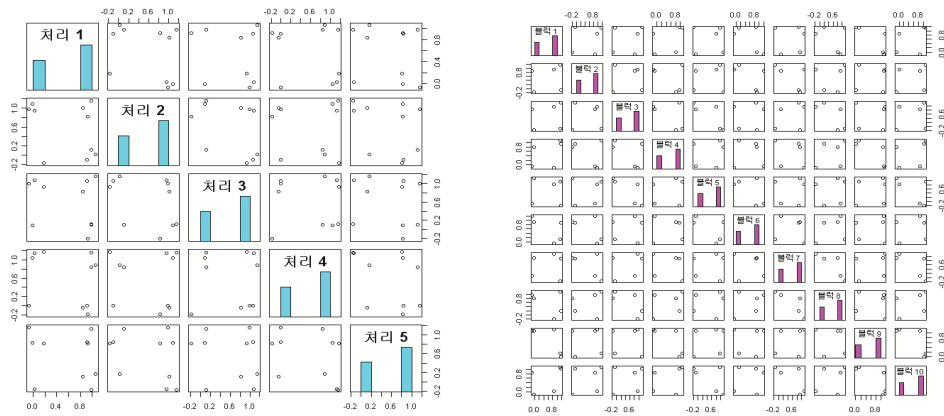


그림 2.16. 불완비블록계획법에 대한 산점도행렬

3. 임의의 두 처리가 동시에 이루어지는 블록의 수는 λ 로 동일하다. 즉, 임의의 두 처리가 같은 블록에서 만나는 횟수는 λ 로 동일하다.

균형불완비블록계획법에서는 다음 관계식이 성립한다.

$$\lambda = \frac{r(k-1)}{t-1}. \tag{2.1}$$

발생행렬(incidence matrix, 결합행렬이라고도 함) N 은 $t \times b$ 행렬으로서 각 행은 대응되는 처리를 나타내고 각 열은 대응되는 블록을 나타낸다. 원소 n_{ij} 는 i 처리가 j 블록에 있으면 1이고, 없으면 0이 된다. 이 발생행렬은 불완비블록계획법을 나타내는 좋은 도구이다. 이러한 발생행렬의 전치행렬을 대상으로 모든 처리들에 대응하는 산점도행렬을 그려보고 발생행렬을 대상으로 모든 블록들에 대응하는 산점도행렬을 그려보면 우리는 불완비블록계획이 균형불완비블록계획이 되는 지를 밝힐 수 있다.

처리의 개수가 5, 블록의 개수가 10인 불완비블록계획법에 대한 발생행렬이

$$N = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

일 때 이러한 발생행렬의 전치행렬을 대상으로 모든 처리들에 대응하는 산점도행렬을 그려보고 발생행렬을 대상으로 모든 블록들에 대응하는 산점도행렬을 그려보면(홀트림(jittering) 사용) 다음 그림 2.16과 같다. 균형불완비블록계획법이 되기 위한 세 가지 조건들이 모두 만족됨을 알 수 있다. 즉 그림 2.16의 왼쪽에 있는 산점도행렬(발생행렬의 전치행렬을 이용한 산점도행렬)에서 각 처리에 대응하는 대각선 패널에 있는 오른쪽 막대그래프들을 비교하여 보면 모든 처리의 반복수가 같음을 알 수 있고, 그림 2.16의 오른쪽에 있는 산점도행렬(발생행렬을 이용한 산점도행렬)에서 각 블록에 대응하는 대각선 패널에 있는 오른쪽 막대그래프들을 비교하여 보면 모든 블록의 크기가 같음을 알 수 있다. 또한 그림 2.16의 왼쪽에 있는 산점도행렬에서 각각 두 개의 처리들 쌍에 대응하는 비대각선 패널의 오른쪽 상단에 표시된 점들의 개수가 $\lambda = 3$ 개로 모두 동일함을 알 수 있다. 그러므로 이 불완비블록계획은 균형불완비블록계획법이 되기 위한 세 가지 조건들을 모두 만족하므로 균형불완비블록계획법이 된다.

3. 결론

실험계획법의 강의내용은 크게 실험계획과 모형분석으로 대별된다. 실험계획법의 강의내용에서 시각화 작업은 주로 모형분석 중심으로 이루어진다. 그러나 실험계획법의 강의내용에서 실험계획을 시각화하는 작업(라틴방격법의 시각화, 초포화계획법의 시각화, 불완비블록계획법의 시각화)도 의미있는 작업이 된다. 실험계획을 시각화하는 작업을 위하여 우리는 계획그림을 제안 할 수 있고 기존의 산점도나 산점도행렬을 사용할 수 있다. 실험계획법의 강의내용상 수식 전개 중심의 강의에 첨가하여 실험계획의 시각화를 통한 교육을 실시한다면 우리는 실험계획법 강의 시 시너지효과를 얻을 수 있다.

참고문헌

- 김영일, 장대홍 (2010). 초포화계획을 평가하기 위한 그래픽방법, <응용통계연구>, **23**, 167-178.
- 장대홍 (2010). 반응표면실험계획을 평가하기 위한 등적분위수그림, <응용통계연구>, **23**, 285-293.
- Bailey, R. A., Cameron, P. J. and Connelly, R. (2008). Sudoku, gerechte designs, resolutions, affine space, spreads, reguli, and Hamming codes, *American Mathematical Monthly*, **115**, 383-404.
- Barton, R. R. (1998). Design-plots for factorial and fractional factorial designs, *Journal of Quality Technology*, **30**, 40-54.
- Barton, R. R. and Schruben, L. W. (1994). Graphical methods for comparing confounding in two or more designs, *Communications in Statistics-Simulation and Computation*, **23**, 953-971.
- Giovannitti-Jensen, A. and Myers, R. H. (1989). Graphical assessment of the prediction capability of response surface designs, *Technometrics*, **31**, 159-171.
- Hyndman, T. (2009). Math art: Latin square, http://tylo42.com/latin_square.
- Khuri, A. I. (1997). Quantile dispersion graphs for analysis of variance estimates of variance components, *Journal of Applied Statistics*, **24**, 711-722.
- Khuri, A. I., Kim, H. J. and Um, Y. (1996). Quantile plots of the prediction variance for response surface designs, *Computational Statistics and Data Analysis*, **22**, 395-407.
- Liu, Y. and Liu, M. (2011). Construction of optimal supersaturated design with large number of levels, *Journal of Statistical Planning and Inference*, **141**, 2035-2043.
- Montgomery, D. C. (2009). *Design and Analysis of Experiments*, 7th ed., Wiley, New York.
- Zahran, A., Anderson-Cook, C. M. and Myers, R. H. (2003). Fraction of design space to assess the prediction capability of response surface designs, *Journal of Quality Technology*, **35**, 377-386.

Visualization for Experimental Designs

Dae-Heung Jang¹

¹Department of Statistics, Pukyong National University

(Received May 2011; accepted August 2011)

Abstract

The lecture of the experimental designs consists of two main part-experimental designs and model analysis. Mostly, the progress of the visualization has been made on a model analysis. As the visualization of experimental designs, we can consider the visualization of Latin squares, supersaturated designs, and balanced incomplete block designs. We can propose the design plots as well as use the scatterplots and the scatterplot matrices for the visualization of experimental designs. Through the visualization of experimental designs, we can use the synergy effect in teaching the lecture of the experimental designs.

Keywords: Experimental designs, visualization, design plot, scatterplot matrix.

¹Professor, Department of Statistics, Pukyong National University, 599-1 Daeyeon-dong, Nam-gu, Busan 608-737, Korea. E-mail: dhjang@pknu.ac.kr