

공간채움 조건을 만족하는 컴퓨터 실험 시나리오의 효율적 생성

임동순^{1†} · 김정훈² · 최봉완³

Efficient Generation of Space Filling Scenarios for Computer Experiments

Dong-Soon Yim · Jung-Hoon Kim · Bong-Whan Choi

ABSTRACT

In general, simulation models are effectively used in the field of engineering design. The experiment with simulation models to obtain optimal design parameters, however, is a time-consuming task and requires a lot of resources. Hence, meta-models representing the relationships between input variables and performance measures are exploited to efficiently determine the value of design parameters. To construct a meta-model, a number of simulation executions with sample scenarios are required. The number and quality of sample scenarios determine not only the level of efficiency in constructing the meta-model but also accuracy of the model. Space-filling condition is regarded to be an important condition for the quality of scenarios. This paper proposes sample scenario generation methods based on space-filling measures such as maxmin, Audze-Eglaiss, and centered L2-discrepancy. The performance of these scenario generation methods are evaluated through experiments.

Key words : Computer Experiments, Meta-models, Space-filling

요약

공학설계분야에서 시뮬레이션 모델을 이용한 실험은 매우 중요한 역할을 한다. 그러나 최적의 설계 파라미터를 구하기 위한 시뮬레이션 실험은 많은 실행시간과 자원의 소요를 필요로 한다. 이를 극복하는 방안으로 입력 변수와 성능척도 간의 관계를 표현한 메타모형이 효과적으로 이용된다. 메타모형을 수립하기 위해서는 샘플 시나리오들을 입력으로 하는 시뮬레이션 실험이 요구된다. 이때 샘플 시나리오의 수와 질이 메타 모형을 수립하는데 걸리는 시간과 메타모형의 정확성을 결정한다. 공간채움 특성은 샘플 시나리오들의 질을 결정하는 중요한 조건이 된다. 이 논문은 maxmin, Audze-Eglaiss, centered L2-discrepancy의 3가지 공간채움 척도에 기초한 샘플 시나리오 생성 방법을 제안하고, 실험을 통해 이들 생성방법에 대한 성능을 분석한 결과를 논의한다.

주요어 : 컴퓨터실험, 메타모델, 공간채움

1. 서론

제품과 공정, 그리고 생산 시스템을 대상으로 하는 공

학설계분야에서 컴퓨터에 의한 실험은 매우 중요한 역할을 한다. 실제적인 환경을 만들어 수행하는 물리적 실험은 많은 비용을 필요로 하므로 이를 대체하는 방안으로 시뮬레이션 모델에 의한 실험이 이용된다. 공학설계를 위한 대부분의 시뮬레이션 실험은 품질, 비용, 납기 등의 관점에서 최적의 설계 파라미터를 결정하기 위한 목적을 가진다. 그러나 설계 파라미터를 포함하는 입력변수의 수와 변수 값에 대한 대안 수가 커지면 많은 수의 시뮬레이션 실험이 필요하게 된다. 더욱이 한 번의 시뮬레이션 실험에 적지 않은 시간이 걸린다면 모든 대안들을 대상으로 하는 실험에는 많은 비용과 시간이 요구된다. 이를 극복하

* 이 연구는 국방과학연구소(6본부,UDD110038DD)지원에 의해 수행된 연구결과임.

접수일(2012년 10월 23일), 심사일(2013년 4월 18일),

게재 확정일(2013년 5월 6일)

¹⁾ 한남대학교 산업경영공학과

²⁾ 국방과학연구소 제6기술연구본부

³⁾ 한남대학교 국방M&S연구센터

주 저 자 : 임동순

교신저자 : 임동순

E-mail; dsyim@hnu.kr

는 방안으로 메타모형을 이용한 실험이 이용된다(Kleijnen, 2005). 즉, 입력 변수의 값에 따라 목적함수가 어떻게 반응하는지에 대한 모형을 수립하여 이모형으로부터 최적의 설계 파라미터를 결정한다. 메타모형을 수립하고 추정하기 위해서는 모든 입력변수의 대안에 대한 실험이 아닌 적절한 수의 샘플 시나리오를 통해 효율적인 실험을 하도록 한다. 이 때 어떠한 샘플 시나리오를 대상으로 실험을 하여야 하는지가 중요하다. 보편적으로 사용되는 실험계획 방법으로는 회귀모형을 추정하기 위한 요인배치법과 Kriging 모형을 추정하기 위한 LHD (Latin hypercube design), CSD (Customized Sequential Disign)가 있다 (Kleijnen, 2005; Wim, 2005).

컴퓨터 실험의 설계에 대한 많은 연구들은 샘플 시나리오가 가져야 하는 두 가지 조건을 언급한다(Johnson et al., 1990; Morris and Mitchell, 1995). 첫 번째는 공간 채움 조건(space-filling)으로 샘플 시나리오들이 전체 시나리오 공간 내에 고르게 분포되도록 한다. 전체 시나리오 공간에 대한 정보를 얻기 위한 샘플 시나리오는 일부 공간에 치우치지 않고, 전체 공간을 대표하여야 하기 때문이다(Pronzato and Muller, 2012). 두 번째 조건은 비축약(non-collapsing)으로 샘플 시나리오들이 동일한 입력변수 값을 공유하지 않도록 한다. 이는 시나리오들의 유사성으로 인하여 비슷한 시뮬레이션 실행결과를 가져올지 모르는 경우를 배제하여 시나리오들 간에 차별화를 하기 위함이다.

LHD는 두 번째 조건인 비 축약을 만족하기 위하여 고안되었다. 첫 번째 조건인 공간 채움을 만족하기 위해서는 시나리오 집합이 어느 정도의 공간 채움을 이루고 있는지에 대한 평가 척도를 정의하여 최적의 샘플 시나리오를 구할 수 있다. 최적의 공간 채움을 이루는 LHD에 대한 연구들은 Columwise-pairwise algorithm (Park, 1994; Ye et al., 2000), 시뮬레이티드 애닐링(Morris and Mitchell, 1995), 유전자 알고리즘(Bates et al., 2003; Liefvendahl and Stocki, 2006)등의 알고리즘을 이용하여 근사 최적해를 구하였다.

LHD는 입력 변수에 대한 값을 구하려고 하는 샘플 시나리오 수만큼으로 나눌 수 있다는 가정을 내포하고 있다. 입력변수의 값이 연속적인 실수값을 가진다면 그 가정은 문제가 되지 않지만, 한정된 범위내의 이산적인 값을 가진다면 비 축약 조건을 만족할 수 없는 경우가 발생한다. 또한, 비 축약 조건은 샘플 시나리오의 수를 제한하여 메타모형의 정확한 추정을 어렵게 할 수 있다. 이러한 관점에서 LHD의 비 축약 조건을 만족하지 않더라도, 공간 채

움의 목적을 달성할 수 있는 샘플 시나리오들을 구하는 것이 의미가 있을 수 있다. 구 패킹 설계(sphere packing design), 균등설계(uniform design), 최대 엔트로피 설계(maximum entropy design) 등의 실험계획 방법들이 이에 속한다(Johnson et al., 2008).

본 논문은 비 축약 조건은 고려하지 않고, 공간 채움 조건을 만족시킬 수 있는 시나리오를 생성하는 방법을 대상으로 한다. 특히, 공간 채움의 척도를 이용한 효과적인 최적화 문제를 정의하고, 이 문제의 만족할 만한 해를 구할 수 있는 휴리스틱 방법을 소개한다. 마지막으로 실험을 통해 공간 채움에 대한 다양한 평가척도의 성능을 분석한 결과를 논의한다.

2. 문제 정의

n 개의 입력 변수가 있고, 입력 변수 v_k 는 L_k 개의 수준을 가지고 있을 때 m 개의 시나리오를 생성한다고 하자. m 개의 시나리오를 포함하는 집합 $S = \{s_1, s_2, \dots, s_m\}$ 에서 시나리오 $s_i = \{x_{i1}, x_{i2}, \dots, x_{im}\}$ 는 각 입력변수의 수준 번호를 포함하는 벡터로 표현된다. x_{ik} 는 S 의 i 번째 시나리오에서 입력변수 v_k 의 수준번호로 1부터 L_k 까지의 정수이다. 공간 채움의 목적을 달성하는 시나리오 집합 S^* 을 구하는 것이 본 연구에서의 근본적인 문제이다.

문제를 구체화하기 위해서는 공간 채움을 위한 목적함수를 정의하여야 한다. 이를 위해 우선 두 시나리오 s_i 와 s_j 간의 거리를 유클리디안 거리인 $d_{ij} = |s_i - s_j|$ 로 표현한다. 시나리오들의 공간 채움을 위한 목적 함수는 공간 채움의 정의에 따라 다르다. 대표적인 목적함수로는 maxmin (Johnson et al., 1990)과 Audze-Eglais (Bates et al., 2003)를 들 수 있다. Maxmin은 S 에 포함된 두 시나리오 간 거리들 중 최소 거리인 다음 식의 평가값을 최대화하는 함수로 정의된다.

$$z(S) = \min\{d_{ij}, s_i, s_j \in S\}$$

Audze-Eglais는 전하를 갖는 입자 간에 작용하는 물리적인 힘과 유사한 기준으로 다음 식으로 표현되는 시나리오 집합 S 의 평가값을 최소화한다.

$$z(S) = \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{1}{d_{ij}^2}, s_i, s_j \in S$$

공간 채움을 위한 또 하나의 척도로는 L_p -discrepancy 가 있다. 특히, centered L_2 -discrepancy (CLP)는 계산 상의 이점으로 인해 균등 설계(uniform design) (Fang et al., 2000)를 위한 척도로 많이 이용된다. Hickernell (1998)가 개발한 CLP 척도의 계산식으로 최소화 되어야 할 식은 다음과 같다.

$$z(S) = -\frac{2}{m} \sum_{i=1}^m \prod_{k=1}^n \left(1 + \frac{1}{2} \left| x'_{ik} - \frac{1}{2} \right| - \frac{1}{2} \left| x'_{ik} - \frac{1}{2} \right|^2 \right) + \frac{1}{m^2} \sum_{i=1}^m \sum_{j=1}^m \prod_{k=1}^n \left(1 + \frac{1}{2} \left| x'_{ik} - \frac{1}{2} \right| + \frac{1}{2} \left| x'_{jk} - \frac{1}{2} \right| - \frac{1}{2} \left| x'_{ik} - x'_{jk} \right| \right)$$

x'_{ik} 는 0 부터 1사이의 정규화된 수로 $x'_{ik} = \frac{x_{ik}-1}{L_k-1}$ 이다.

전체 시나리오 수는 $\prod_{k=1}^n L_k$ 이고, 이 중 m 개의 시나리오를 뽑는 경우의 수가 위 문제의 해 공간 크기에 해당하여 입력변수의 수와 수준의 수가 증가하면 해공간이 커져 최적해를 구하기 위해 많은 시간이 요구된다.

시간 복잡성을 완화할 수 있는 방법 중의 하나는 전체 시나리오 보다는 M 개의 한정된 시나리오를 대상으로 하여 이 중 최적의 m 개 시나리오를 선택하는 것이다. 이 때 M 개 시나리오 역시 입력 변수로 구성된 전체 공간을 공평하게 채우고 있는 시나리오이어야 한다. 이를 위한 1차적인 조건은 M 개 시나리오에 포함된 각 입력 변수의 각 수준번호 빈도수가 균등분포에 따르도록 하는 것이다 (Fang and Linn, 2003).

본 연구에서는 이렇게 균등분포에 따라 랜덤으로 생성된 M 개의 시나리오 중 공간 채움의 목적을 위한 m 개 시나리오를 선정하는 문제를 다룬다. 공간 채움의 척도로는 maxmin, Audze-Eglais, 그리고, CLP를 이용한다.

3. 시나리오 생성방법

소개된 문제의 해를 효율적으로 구하기 위한 방법으로 휴리스틱 알고리즘을 고안하였다. Audze-Eglais 또는 CLP 평가함수를 적용한 최소화 문제에 적용할 때의 알고리즘은 다음과 같다.

3.1 최소화 적용 알고리즘

- 1) M 개의 시나리오 중 임의의 m 개를 선택하여 현재 해 S 로 하고, 평가값 $z(S)$ 를 구한다.
- 2) 다음 절차를 반복한다.
 - 2-1) 각 $s_i \in S$, 각 $s_k \notin S$ 들에 대하여 S 에서 s_i 를 제

거하고, s_k 를 삽입하여 시나리오 집합 $S(i,k)$ 을 생성하고, $z(S(i,k))$ 를 구한다.

2-2) $z(S') = \min_{i,k} \{z(S(i,k))\}$ 를 구한다.

2-2-1) $z(S')$ 가 $z(S)$ 보다 작으면 $S=S'$ 로 한다.

2-2-2) $z(S')$ 가 $z(S)$ 보다 크거나 같으면 정지한다.

알고리즘은 TSP (Traveling salesman problem) 문제에 적용되는 1-relocate 향상적 휴리스틱과 유사하다. 그러나 하나의 도시를 현 위치에서 다른 위치로 바꾸는 1-relocate와 다르게 초기해 S 를 구성한 후 S 에서 하나의 시나리오를 제거하고, 새로운 시나리오를 삽입하는 과정을 반복한다. 단계 2-1)에서 한 시나리오의 제거와 삽입에 대한 모든 대안을 고려한다. 2-2)에서는 모든 대안 중 가장 좋은 해를 선택한다. 가장 좋은 해의 평가값이 현재해의 평가값보다 좋다면 이 해를 현재해로 변경한다. 이러한 반복은 평가값이 향상되는 한 계속되어 최종해는 부분 최적해가 된다. 때문에 위 알고리즘은 부분 최적화 알고리즘(local optimization algorithm)에 속한다. 알고리즘의 단계 2)에서 한 번의 반복에 $m \times (n-m)$ 개의 $S(i,k)$ 를 생성한다. 그러나 반복 횟수를 예측할 수 없어 알고리즘의 이론적인 복잡성을 유도하기는 불가능하다.

위 알고리즘에 maxmin 평가값을 적용하고자 한다면 단계 B)를 최대화 문제에 맞게 수정하여야 한다. 그러나 하나의 시나리오를 삭제하고, 새로운 시나리오를 삽입하는 알고리즘은 maxmin 평가값의 경우에 적합하지 않을 수 있다. 현재해에서 최소 거리를 이루는 두 시나리오들이 하나가 아닌 다수가 존재한다면 하나의 시나리오만을 삭제하고 삽입하는 알고리즘은 향상된 해를 생성할 가능성이 많지 않기 때문이다. 최소 거리를 이루는 모든 시나리오들을 삭제하고, 새로운 다수의 시나리오들을 삽입할 때 향상된 해를 생성할 가능성이 더 많아진다. 이와 같은 문제점을 해결하기 위해 다음과 같이 maxmin 평가값에 맞게 알고리즘을 수정하였다.

3.2 Maxmin 적용 알고리즘

- 1) M 개의 시나리오 중 임의의 m 개를 선택하여 현재 해 S 로 하고, 평가값 $z(S)$ 를 구한다. $z(S)$ 와 동일한 값의 거리를 갖는 S 의 두 시나리오들의 개수 $p(S)$ 를 구한다.
- 2) 다음 절차를 반복한다.
 - 2-1) 각 $s_i \in S$, 각 $s_k \notin S$ 들에 대하여 S 에서 s_i 를 제거하고, s_k 를 삽입하여 시나리오 집합 $S(i,k)$

을 생성하고, $z(S(i,k))$ 를 구한다. $z(S(i,k))$ 와 동일한 값의 거리를 갖는 $S(i,k)$ 의 두 시나리오 개수 $p(S(i,k))$ 를 구한다.

2-2) $z(S') = \min_{i,k}\{z(S(i,k))\}$ 를 구한다.

2-2-1) $z(S')$ 가 $z(S)$ 보다 작으면 $S = S'$ 로 한다.

2-2-2) $z(S')$ 가 $z(S)$ 와 같고, $p(S')$ 이 $p(S)$ 보다 작으면 $S = S'$ 로 한다.

2-2-2) 위 두 조건에 속하지 않으면 정지한다.

이전 알고리즘과 마찬가지로 단계 2-A)에서 한 시나리오의 제거와 삽입에 대한 모든 대안을 고려한다. 가장 좋은 해의 평가값이 현재해의 평가값보다 좋다면 이 해를 현재해로 변경한다. 이에 부가하여 가장 좋은 해의 평가값이 현재해의 평가값과 같은 경우에도 현재해를 변경할 수 있는 가능성을 갖도록 하였다. 즉, 해의 평가값인 최소

거리를 유발한 시나리오 쌍의 수가 현재해에서 보다 작다면 역시 이 해를 현재해로 변경한다. 이러한 반복은 평가값의 향상 또는 평가값이 동일한 경우 가장 작은 거리의 시나리오 쌍 개수의 감소가 발생하는 한 계속되어 최종해는 부분 최적해가 된다.

4. 실험 및 분석

4.1 실험 대상 시나리오

설명된 3가지 평가 함수인 maxmin, Audze-Eglais, 그리고 CLP에 기초한 방법을 대상으로 실험을 수행하였다. 알고리즘들은 자바 프로그램 언어로 구현되어 인텔 2.4 GHZ 의 CPU를 가진 PC에서 실행되었다. 생성되어야 할 시나리오는 7개 입력 변수로 구성되어 Table 1과 같은 수준수를 갖고 있는 것으로 가정하였다.

4.2 (1000,100) 실험 결과

7개 입력 변수의 수준 수를 고려하면 모든 시나리오 수는 51840이다. 이 중 1000개 시나리오를 랜덤으로 생성하였고, 생성된 시나리오에 포함된 입력변수의 수준값이 균등분포에 따르도록 하였다. Table 2는 생성된 1000개의 시나리오 중 100개를 선택하여 시나리오 집합으로 하는 경우 각 평가함수에 의한 알고리즘의 실행 결과를 나타낸다. 초기 난수를 변경하여 10회를 실행하였고, 각 실행에서 3가지 평가함수에 의한 알고리즘들은 동일한 1000개 시나리오를 입력으로 하였다. Table 2에서 Z1, Z2, Z3는 각 평가함수에 기초한 방법으로 생성된 해에 대

Table 1. No. of levels for each input variable

variables	no. of levels
V1	6
V2	5
V3	3
V4	6
V5	6
V6	4
V7	4

Table 2. Experiment results in case of (1000, 100)

No	Maxmin				Audze-Eglais				CLP			
	Z1	Z2	Z3	Rep	Z1	Z2	Z3	Rep	Z1	Z2	Z3	Rep
1	8	8.9	-1.5	143	5	6.7	-1.4	102	2	20.2	-1.7	72
2	8	9.1	-1.5	126	5	6.7	-1.4	97	2	20.8	-1.7	70
3	8	8.5	-1.4	133	5	6.5	-1.4	96	2	19.5	-1.7	70
4	8	8.8	-1.5	152	5	6.5	-1.4	97	1	21.2	-1.7	60
5	8	8.8	-1.5	120	6	6.6	-1.4	104	2	19.9	-1.7	70
6	8	8.9	-1.5	132	5	6.6	-1.3	101	3	20.0	-1.7	53
7	8	8.8	-1.5	141	5	6.6	-1.4	98	2	19.9	-1.7	62
8	8	8.6	-1.5	150	5	6.4	-1.3	98	2	20.2	-1.7	59
9	8	8.8	-1.5	136	6	6.6	-1.4	96	1	20.9	-1.7	57
10	8	8.6	-1.4	125	5	6.3	-1.3	94	1	21.1	-1.7	55
Ave	8.0	8.8	-1.5	135.8	51.2	6.6	-1.4	98.3	1.8	20.4	-1.7	62.8

한 최소거리, Audze-Eglais 평가값, CLP 평가값을 나타낸다. Rep는 알고리즘 실행에서 반복된 단계 2)의 횟수를 나타낸다.

Maxmin 평가함수에 의한 방법은 10회의 실행 모두 최소 거리 8을 달성하여 세 방법 중 가장 큰 최소거리를 달성하는 자연스러운 결과를 나타낸다. 마찬가지로 Audze-Eglais 평가함수에 의한 방법은 가장 작은 Z2, CLP 평가함수에 기초한 방법은 가장 작은 Z3를 나타낸다. Maxmin, Audze-Eglais 평가 기준으로 보면 CLP 평가함수에 기초한 방법이 Z1 평균 1.8, Z2 평균 20.4로 가장 나쁜 결과를 나타내고, CLP 평가 기준에서는 Audze-Eglais 평가함수에 기초한 방법이 Z3 평균 -1.4로 가장 나쁜 결과를 나타낸다. 알고리즘의 반복수에서는 CLP 평가기준의 방법이 평균 62.8의 가장 적은 횟수를 보여 가장 많은 반복수를 보인 maxmin 기준 방법의 135.8에 비해 반 정도가 된다.

세 평가함수에 의한 해는 다소 극한적인 점(extreme points)을 포함한다. Fig. 1은 Table 2의 첫번째 실행에서 생성된 최종해에 해당하는 100개 시나리오를 대상으로

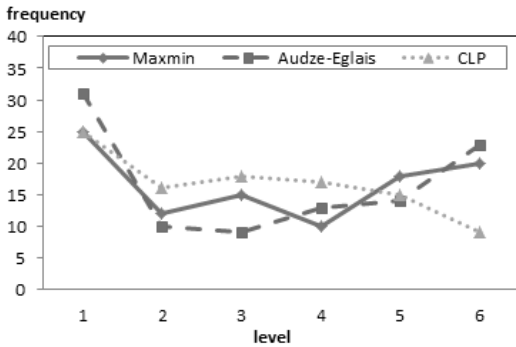


Fig. 1. Frequency of each level in V1 under three evaluation functions

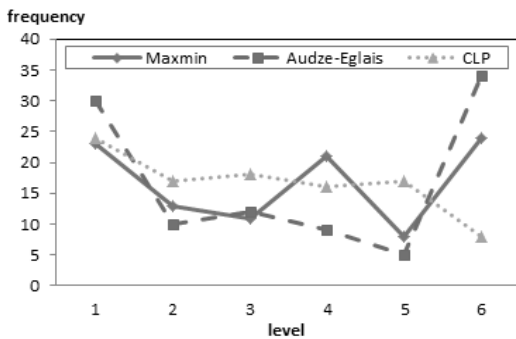


Fig. 2. Frequency of each level in V4 under three evaluation functions

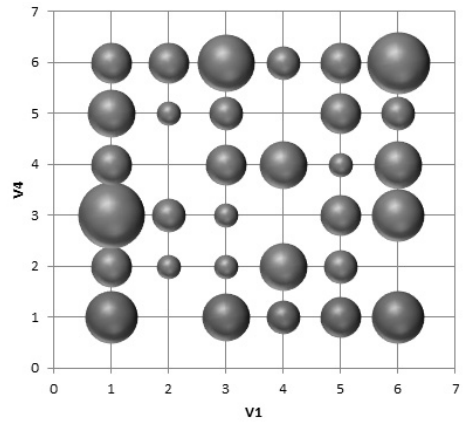


Fig. 3. Frequency of (V1,V4) in the solution under Maxmin criteria

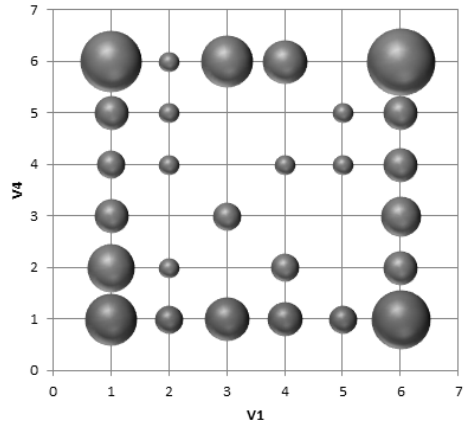


Fig. 4. Frequency of (V1,V4) in the solution under Audze-Eglais criteria

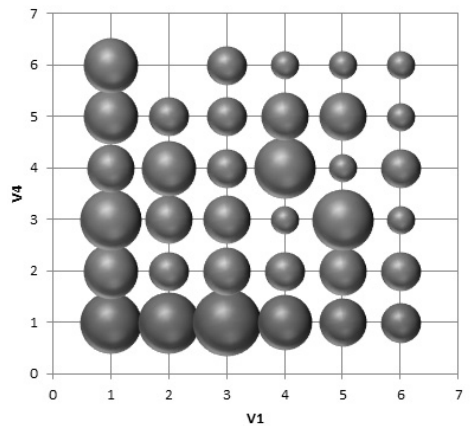


Fig. 5. Frequency of (V1, V4) in the solution under CLP criteria

변수 V1의 각 수준에 대한 빈도수를 나타낸다. Maxmin과 Audze-Eglais 평가함수에 의한 시나리오들은 가운데 수준 번호보다 양 끝의 수준 번호인 1과 6을 상대적으로 많이 포함하고 있다. Fig. 2의 변수 V4에서도 동일한 결과를 나타낸다. Fig. 1, 2에 의하면 Audze-Eglais 평가함수에 의한 해가 maxmin 보다 더 극한적인 점을 포함하고 있음을 알 수 있다. CLP 평가함수 방법에 의한 시나리오들은 큰 수준 번호보다는 작은 수준 번호를 많이 포함하고 있어 수준 번호1이 가장 많은 빈도수를 가지고, 수준 번호 6이 가장 적은 빈도수를 가진다.

Fig. 3, 4, 5는 각 방법의 해에서 시나리오들이 포함하고 있는 두 변수인 (V, V4) 조합에 대한 빈도수를 나타낸다. Fig 3, 4, 5에서 원은 빈도수의 크기를 지름으로 한다. 결과는 하나의 변수에 대한 결과와 유사하다. Maxmin 평가함수에 의한 해는 상대적으로 적은 치우침을 보이지만 Audze-Eglais 평가함수에 의한 해는 가장자리 극한 점에 치우치고, CLP 평가함수에 의한 해는 작은 값의 조합에 치우치는 분포를 나타내고 있다. 이러한 치우침은 다른 실행의 해에서도 동일한 결과를 보인다.

각 평가함수에 의한 시나리오 해들이 일부 값에 치우치는 정도에 대한 정량적 평가를 위해 균등분포 검정을 위한 카이 제곱 통계량을 이용하였다. 우선, 수준 수가 6인

변수들 V1, V4, V5 만을 대상으로 하여 각 변수의 구간 수를 3으로 하였다. 이에 따라 총 27개 구간이 설정되어 각 구간에 균등하게 포함될 시나리오 수는 100/27개 이다. 각 구간에 포함될 시나리오 수에 대한 카이 제곱 통계량을 구한 결과는 Table 3과 같다. 각 평가함수에 의한 방법에서 10회 실행하여 각 실행 결과로부터 구한 카이 제곱 통계량의 최소, 최대, 평균을 나타내었다. $\chi^2(26,0.05)$ 값이 38.89 이므로 5% 유의수준에서 maxmin과 CLP에 의한 각 10개의 해 모두 균등분포에 따른다고 볼 수 있다. 반면, Audze-Eglais에 의한 일부 해는 5% 유의수준에서 균등분포에 따른다고 볼 수 없다. Maxmin에 의한 해들은 3 평가함수에 의한 해 중 가장 작은 값을 가져 균등 분포에 가장 잘 따른다.

4.3 M의 변화에 따른 실험결과

각 평가함수에 기초한 방법에 의해 생성되는 해들은 전체 시나리오가 아닌 보다 축소된 M개의 시나리오를 대상으로 한다. 너무 적은 수의 M은 우수한 해를 포함하지 못할 확률이 커진다. Table 4는 M의 값을 200에서 1000까지 변화시켜 이 중 100개의 시나리오를 선택한 실험결과이다. 이전과 마찬가지로 각 M의 값에서 10번의 실행을 하여 평균을 구하였다.

Table 4의 결과는 M의 값이 증가함에 따라 각 평가함수 측면에서 좋은 해를 생성하는 결과를 보인다. Maxmin에 기초한 방법에서는 Z1의 평균이 M의 증가에 따라 5.8에서 8.0까지로 상승하여 향상된 해를 생성함을 나타낸다. Audze-Eglais 역시 Z2의 평균이 M의 증가에 따라 10.6에서 6.6까지 감소하여 향상된 값을 나타낸다. CLP 경우에는 Z3 평균 값이 모든 M에서 -1.7로 동일하여 M

Table 3. Chi square statistics

	Maxmin	Audze-Eglais	CLP
Min	5.3	28.52	16.64
Max	18.26	55.52	30.14
Ave	11.94	38.73	23.01

Table 4. Experiment results with the increase of M

M	Maxmin			Audze-Eglais			CLP		
	Z1	Z2	Z3	Z1	Z2	Z3	Z1	Z2	Z3
200	5.8	12.9	-1.5	3.8	10.6	-1.5	1.3	22.4	-1.7
300	6.2	11.4	-1.5	4.6	8.6	-1.5	1.5	21.8	-1.7
400	7.0	10.6	-1.4	4.9	8.0	-1.5	1.7	21.2	-1.7
500	7.0	10.3	-1.5	4.9	7.6	-1.4	1.5	21.2	-1.7
600	7.1	10.2	-1.5	5.0	7.2	-1.4	1.7	20.7	-1.7
700	7.6	9.5	-1.5	5.1	6.9	-1.5	1.6	20.8	-1.7
800	7.6	9.5	-1.5	5.2	6.9	-1.4	1.4	21.5	-1.7
900	7.7	9.5	-1.5	5.3	6.7	-1.5	1.7	21.3	-1.7
1000	8.0	8.8	-1.5	5.2	6.6	-1.4	1.8	20.4	-1.7

의 증가에 따라 향상된 해를 생성한다고 볼 수 없으나, Z1, Z2 값은 모두 M 의 증가에 따라 향상되었음을 나타낸다.

Fig. 6은 M 의 증가에 따른 알고리즘의 실행시간을 나타낸다. 실행시간 측면에서는 maxmin 방법이 M 의 증가에 따라 기하적으로 증가하는 형태를 보여 가장 많은 실행시간을 필요로 하였고, Audze-Eglais 방법은 가장 적은 시간을 필요로 하였다. 그러나 maxmin의 경우 1000개의 M 에서 100개의 시나리오를 뽑을 때 약 5분이 걸려 이 정도의 실행시간이 큰 부담이 된다고 볼 수 없어 실험 문제에 대해 알고리즘들은 대체로 효율적임을 나타낸다.

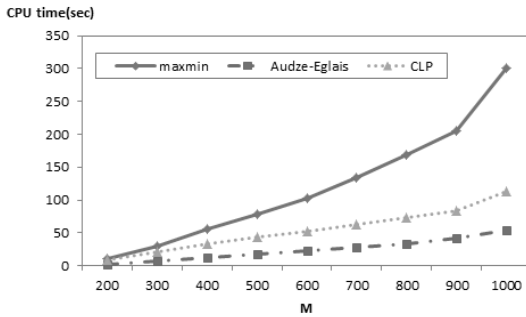


Fig. 6. Computation time with the increase of M

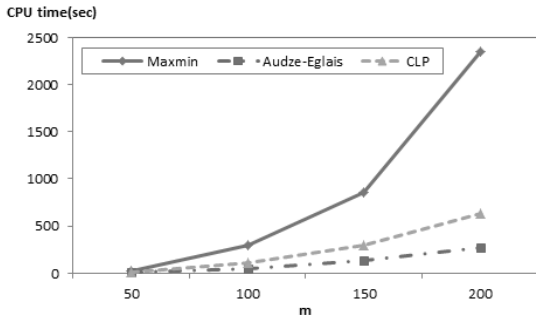


Fig. 7. Computation time with the increase of m

4.4 m 의 변화에 따른 실험결과

해에 포함되는 시나리오의 수인 m 의 값 역시 해의 구성에 큰 영향을 미친다. Table 5는 M 을 1000으로 하고, m 의 값을 50에서 200까지 변화시킨 실험결과이다. 이전과 마찬가지로 각 m 의 값에서 10번의 실행을 하여 평균을 구하였다. 해에 포함되는 시나리오의 수가 작을수록 각 평가함수 측면에서 좋은 결과를 가져온다. Maxmin의 예를 들면 시나리오들의 수가 작을수록 시나리오 간의 거리가 커져 자명한 결과를 가져온다. 그러나 CLP 경우는 예외로 평가척도인 Z3의 값이 -1.7로 변화가 없다.

알고리즘의 실행시간을 나타내는 Fig. 7에 따르면 M 의 변화에서와 마찬가지로 m 의 증가에 따라 실행시간은 기하적으로 증가한다. 특히 maxmin의 증가폭이 매우 커 m 의 변화에 따른 실행시간 민감도가 가장 크다. m 이 200일 때 실행시간은 39분으로 적지 않은 시간이 소요되었다고 볼 수 있다.

5. 결론

컴퓨터 시뮬레이션 프로젝트의 성공적인 수행을 위해 효과적인 실험계획이 필수적이다. 실험계획의 일반적인 절차는 입력변수의 값에 따른 메타모형을 수립하고 메타 모형의 파라미터 추정을 위한 샘플 시나리오를 생성하는 것이다. 이 때 전체 시나리오 공간을 대표할 수 있는 샘플 시나리오를 생성하여야 정확한 메타 모형을 수립할 수 있게 된다. 본 연구에서는 공간 채움 조건이 샘플 시나리오 생성을 위한 가장 중요한 조건으로 간주하여 이에 속하는 3가지 평가척도인 maxmin, Audze-Eglais, 그리고 centered L_2 -discrepancy의 성능을 비교하였다. 보다 효율적인 시나리오 생성을 위해 전체 시나리오가 아닌 M 개의 랜덤 시나리오를 생성하고, 이 중 m 개의 시나리오를 선정하는 방법을 택하였다. 한 시나리오의 삽입과 삭제 메커니즘에 기초한 부분 최적화 알고리즘을 적용하여 3가지 평가척도들에 대한 실험을 수행하였다.

Table 5. Experiment results with the increase of m

m	Maxmin			Audze-Eglais			CLP		
	Z1	Z2	Z3	Z1	Z2	Z3	Z1	Z2	Z3
50	11.3	1.8	-1.5	9.1	1.2	-1.3	2.9	4.5	-1.7
100	8.0	8.8	-1.5	5.2	6.6	-1.4	1.8	20.4	-1.7
150	6.5	24.1	-1.4	4.2	17.3	-1.4	0.9	49.7	-1.7
200	5.7	45.7	-1.5	3.4	34.4	-1.4	0.9	91.0	-1.7

실험 결과 3가지 평가척도에 의한 근사 최적해들은 전체 시나리오 공간에 고르게 분포되기 보다는 일부 공간에 편중되어 있는 형태를 보인다. 특히, Audze-Eglais에 의한 해는 LHD와 유사하게 시나리오 공간의 극한점에 위치하고, centered L_2 -discrepancy에 의한 해는 시나리오 공간에서 각축의 작은 값들에 편중되어 위치한다. Maxmin에 의한 해는 3가지 평가 함수에 의한 해 중 시나리오 공간에 가장 고르게 분포되어 있고, Audze-Eglais, centered L_2 -discrepancy 평가 값에서도 대체로 만족할 만한 결과를 보인다. 실행시간 측면에서는 Maxmin 방법을 제외한 두 방법들이 대체로 적은 시간이 걸리는 반면, maxmin은 M 과 m 의 증가에 따라 기하적으로 증가하여 보다 많은 시간을 필요로 한다.

일반적으로 샘플 시나리오의 수 m 은 메타모형의 정확도와 시뮬레이션 실행시간과의 관계에 따라 결정된다. 너무 적은 시나리오로는 정확한 메타모형을 추정하기 어렵고, 너무 많은 수의 시나리오는 시뮬레이션 모형의 실행 시간을 증가시킨다. 이 두 요소간의 상충관계를 고려하여 결정된 m 에 따라 전체 시나리오 해 공간의 수 M 을 결정하여야 한다. 가장 바람직한 성능을 보이는 maxmin에 기초한 방법을 사용하기 위해서는 또 다른 상충관계인 시나리오 해의 질과 시나리오 생성방법의 계산시간을 고려하여야 한다.

Reference

1. Audze, P. and Eglais, V. (1977), "New approach to planning out of experiments", Problems of Dynamics and Strength, Vol. 35, pp. 104-107.
2. Bates, S. J., Sienz, J, Langley, D. S. (2003), "Formulation of the Audze-Eglais uniform Latin hypercube design of experiments", Advances in Engineering Software, Vol. 34, pp. 493-506.
3. Fang, K., Lin, D., Winker P., Zhang, Y. (2000), "Uniform Design: Theory and Application", Technometrics, Vol. 42, No. 3, pp. 237-248.
4. Fang, K and Lin D. K. (2003), "Uniform experimental design and their applications in industry", Handbook of Statistics, Vol. 22.
5. Hickernell, F. J. (1998), "A generalized discrepancy and quadrature error bound", Mathematics of Computation, Vol. 67, pp. 299-322.
6. Johnson, M. E., Moore, L. M., and Ylvisaker, D. (1990), "Minimax and maxmin distance design", Journal of Statistical Planning and Inference, Vol. 26, pp. 131-148.
7. Johnson, R. T., Jones, B., and Fowler, J. W. (2008), "Comparing Designs for Computer Simulation Experiments", Proceedings of the 2008 Winter Simulation Conference, pp. 463-470.
8. Kleijnen, J. P. (2005), "An Overview of the design and analysis of simulation experiments for sensitivity analysis", European Journal of Operational Research, Vol. 164, No. 2, pp. 287-300.
9. Liefvendahl, M. and Stocki, R. (2006), "A study on algorithms for optimization of Latin hypercubes", Journal of Statistical and Inference, Vol. 136, pp. 3231-3247.
10. Morris, M. D. and Mitchell, T. J. (1995), "Exploratory designs for computer experiments", Journal of Statistical Planning and Inference, Vol. 43, pp. 381-402.
11. Park, J. S. (1994), "Optimal Latin-hypercube designs for computer experiments", Journal of Statistical Planning and Inference, Vol. 39, pp. 95-111.
12. Pronzato, L. and Muller, W. G. (2012), "Design of Computer Experiments: Space Filling and Beyond", Statistics and Computing, Vol 22, pp. 681-701.
13. Wim, C. M. (2005), "Kriging Metamodeling in Discrete-Event Simulation: An Overview", Proceedings of the 2005 Winter Simulation Conference, pp. 202-208.
14. Ye, K. Q., Li, W., Sudjianto, A. (2000), "Algorithmic construction of optimal symmetric Latin hypercubes", Journal of Statistical Planning and Inference, Vol. 90, pp. 145-159.



임 동 순 (dsyim@hnu.kr)

1983 한양대학교 산업공학과 학사
1986 한국과학기술원 산업공학과 석사
1991 아이오와주립대학 산업생산시스템공학과 박사
현재 한남대학교 산업경영공학과 교수

관심분야 : 생산시스템 시뮬레이션, 최적화 알고리즘



김 정 훈 (kimjh@add.re.kr)

1998 서울대학교 해양학과 학사
2001 서울대학교 지구환경과학부 석사
2001 ~ 현재 국방과학연구소 제6기술연구본부 선임연구원

관심분야 : 해군무기체계 전투실험, 합성전장환경 모의, 시뮬레이션 기반 최적화



최 봉 완 (bwchoi721@hnu.kr)

1981 해군사관학교 O.R. 학사
1985 서울대학교 산업공학과 학사
1988 아이오와주립대학 산업생산시스템공학과 석사
1994 아이오와주립대학 산업생산시스템공학과 박사

관심분야 : 군사 O.R., 체계공학(SE), 모델링&시뮬레이션