

&lt;학술논문&gt;

DOI <https://doi.org/10.3795/KSME-A.2017.41.5.381>ISSN 1226-4873(Print)  
2288-5226(Online)

## 크리깅 근사모델 기반의 중요도 추출법을 이용한 고장확률 계산 방안

이 승 규\* · 김 재 훈\*\*†

\* 한국항공우주연구원, \*\* 충남대학교 기계설계공학과

### Failure Probability Calculation Method Using Kriging Metamodel-based Importance Sampling Method

Seunggyu Lee\* and Jae Hoon Kim\*\*†

\* Korea Aerospace Research Institute

\*\* Dept. of Mechanical Engineering, Chungnam Nat'l Univ.

(Received October 4, 2016 ; Revised December 16, 2016 ; Accepted January 13, 2017)

**Key Words:** Importance Sampling(중요도 추출법), Markov Chain Simulation(마르코프 체인 시뮬레이션), Kernel Density(커널 밀도함수), Kriging Metamodel(크리깅 근사모델)

**초록:** 마르코프체인 시뮬레이션으로 추출한 점을 기반으로 커널 밀도함수를 구성하고 중요도 추출함수로 가정하였다. 크리깅 근사모델은 한계상태식 근방에서 상세히 구성되었다. 고장확률은 크리깅 근사모델에 대해 중요도 추출법을 수행하여 계산하였다. 커널 밀도함수가 한계상태식의 근방에서 더 많은 점을 추출할 수 있도록 기존의 방법을 개선하였다. 커널 밀도함수의 파라미터를 찾기 위한 안정적인 수치 계산 방안이 제시된다. 크리깅 근사모델의 불확실성으로 인해 계산된 고장확률이 변경될 가능성을 계산하여, 크리깅 근사모델의 완성도를 평가하였다.

**Abstract:** The kernel density was determined based on sampling points obtained in a Markov chain simulation and was assumed to be an important sampling function. A Kriging metamodel was constructed in more detail in the vicinity of a limit state. The failure probability was calculated based on importance sampling, which was performed for the Kriging metamodel. A pre-existing method was modified to obtain more sampling points for a kernel density in the vicinity of a limit state. A stable numerical method was proposed to find a parameter of the kernel density. To assess the completeness of the Kriging metamodel, the possibility of changes in the calculated failure probability due to the uncertainty of the Kriging metamodel was calculated.

## 1. 서 론

설계 시 불확실성(Uncertainty)을 확률변수로 가정하고, 고장확률(제품이 의도한 기능을 수행하지 못 할 확률)을 계산하는 방안에 대한 연구가 이루어져 왔다. 제품의 상태를 나타내는 성능함수(Performance Function,  $g$ )는 안전한 상태에서 양의 값, 고장 상태에서 음의 값을 가진다. 그리고 영(Zero)일 때, 한계상태식(Limit State Equation)이라고

한다.

불확실성을 확률변수벡터인  $\tilde{\theta} = [\theta_1, \theta_2, \dots, \theta_n]^T$ 로 정의하고  $\tilde{\theta}$ 의 결합확률밀도를  $p(\tilde{\theta})$ 라고 할 때, 고장확률( $P_f$ )은 식 (1)에 의해 계산된다.  $I_g(\tilde{\theta})$ 는 성능함수가 음의 값일 때 1이며, 그 외에는 영(Zero)인 지시함수(Indicator Function)이다.  $\Theta$ 는  $\tilde{\theta}$  전체 집합이다. 대부분의 경우, 식 (1)은 해석적으로 계산하기 힘들며, 수치계산을 통해 근사값을 계산한다.

$$P_f = \int_{g < 0} p(\tilde{\theta}) d\tilde{\theta} = \int_{\Theta} I_g(\tilde{\theta}) p(\tilde{\theta}) d\tilde{\theta} \quad (1)$$

† Corresponding Author, kimjhoon@cnu.ac.kr

© 2017 The Korean Society of Mechanical Engineers

식 (1)의 근사값을 계산하기 위한 방안 중에 하나가 몬테칼로 시뮬레이션(Monte Carlo Simulation)이다. 몬테칼로 시뮬레이션은 비선형 한계상태식에 강하지만, 많은 계산시간이 필요하다.<sup>(1)</sup>

이를 개선하기 위한 연구는 크게 2가지 방향으로 이루어졌다. 첫 번째가 성능함수 계산 횟수를 줄이는 방법이다. 중요도 추출법(Importance Sampling), 서브셋 시뮬레이션(Subset Simulation), 디렉셔널 시뮬레이션(Directional Simulation) 등이 대표적인 예이다. 중요도 추출법의 효율은 중요도 추출함수(Importance Sampling Function)에 따라 결정되며, Au<sup>(2)</sup>는 마르코프 체인 시뮬레이션(Markov Chain Simulation)으로 고장영역의 점을 추출하고 이를 기반으로 커널 밀도함수(Kernel Density) 형태의 중요도 추출함수를 구성하였다. 두 번째가 성능함수 1회 계산 시간을 줄이는 방안이다. 유한요소해석 등 대형모델의 해석을 통해 성능함수를 계산할 경우, 계산 시간이 많이 소요된다. 또한, 몬테칼로 시뮬레이션으로 의미있는 고장확률을 계산하기 위해서는 수천번에서 수십만번 이상의 계산이 필요하다. 이를 위해 Bucher,<sup>(3)</sup> Kim<sup>(4)</sup> 등은 성능함수를 다항식 기반의 반응평면(Response Surface)으로 근사하고 이를 고장확률 계산에 적용하였다.

Sacks<sup>(5)</sup>은 전산실험에 좀 더 적절한 크리깅 근사모델(Kriging Metamodel)을 제시하였으며, 주병현<sup>(6)</sup>은 크리깅 근사모델을 고장확률 계산을 위한 성능함수 근사에 적용하였다. 초기에는 성능함수 전 영역에 대해 크리깅 근사모델을 정확히 구성하는 연구가 진행되었으며, 최근에는 한계상태식 근방 위주로 크리깅 근사모델을 구성하는 연구가 진행되고 있다. 또한, 크리깅 근사모델을 통해 계산되는 근사값 자체의 정확도에 관심을 갖기 보다는 크리깅 근사모델이 성능함수 값의 음/양을 분류하는 정확도에 대한 연구가 진행되고 있다. 즉, 고장 여부만 정확히 구분하여, 크리깅 근사모델 구성에 소요되는 시간을 줄이고자 하는 방향으로 연구가 진행되고 있다.

본 연구에서는 마르코프 체인 시뮬레이션, 커널 밀도함수, 크리깅 근사모델을 적용한 중요도 추출법 수행 방안을 제시한다. 중요도 추출법 수행 시 한계상태식 근방에서 더 많이 점이 추출될 수 있도록 커널 밀도함수를 조정하는 방안과 커널 밀도함수의 파라미터(Parameter)를 찾을 수 있는 안정적인 수치계산 방안이 제시된다. 또한, 크리깅 근사모델의 불확실성으로 인해 계산된 고장확률이 변경될 가능성을 계산하여, 크리깅 근사모델의 완

성도를 평가하는 방안을 제시한다.

## 2. 본 론

### 2.1 중요도 추출법

몬테칼로 시뮬레이션의 계산 횟수를 줄이기 위해 중요도 추출법이 적용되어 왔다. 중요도 추출법은 좀 더 쉽게 고장영역의 점을 추출할 수 있는 중요도 추출함수를 이용하고 이를 보상해 주는 가중치(Weight Factor)를 곱한다. 식 (1)에서의  $p(\tilde{\theta})$  대신  $h(\tilde{\theta})$  분포에서 점을 추출하고 식 (2)와 같이 고장확률을 계산한다. 식 (2)에서  $p(\tilde{\theta})/h(\tilde{\theta})$ 의 값이 가중치이며,  $h(\tilde{\theta})$ 가 중요도 추출함수이다.

$$P_f = \int_{\Theta} I_g(\tilde{\theta}) \frac{p(\tilde{\theta})}{h(\tilde{\theta})} h(\tilde{\theta}) d\tilde{\theta} \quad (2)$$

중요도 추출함수가 식 (3)과 같을 때 예측된 고장확률의 분산이 영(Zero)이므로 가장 이상적이지만, 식 (3)에 포함된  $P_f$  값이 식 (1)에서 궁극적으로 알고자 하는 값이므로  $h_{opt}(\tilde{\theta})$ 를 알 수 없다.<sup>(2)</sup> 따라서,  $h_{opt}(\tilde{\theta})$ 에 가까운 분포를 가정하여 식 (2)에 적용한다.

$$h_{opt}(\tilde{\theta}) = I_g(\tilde{\theta})p(\tilde{\theta})/P_f \quad (3)$$

### 2.2 적응적 중요도 추출법(Adaptive Importance Sampling)

본 연구에서는 Au<sup>(2)</sup>가 제시한 적응적 중요도 추출법을 적용하였다. Au<sup>(2)</sup>는 메트로폴리스 알고리즘(Metropolis Algorithm)을 통해 마르코프 체인 시뮬레이션을 수행하여 고장영역의 점들을 용이하게 추출하였다. 또한, 이 점들을 중심으로 하여 2.3절의 커널 밀도함수를 구성하고 이를 식 (2)의 중요도 추출함수로 가정하였다. 메트로폴리스 알고리즘은 개략적으로 다음과 같은 절차로 수행한다.

#### 메트로폴리스 알고리즘

(1) 고장영역의 임의의 점을 시작점으로 선택한다. 이 점은 연구자의 직관에 의해 선택할 수 있으며, 결과에 거의 영향을 주지 않는 것으로 알려져 있다.

(2) 현재 점이  $\tilde{\theta}^{(i)}$ 라고 한다면,  $\tilde{\theta}^{(i)}$ 를 중심으로 하는 균일분포에서 한 개의 점( $\xi$ )을 추출한다. 균일분포의 길이는 Au<sup>(2)</sup>의 논문 식 (12)를 적용하였다.

(3)  $\eta = I_g(\xi)p(\xi)/I_g(\tilde{\theta}^{(i)})p(\tilde{\theta}^{(i)})$ 를 계산한다.

(4)  $\eta \geq 1$ 이면  $\xi$ 를 마르코프 체인 시뮬레이션의 다음 점( $\tilde{\theta}^{(i+1)}$ )으로 선택한다.

(5)  $\eta < 1$ 이면,  $\eta$ 의 확률로  $\xi$ 를 마르코프 체인 시뮬레이션의 다음 점( $\tilde{\theta}^{(i+1)}$ )으로 선택한다.

상기 절차에서 마르코프 체인 시뮬레이션의 다음 점으로 선택된 점을 “수락점(Accept Points)”라고 하며, 본 연구에서는  $M$ 개를 추출하였다. 다음 점으로 선택되지 못한 점을 “기각점(Reject Points)”라고 하고, 그 개수는  $M_R$ 라고 하였다.

### 2.3 커널 밀도함수

본 연구에서는  $Au^{(2)}$ 가 적용한 방법을 응용하여 식 (5)의 정규분포인 커널(Kernel)의 선형결합으로 식 (4)의 커널 밀도함수를 구성하였다. 마르코프 체인 시뮬레이션에서 추출한  $M$ 개 점의 공분산 행렬(Covariance Matrix)를 식 (5)의  $\mathbf{S}$ 에 적용하였다. 식 (4), (5)에서  $n$ 은 확률변수의 개수이다.

$Au^{(2)}$ 는 모든 커널에서 균일하게 점들을 추출할 것을 제안하였다. 하지만, 본 연구에서는 일부 한 계상대식 근방의 커널에서만 점을 추출하였다. 이를 위해, 마르코프 체인 알고리즘으로 생성된 점들에 대해 원점과의 거리의 평균( $\mu$ )과 표준편차( $\sigma$ )를 계산하고  $\mu + \sigma$  이내의 점들을 중심으로 하는 커널에서만 점을 추출하였다.

본 연구에서는 전체  $M$ 개의 마르코프 체인 시뮬레이션의 추출점 중에서 커널 밀도함수 구성에 적용된 점의 수를  $M_k$ 라고 하였다. 본 연구에서는 커널 밀도함수에서 추출된 점을 기반으로 크리깅 근사모델의 완성도를 평가한다. 원점과 멀리 떨어진 점은 고장확률에 미치는 영향이 적다. 또한, 크리깅 근사모델을 상세화하기 위해서는 성능함수 계산횟수가 증가한다. 따라서, 중심이 원점에서  $\mu + \sigma$ 보다 떨어진 커널에서는 점들을 추출하지 않았다.

$$\mathbf{k}(\tilde{\theta}) = \frac{1}{M_k} \sum_{i=1}^{M_k} \mathbf{K}_i(\tilde{\theta}) \quad (4)$$

$$\mathbf{K}_i(\tilde{\theta}) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_i|}} \exp \left[ -\frac{1}{2} (\tilde{\theta} - \tilde{\mu}_i)^T \Sigma^{-1} (\tilde{\theta} - \tilde{\mu}_i) \right] \quad (5)$$

where,  $\Sigma_i = \mathbf{S} \cdot \mathbf{w}^2 \cdot \lambda_i^2$

식 (5)에서  $w$ 는 모든 커널의 공분산을 조정해 주기 위한 값이며,  $Au^{(2)}$ 가 제시한 방안을 적용하여 최적의 값을 찾았다.  $\lambda_i$ 는 각 커널의 공분산을 조정해 주기 위한 값이다. 본 연구에서는  $Au^{(2)}$ 가 제시한  $\lambda_i$  계산 방안을 수치 계산에 좀더 적합한 방법으로 개선하여 적용하였다.

$$\lambda_i = \left\{ \frac{[\prod_{j=1}^M p(\tilde{\theta}^{(j)})]^{1/M}}{p(\tilde{\theta}^{(i)})} \right\}^\alpha \quad (6)$$

$Au^{(2)}$ 는 식 (6)을 통해 식 (5)의  $\lambda_i$ 를 계산하였다.  $p(\tilde{\theta})$ 가 작고,  $M$ 이 클 경우, 식 (6)의 수치 계산값은 영(Zero)이 된다. 이러한 오류를 방지하기 위해 본 연구에서는 log를 적용한 식 (7)을 통해  $\lambda_i$ 를 계산하였다.  $\alpha$ 는  $Au^{(2)}$ 가 제시한 0.5를 적용하였다.

$$\begin{aligned} \lambda_i^* &= \log(\lambda_i) \\ &= \alpha \left[ \frac{1}{M} \sum_{j=1}^M \log\{p(\tilde{\theta}^{(j)})\} - \log\{p(\tilde{\theta}^{(i)})\} \right] \\ \lambda_i &= 10^{\lambda_i^*} \end{aligned} \quad (7)$$

### 2.4 크리깅 근사모델<sup>(5-7)</sup>

성능함수 계산 시간을 줄이기 위해, 크리깅 근사모델을 적용하였다. 크리깅 근사모델은 다항식 근사모델에 편차값을 더한 것으로 식 (8)과 같다. 식 (8)의  $\mathbf{F}$ 는 식 (9), (10)과 같다.  $\tilde{f}$ 는  $q$ 개의 함수로 이루어진 벡터이며,  $\mathbf{F}$ 는  $s$ 개의 실험점(성능함수의 입력값과 출력값을 모두 알고 있는 점)이 있을 경우,  $s \times q$  행렬이다.  $\tilde{\beta}$ 는 최소자승법의 결과로 나오는 계수와 같다.  $\tilde{Z}$ 는 가우시안 프로세스(Gaussian Process)이며, 평균은 영(Zero)이고, 분산은  $\sigma_z^2$ 인 정규분포이다.  $\tilde{Z}$ 의 공분산은 식 (11)과 같다.

$$\tilde{Y} = \mathbf{F}\tilde{\beta} + \tilde{Z} \quad (8)$$

$$\tilde{f} = [\zeta_1(\tilde{x}), \zeta_2(\tilde{x}), \dots, \zeta_q(\tilde{x})]^T \quad (9)$$

$$\mathbf{F} = \begin{bmatrix} \tilde{f}^T(\tilde{x}_1) \\ \tilde{f}^T(\tilde{x}_2) \\ \vdots \\ \tilde{f}^T(\tilde{x}_s) \end{bmatrix} \quad (10)$$

$$\text{Cov}(\tilde{Z}(\tilde{x}_i), \tilde{Z}(\tilde{x}_j)) = \sigma_z^2 \mathbf{C} \quad (11)$$

여기서,  $\mathbf{C}$ 는 식 (12)의 상관함수(Correlation Function)  $R(\tilde{x}_i, \tilde{x}_j)$ 를  $i, j$  번째 요소로 갖는 행렬이다. 실험점이  $s$ 개라면,  $\mathbf{C}$ 의 크기는  $s \times s$ 이다. 상관함수는 크리깅 근사모델의 부드러운 정도(Smoothness)에 영향을 주며,  $d$  차원의 확률변수에 대해 식 (12)의 가우시안 상관함수(Gaussian Correlation Function)가 적용된다.

$$\mathbf{R}(x_i, x_j) = \exp \left( -\sum_{\gamma=1}^d \rho_\gamma |x_\gamma^i - x_\gamma^j|^2 \right) \quad (12)$$

$x_\gamma^i$ 는  $i$  번째 실험점의  $\gamma$  번째 요소이다.  $\rho_\gamma$ 는 스케일 파라미터(Scale Parameter)이다.  $\hat{\beta}$ 와  $\hat{\sigma}_z^2$ 의 추정량인  $\hat{\beta}$ 와  $\hat{\sigma}_z^2$ 는 식 (13), (14)과 같다.

$$\hat{\beta} = (\mathbf{F}^T \mathbf{C}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{C}^{-1} \tilde{\mathbf{y}} \quad (13)$$

$$\hat{\sigma}_z^2 = 1/s (\tilde{\mathbf{y}} - \mathbf{F} \hat{\beta})^T \mathbf{C}^{-1} (\tilde{\mathbf{y}} - \mathbf{F} \hat{\beta}) \quad (14)$$

$\tilde{\mathbf{y}} = [y_1, y_2, \dots, y_s]^T$ 는 실험점에서 성능함수의 출력값을 요소로 갖는 벡터이다.  $\hat{\beta}$ 와  $\hat{\sigma}_z^2$ 는 식 (12)의  $\rho_\gamma$ 가 결정된 이후, 결정할 수 있다.  $\rho_\gamma$ 의 추정량은 최대우도추정(Maximum Likelihood Estimation)법에 의해 결정되며, 식 (15)와 같다.

$$\begin{aligned} \text{Max. } L(\rho) &= -\frac{1}{2} [s \cdot \ln(\hat{\sigma}_z^2) + \ln\{\det(\mathbf{C})\}] \\ \text{Subject to } \rho_\gamma &\geq 0 \text{ for } \gamma = 1, \dots, d \end{aligned} \quad (15)$$

임의의 입력값에 대한 성능함수의 추정값은 식 (16)과 같다.  $\tilde{\mathbf{r}}_x$ 는 임의의 입력값과 실험점 사이의 상관벡터이며, 식 (17)과 같다.

$$\hat{\mathbf{y}}(x) = \tilde{\mathbf{f}}_x^T \hat{\beta} + \tilde{\mathbf{r}}_x^T \mathbf{C}^{-1} (\tilde{\mathbf{y}} - \mathbf{F} \hat{\beta}) \quad (16)$$

$$\tilde{\mathbf{r}}_x = [\mathbf{R}(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}), \mathbf{R}(\tilde{\mathbf{x}}_2, \tilde{\mathbf{x}}), \dots, \mathbf{R}(\tilde{\mathbf{x}}_s, \tilde{\mathbf{x}})]^T \quad (17)$$

크리깅 근사모형을 통해 임의의 점에 대한 성능함수의 추정값과 평균제곱오차(MSE, Mean Square Error)인 식 (18)을 계산할 수 있다.

$$\hat{\sigma}_y^2 = \hat{\sigma}_z^2 \left( \mathbf{1} - [\tilde{\mathbf{f}}_x^T \quad \tilde{\mathbf{r}}_x^T] \begin{bmatrix} \mathbf{0} & \mathbf{F}^T \\ \mathbf{F} & \mathbf{C} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{f}}_x \\ \tilde{\mathbf{r}}_x \end{bmatrix} \right) \quad (18)$$

## 2.5 크리깅 근사모형의 상세화 및 종료 기준

Echard<sup>(8)</sup>는 식 (19)의 학습함수(Learning Function)를 정의하고, 학습함수값이 가장 작은 점을 실험점으로 추가하여 크리깅 근사모형을 상세화하는 방안을 제시하였다.  $\hat{g}(\tilde{\theta})$ 는 점  $\tilde{\theta}$ 에서 크리깅 근사모형의 추정값이다.  $\sigma_{\hat{g}}(\tilde{\theta})$ 는 점  $\tilde{\theta}$ 에서 크리깅 근사모형 추정값의 불확도이며, 식 (18)의 제곱근이다.

$$U(\tilde{\theta}) = \frac{|\hat{g}(\tilde{\theta})|}{\sigma_{\hat{g}}(\tilde{\theta})} \quad (19)$$

Echard<sup>(8)</sup>는 크리깅 근사모형을 통해 추정하는 모든 점들의 학습함수값 중 최소값이 2 이상이 될 때까지 크리깅 근사모형을 상세화할 것을 제시하

였다. 본 연구에서는 크리깅 근사모형이 예측하는 고장확률이 변경될 가능성을 계산하고 이를 크리깅 근사모형의 상세화 종료 기준으로 삼았다.

크리깅 근사모형은 가우시안 프로세스이다. 가우시안 프로세스는 추정하는 모든 값이 정규분포를 갖는다고 가정하며, 추정되는 값은 정규분포의 평균값이다.<sup>(7)</sup> 식 (19)의 학습함수값은 크리깅 근사모형 추정값이 영(Zero)으로부터 떨어진 통계적 거리이다. 이는 크리깅 근사모형이 성능함수의 음/양을 예측하는 능력에 대한 불확도로 볼 수 있다. 따라서, 식 (20)은 크리깅 근사모형이 추정하는 모든 점들에서 추정값의 부호가 변경될 확률의 평균값이다. 본 연구에서는 크리깅 근사모형 상세화를 종료하는 기준으로 식 (20)의 값을 적용하였다.

$$T = 1/N_T \sum_{i=1}^{N_T} 1 - \Phi(u_{\theta_i})$$

where,

$\Phi(\cdot)$  : Cumulative distribution function of normal distribution

$u_{\theta_i}$  : Value of learning function at point  $\tilde{\theta}_i$  (20)

$N_T$  : Number of total point of kriging metamodel prediction

$T$ 값이 작을수록 크리깅 근사모형의 정확도는 높아지지만, 수치비용이 증가한다. 따라서, 본 연구에서는  $T$ 값의 기준으로 0.005 또는 0.01의  $T$ 값을 적용하였으며, 이 때 기존 연구결과와 동등하거나 개선된 결과를 얻을 수 있음을 수치예제를 통해서 확인하였다.

## 2.6 고장확률 계산 절차

본 연구에서는 Au<sup>(2)</sup>가 제시했던 계산 절차를 개선하고 크리깅 근사모형을 구성하는 절차를 추가하여 고장확률을 계산하였다.

**Step 1)** 2.2절에 제시된 메트로폴리스 알고리즘에 따라  $M$ 개의 점을 추출한다.

**Step 2)** 2.3절에 제시된 커널 밀도함수를 구성하고 점들을 추출한다. 각 커널에서  $N$ 개씩 총  $M_k \times N$ 개를 추출한다.

**Step 3)** 2.4절에 제시된 크리깅 근사모형을 구성한다. 크리깅 근사모형의 실험점은 Step1)에서 생성된  $M$ 개의 수락점과  $M_R$ 개의 기각점이다.

**Step 4)** Step2)에서 추출된  $M_k \times N$ 개의 점들에서 크리깅 근사모형을 계산하고 식 (20)에 제시된  $T$ 를 계산한다.  $T$ 가 일정한 기준값보다 크면,  $M_k \times N$ 개의 점들 중에서 식 (19)의 학습함수값이 가장 작은 값을 실험점으로 추가하여 크리깅 근사모형을

상세화 한다.  $T$ 가 기준값을 만족할 때까지 이 과정을 반복한다. 본 연구에서는  $T$ 가 0.005 또는 0.01보다 작을 때까지 이 과정을 반복하였다.

**Step 5)** 고장확률은 식 (21), (22)에 따라 계산하고, 분산은 식 (23), (24)에 따라 계산한다.  $\hat{\theta}_j^i$ 는  $i$ 번째 커널에서 추출한  $j$ 번째 점이다. 식 (25)의 변동계수(C.O.V, Coefficient Of Variance)가 0.05보다 클 경우, Step 2)에서  $N$ 을 증가시켜서 다시 절차에 따라 고장확률을 계산하였다.

$$\hat{P}_f = \frac{1}{M_k} \sum_{i=1}^{M_k} \hat{P}_{k,i} \quad (21)$$

$$\hat{P}_{k,i} = \frac{1}{N} \sum_{j=1}^N I_g(\hat{\theta}_j^i) \frac{p(\hat{\theta}_j^i)}{K_i(\hat{\theta}_j^i)} \quad (22)$$

각 커널의 분산을 구한 다음 전체의 분산을 구한다. 중심극한정리 및 확률변수 선형합의 분산 정리에 따라 고장확률은 분산은 식 (23), (24)와 같다. 단, 각 추출된 점들 사이의 공분산은 무시되었다.

$$Var(\hat{P}_f) = \frac{1}{M_k^2} \sum_{i=1}^{M_k} Var(\hat{P}_{k,i}) \quad (23)$$

$$Var(\hat{P}_{k,i}) = \frac{1}{N} \frac{1}{N-1} \sum_{j=1}^N \left[ I_g(\hat{\theta}_j^i) \frac{p(\hat{\theta}_j^i)}{K_i(\hat{\theta}_j^i)} - \hat{P}_{k,i} \right]^2 \quad (24)$$

$$C.O.V = \sqrt{Var(\hat{P}_f) / \hat{P}_f} \quad (25)$$

식 (25)의 분모는 Strong Law of Large Numbers에 따라  $N$ 이 증가함에 따라 고장확률에 수렴하여 거의 변하지 않으며, 식 (25)의 분자는  $N$ 이 증가함

에 따라 감소하므로, 식 (25)의 C.O.V는  $N$ 이 증가함에 따라 감소한다.

### 3. 수치예제

#### 3.1 수치예제 1

식 (26)은 Dubourg<sup>(9)</sup> 논문에서 실린 예제이다. 확률변수는 평균이 1이고 표준편차가 0.2인 대수정규(Lognormal) 분포이며, 서로 독립이다. 한계상태식은 선형이지만, 확률변수가 대수정규분포이므로, 기존의 FORM(First Order Reliability Method) 방법을 적용할 경우, 비선형을 유발한다.

결과는 Table 1과 같다.  $N_{LSE}$ 는 성능함수의 계산 횟수이며, C.O.V는 식 (25)에 정의된 값이다. Meta-IS는 Dubourg<sup>(9)</sup>가 제시하였으며, 적응적 중요도 추출법과 크리깅 근사모델을 접목한 방법이다. Meta-IS는 본 논문에서와 달리 “Cross-validation” 기법을 통해 크리깅 근사모델의 완성도를 평가한다.

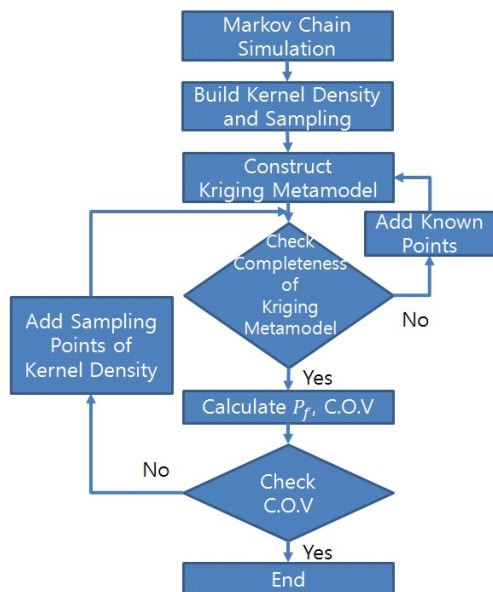
$$g = 2 + 3 \times 0.2\sqrt{2} - \theta_1 - \theta_2 \quad (26)$$

본 논문에서 제시하는 방법이 좀 더 적은 성능함수 계산량으로 고장확률을 계산한 것을 확인할 수 있다. 크리깅 근사모델에서 식 (8)의  $\mathbf{F}$ 는 상수를 가정하였다.

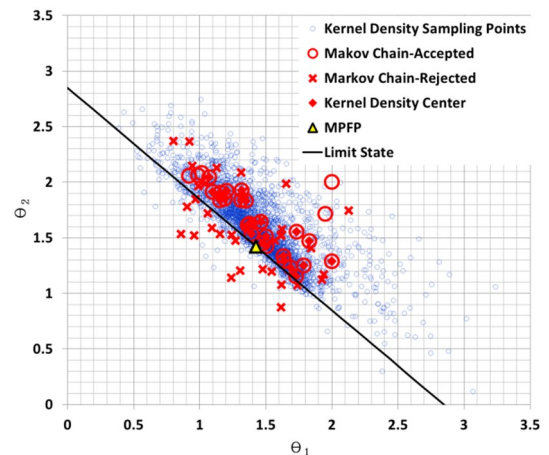
추출된 점은 Fig. 2와 같다. 마르코프 체인 시물

**Table 1** Result of example 1

	Reference		Proposed method
	Monte carlo	Meta-IS	
$N_{LSE}$	500,000	121	65
$P_f$	$4.98 \times 10^{-3}$	$5.02 \times 10^{-3}$	$5.04 \times 10^{-3}$
C.O.V	<2%	<2%	<2%



**Fig. 1** Procedure



**Fig. 2** Sampling result of example 1

레이션에 의해 고장영역 중에서도 한계상태식과 MPFP(Most Probable Failure Point) 근방의 점들이 대부분 추출된 것을 확인할 수 있다. MPFP는 표준 정규분포 확률변수 공간에서 원점과 가장 가까운 한계상태식 위의 점으로, MPFP 주변 영역이 고장 확률에 미치는 영향이 큰 것으로 알려져 있다. 한계상태식에서 상대적으로 멀리 떨어진 일부 점들은 제외하고 커널의 중심으로 선정된 것을 볼 수 있다. 이는 고장확률에 영향력이 큰 한계상태식 근방의 점들만 고려하여 수치비용을 줄이는 효과가 있다.

마르코프 체인 시뮬레이션 결과 중 수락점과 기각점 모두 크리깅 근사모델의 실험점으로 활용되었다. 커널 밀도함수에서 추출된 점은 대부분 고장영역에 분포된 것을 확인할 수 있다. 식 (3)의 이상적인 중요도 추출함수는 지시함수를 포함하고 있으므로, 고장영역의 점들만 추출하게 되어 있다. 이러한 관점에서 볼 때, Fig. 2의 커널 밀도함수는 이상적인 중요도 추출함수를 잘 모사하고 있는 것을 확인할 수 있다.

3.2 수치예제 2

식 (27)는 Cadini<sup>(10)</sup> 논문에서 실린 예제이다. 확률변수는 서로 독립인 표준정규분포이다. MPFP가 1개 존재하는 비선형 한계상태식이다.

$$g = 0.5(\theta_1 - 2)^2 - 1.5(\theta_2 - 5)^3 - 3 \quad (27)$$

Table 2 Result of example 2

	Reference		Proposed Method
	Monte carlo	Cadini(2014)	
$N_{LSE}$	$5 \times 10^7$	28	17
$P_f$	$2.85 \times 10^{-5}$	$2.87 \times 10^{-5}$	$2.86 \times 10^{-5}$
C.O.V	2.64 %	2.39 %	2.15 %

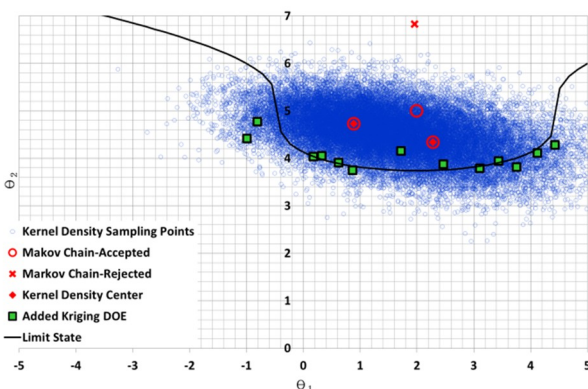


Fig. 3 Sampling result of example 2

결과는 Table 2와 같다. 좀 더 적은 성능함수의 계산 횟수로 기존 연구결과와 유사한 결과를 얻었다. Cadini<sup>(10)</sup>의 방법이 중요도 추출법에 크리깅 근사모델을 접목시킨 것은 본 연구와 같지만, Cadini<sup>(10)</sup>는 식 (3)의 중요도 추출함수를 모사하기 위해 크리깅 근사모델을 적용하였다.

추출된 점은 Fig. 3과 같다. 마르코프 체인 시뮬레이션 결과 중 수락점과 기각점 모두가 크리깅 근사모델의 실험점으로 활용되었다. 크리깅 근사모델 실험점으로 추가된 점들이 주로 고장확률에 영향이 큰 한계상태식 근방에서 선택된 것을 확인할 수 있다.

3.3 수치예제 3

식 (28)는 Au<sup>(2)</sup> 논문에서 실린 예제이다. 확률변수는 표준정규분포이며, 서로 독립이다. 식 (28)에는 3개의 MPFP가 존재하지만, Au<sup>(2)</sup>, Zhao<sup>(11)</sup>는 2개의 MPFP에 대한 고장확률을 계산하였다. 기존 연구와 비교하기 위해 본 논문에서도 2개의 MPFP에 대한 고장확률을 계산하였다. FORM 기반의 방법으로는 2개 이상의 MPFP를 쉽게 찾기 힘들다고 알려져 있다. 반면, 본 논문에서와 같이 마르코프 체인 시뮬레이션을 활용하면, 2개 이상의 MPFP를 좀 더 쉽게 찾고, 이를 고장확률 계산에 반영할 수 있다.

크리깅 근사모델에서 식 (8)의 **F**는 상수를 가정하였다. 본 논문에서 제시하는 방법이 기존 연구

Table 3 Result of example 3

	Reference		Proposed Method
	Au(1999)	Zhao(2015)	
$N_{LSE}$	500	163	110
$P_f$	$2.68 \times 10^{-3}$	$2.69 \times 10^{-3}$	$2.62 \times 10^{-3}$
C.O.V	7%	5%	3%

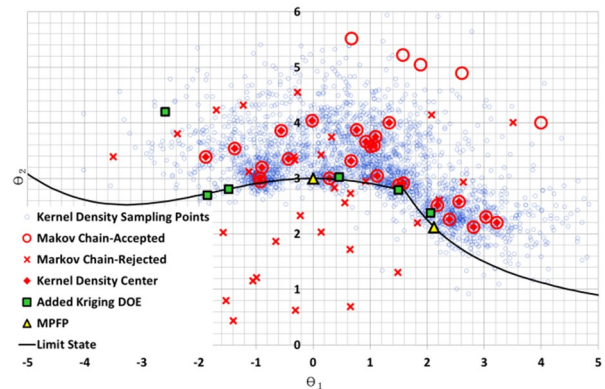


Fig. 4 Sampling result of example 3

결과에 비해 적은 성능함수 계산 횟수로 고장확률을 계산한 것을 확인할 수 있다.

$$\begin{aligned}
 g &= \min(g_1, g_2) \\
 g_1 &= 2 - \theta_2 + \exp\left(-\theta_1^2/10\right) + \left(\theta_1/5\right)^4 \\
 g_2 &= 4.5 - \theta_1\theta_2
 \end{aligned}
 \tag{28}$$

추출된 점은 Fig. 4와 같다. 마르코프 체인 시뮬레이션에 의해 고장영역 중에서도 경계와 MPFP 근방의 점들이 대부분 추출된 것을 확인할 수 있다. 일부 한계상태식에서 상대적으로 멀리 떨어진 점들은 제외하고 커널의 중심으로 선정된 것을 볼 수 있다. 이는 고장확률에 영향력이 한계상태식 근방의 점들만 고려하여 수치비용을 줄이는 효과가 있다. Au<sup>(2)</sup>에 따르면, MPFP는 (0,3), (3/√2, 3/√2)이다. MPFP 주위에 커널 중심이 잘 분포된 것을 볼 수 있다. 이로 인해 좀 더 정확한 고장확률을 계산할 수 있었다.

마르코프 체인 시뮬레이션 결과 중 수락점과 기각점 모두가 크리깅 근사모델의 실험점으로 활용되었다. 크리깅 근사모델 실험점으로 추가된 점들이 주로 고장확률에 영향이 큰 한계상태식 근방에서 선택된 것을 확인할 수 있다.

다른 예제들과 같이 커널 밀도함수에서 추출된 점은 대부분 고장영역에 분포된 것으로 이상적인 중요도 추출함수를 잘 모사하고 있는 것을 확인할 수 있다.

3.4 수치예제 4

식 (29)을 통해 변형률이 주어졌을 때, 피로 수명을 계산할 수 있다. 조태민<sup>(12)</sup> 등은 피로수명의 변동을 반영하기 위해 식 (29)의 재료상수를 확률변수로 가정하고, 이미 정해진 피로수명을 만족하지 못할 확률을 계산하였다. Meggiolaro<sup>(13)</sup>는 시험을 통하여 금속재에 대해서 식 (29)의 피로재료상수는 Log-logistic 분포와 가장 유사하다는 연구결

과를 발표하였다. 본 연구에서는 피로재료상수는 Log-logistic 분포와 유사하다고 알려진 대수정규분포로 가정하고 식 (31)의 성능함수에 대한 고장확률을 계산하였다. 응력의 크기를 상수로 가정하고 식 (30)을 통해 변형률 크기를 계산하였다. 이를 식 (29)에 대입하여 수명을 계산하였다. 본 예제의 식 (29), (30)은 피로수명 계산을 위한 상용 코드로 대체될 수도 있다.

본 연구에서 제시한 방안으로 성능함수를 676번 계산하고 고장확률 0.0945%를 얻었다. 몬테칼로 시뮬레이션 결과와 C.O.V를 고려할 때, 본 연구에서 제시한 방안으로 얻은 고장확률이 합리적인 결과임을 확인할 수 있다.

$$\frac{\Delta\varepsilon}{2} = \frac{\sigma_f'}{E} (2N_f)^{-b} + \varepsilon_f' (2N_f)^{-c}
 \tag{29}$$

$$\Delta\varepsilon = \frac{\Delta\sigma}{E} + 2 \left(\frac{\Delta\sigma}{2K'}\right)^{\frac{1}{n'}}
 \tag{30}$$

$$g = N_f - 1000
 \tag{31}$$

3.5 수치예제 5

본 예제에서는 식 (6)을 통해 커널 밀도함수의 파라미터인 벡터 λ<sub>i</sub>를 계산할 때, 벡터 λ<sub>i</sub>의 모든 요소가 영(Zero)이 되는 오류가 발생하지 않는 빈도와 식 (7)의 효용성을 살펴보고자 한다. 식 (6)을 통해 계산된 벡터 λ<sub>i</sub>의 모든 요소가 영(Zero)이 되는 이유는 식 (6)에 포함된 항인 식 (32)의 수치계산 결과가 영(Zero)되기 때문이다. 따라서, 식 (32)의 수치계산 결과가 영(Zero)이 되지 않는 빈도를 조사하였다.

$$\prod_{j=1}^M p(\tilde{\theta}^{(j)})
 \tag{32}$$

식 (32)의 M이 30, 40, ..., 100 일 때에 대해 수치계산을 수행하였다. 식 (32)는 마르코프 체인 시뮬레이션에 의해 고장영역에서 추출된 M개의 점들에 대해 적용된다. 마르코프 체인 시뮬레이션은 난수 추출 알고리즘이므로 수행할 때마다, 추출된 점들이 다르다. 따라서, 각 M에 대해 100번의 마르코프 체인 시뮬레이션 및 식 (32)의 수치계산 수행 후 영(Zero)이 되지 않는 빈도를 조사하였다.

Table 5 Result of example 4

	Monte carlo	Proposed method
N <sub>LSE</sub>	5 × 10 <sup>5</sup>	676
P <sub>f</sub>	1.00 × 10 <sup>-3</sup>	0.945 × 10 <sup>-3</sup>
C.O.V	< 5%	< 5%

Table 4 Random variables of example 4

Random Variable	Distribution	Mean	Standard Deviation
σ <sub>f</sub> '	Lognormal	709 MPa	Mean × 0.05
b	Lognormal	0.056	
ε <sub>f</sub> '	Lognormal	0.12	
c	Lognormal	0.75	
E	Normal	71.7 GPa	
Δσ	Deterministic	820 MPa	0
K'	Deterministic	787 MPa	0
n'	Deterministic	0.07	0

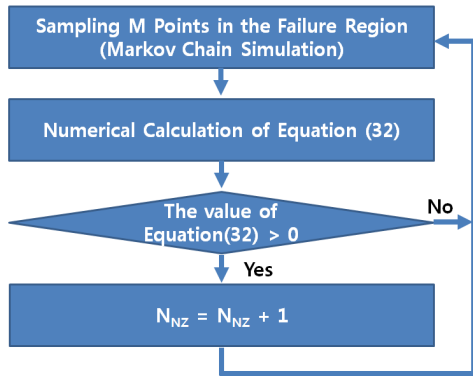


Fig. 5 Procedure of example 5

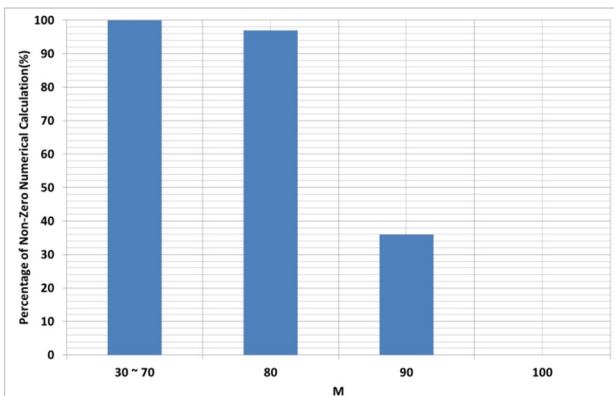


Fig. 6 Percentage of non-zero numerical calculation result

상세한 절차는 Fig. 5와 같다.

Fig. 6에서 Y축은  $N_{NZ}/100$  (%)이다.  $M$ 이 70보다 커지면서 영(Zero)이 아닌 값을 얻는 비율이 점점 낮아지는 것을 확인할 수 있다.  $M$ 이 증가하는 것은 식 (32)에서 작은  $p(\cdot)$  값이 곱해지는 횟수가 증가하는 것을 의미하므로, 타당한 결과이다. 식 (7)을 적용했을 때는 모든  $M$  값에 대해 100% 영(Zero)이 아닌 파라미터를 얻을 수 있었다.

#### 4. 결 론

최근 들어, 마르코프 체인 시뮬레이션과 크리깅 근사모형을 중요도 추출법에 적용한 고장확률 계산 방안이 연구되어 왔다. 마르코프 체인 시뮬레이션은 다중 MPFP를 좀 더 쉽게 찾는 수단이며, 중요도 추출법과 크리깅 근사모형은 수치계산 비용을 줄이기 위한 방안이다.

본 연구에서는 아래의 사항들을 고려하여, 기존의 평가 방안을 개선할 수 있음을 확인하였다.

(1) 마르코프 체인 시뮬레이션으로 추출된 점들의 통계값을 바탕으로 한계상태식에 좀 더 근접한

점들만 고장확률 계산에 반영하였다.

(2) 중요도 추출함수를 구성하기 위해 필요한 커널 밀도함수의 파라미터  $\lambda_i$ 를 안정적으로 수치 계산할 수 있는 방안을 적용하였다.

(3) 고장확률이 변화될 가능성을 기준으로 크리깅 근사모형의 완성도를 평가하였다.

동일한 수치예제에 대해서 중요도 추출법과 크리깅 근사모형을 접목시킨 기존의 연구결과와 비교하고 성능함수 계산 횟수가 줄어든 것을 확인할 수 있었다. 2차원 문제에서 추출된 점들을 볼 때, 개선된 방안으로 계산된 파라미터  $\lambda_i$ 가 반영된 커널 밀도함수가 이상적인 중요도 추출함수를 잘 묘사하는 것을 확인할 수 있었다.

#### 참고문헌 (References)

- (1) Bichon, B. J., Eldred, M. S., Swiler, L. P., Mahadevan, S. and McFarland, M., 2008, "Efficient Global Reliability Analysis for Nonlinear Implicit Performance Functions," *AIAA Journal*, Vol. 46, No. 10, pp. 2459~2468.
- (2) Au, S. K. and Beck, J. L., 1999, "A New Adaptive Importance Sampling Scheme for Reliability Calculations," *Structural Safety*, Vol. 21, pp. 135~158.
- (3) Bucher, C. G. and Bourgund, U., 1990, "A Fast and Efficient Response Surface Approach for Structural Reliability Problems," *Structural Safety*, Vol. 7, pp. 75~66.
- (4) Kim, S.-H. and Na, S.-W., 1997, "Response Surface Method using Vector Projected Sampling Points," *Structural Safety*, Vol. 19, No. 1, pp. 3~19.
- (5) Jerome Sacks, Susannah B. Schiler and William J. Welch, 1989, "Designs for Computer Experiments," *Technometrics*, Vol. 31, No. 1, pp. 41~47.
- (6) Ju, B.-H., 2008, "Reliability Based Design Optimization Using a Kriging Metamodel and a Moment Method," *Korea Advanced Institute of Science and Technology*.
- (7) Santner, T. J., Williams, B. J. and William, I. N., 2003, "The Design and Analysis of Computer Experiments," *Springer*, Chapter 2~4.
- (8) Echard, B., Gayton, N. and Lemaire, M., 2011, "AK-MCS: An Active Learning Reliability Method Combining Kriging and Monte Carlo Simulation," *Structural Safety*, Vol. 33, pp. 145~154.
- (9) Dubourg, V., Sudret, B. and Deheeger, F., 2013, "Metamodel-based Importance Sampling for Structural Reliability Analysis," *Probabilistic Engineering Mechanics*, Vol. 33, pp. 47~57.
- (10) Cadini, F., Santos, F. and Zio, E., 2014, "An Improved



- Adaptive Kriging-based Importance Technique for Sampling Multiple Failure Regions of Low Probability,” *Reliability Engineering and System Safety*, Vol. 131, pp. 109~117.
- (11) Zhao, H., Yue, Z., Liu, Y., Gao, Z. and Zhang, Y., 2015, “An Efficient Reliability Method Combining Adaptive Importance Sampling and Kriging Metamodel,” *Applied Mathematical Modelling*, Vol. 39, pp. 1853~1866.
- (12) Cho, T.-M., Ju, B.-H., Jung, D.-H. and Lee, B.-C., 2006, “Reliability Estimation Using Two-Stage Kriging Metamodel and Genetic Algorithm,” *Trans. Korean Soc. Mech. Eng. A*, Vol. 30, No. 9, pp. 1116~1123.
- (13) Meggiolaro, M. A. and Castro, J. T. P., 2004, “Statistical Evaluation of Strain-life Fatigue Crack Initiation Predictions,” *International Journal of Fatigue*, Vol. 26, pp. 463~476.