

〈 논문 〉

보간과 회귀를 위한 일반크리깅 모델

이 태 희[†] · 정 재 준^{*}

(2004년 8월 4일 접수, 2004년 12월 27일 심사완료)

Generalized Kriging Model for Interpolation and Regression

Tae Hee Lee and Jae Jun Jung

Key Words : Generalized Kriging Model(일반크리깅 모델), Interpolating Kriging Model(보간크리깅 모델), Maximum Likelihood Estimation(최우량추정법)

Abstract

Kriging model is widely used as design analysis and computer experiment (DACE) model in the field of engineering design to accomplish computationally feasible design optimization. In general, kriging model has been applied to many engineering applications as an interpolation model because it is usually constructed from deterministic simulation responses. However, when the responses include not only global nonlinearity but also numerical error, it is not suitable to use kriging model that can distort global behavior. In this research, generalized kriging model that can represent both interpolation and regression is proposed. The performances of generalized kriging model are compared with those of interpolating kriging model for numerical function with error of normal distribution type and trigonometric function type. As an application of the proposed approach, the response of a simple dynamic model with numerical integration error is predicted based on sampling data. It is verified that the generalized kriging model can predict a noisy response without distortion of its global behavior. In addition, the influences of maximum likelihood estimation to prediction performance are discussed for the dynamic model.

기호설명

- $\hat{Y}(\mathbf{x})$: 크리깅 모델
- σ_z^2 : 편차의 분산
- σ_e^2 : 랜덤오차의 분산
- θ : 상관계수 벡터
- α : 상관파라미터 벡터

1. 서론

크리깅 모델은 공간 분석 및 공간 자료의 연관 관계를 규명하기 위해 지질 통계학에서 처음으로 이용되었다.⁽¹⁾ 공학분야에서는 계산 비용이 많이 요구되는 시뮬레이션 모델을 대신하여 근사적절

계 과정에서 근사모델로 이용되거나 비선형적인 응답을 효과적으로 표현하기 위한 전산실험 모델 (design analysis and computer experiment model: DACE model)로서 폭넓게 활용되고 있다.

Sacks 는 VLSI 회로설계에 크리깅 모델을 처음으로 적용하였으며,⁽²⁾ Simpson 은 비행기 스파이크 노즐(aerospike nozzle)에 대한 시뮬레이션 응답을 크리깅 모델로 근사화하여 효과적인 다분야통합 최적설계를 수행하였다.⁽³⁾ Lee 등은 CRT 패널의 형상 최적설계와 연약 지반 위를 운행하는 심해저 집광 차량의 비선형적인 동적 응답을 예측하는데 크리깅 모델을 이용하였다.^(4,5) 이와 같이 전산실험 모델로서의 크리깅 모델은 같은 입력에 대해서 동일한 응답을 주는 결정적(deterministic)인 응답을 바탕으로 하기 때문에 보간 모델로서의 활용성이 중요하게 인식되어 왔다.

그러나 전산실험이라 하더라도 비선형성을 가지면서 수치오차가 포함된 시뮬레이션 응답에 대해서는 기존의 보간기능을 가진 크리깅모델을 적용하는 것은 실제함수의 특성을 왜곡시킬 위험이

[†] 책임저자, 회원, 한양대학교 기계공학부
E-mail : thlee@hanyang.ac.kr
TEL: (02)2290-0449, FAX (02)2298-4634

^{*} 한양대학교 대학원 기계설계학과

있다. 따라서 이러한 문제에서는 랜덤오차의 효과를 완화하면서 전역적인 비선형을 표현할 수 있는 근사화 기법을 적용하는 것이 타당하다. 이와 관련된 연구로, Sacks 는 측정오차가 존재할 경우 오차의 분산항을 고려함으로써 크리깅 모델을 확장할 수 있음을 언급하였다.⁽⁶⁾ Mitchell 과 Morris 는 크리깅 모델을 반응표면모델의 보간 특성을 지닌 모델로 보고, 랜덤응답에 대한 회귀모델로 사용될 수 있음을 지적하였다.⁽⁷⁾ Sasena 는 지질통계학에서 사용되는 너겟 파라미터(nugget parameter)를 이용하여 비보간(non-interpolation) 특성을 지닌 평활화 크리깅(smoothed Kriging)을 연구하였다.⁽⁸⁾

그러나 랜덤오차를 고려할 수 있는 크리깅 모델에 대해서 회귀특성에 영향을 주는 인자와 그에 측정능과의 관련성을 분석한 연구는 없었으며, 무엇보다도 랜덤오차를 고려할 수 있는 크리깅 모델의 확장성을 중요하게 인식하고 실제적인 공학 문제에 적용한 사례는 전무한 실정이다.

본 연구에서는 통계개념을 바탕으로 일반크리깅 모델을 해석적으로 유도하고, 이를 구현하기 위한 최우량 추정과정과 예측성능에 대해 논의한다. 본 연구에서 사용된 크리깅 모델은 보간법을 이용하는 보간크리깅 모델(interpolating Kriging model)과 구별하기 위해 특별히 일반크리깅 모델(generalized Kriging model)이라고 명명한다. 정규분포와 삼각함수 형태의 오차를 갖는 수확함수에 대해서 일반크리깅 모델을 적용하고 그 예측 성능을 보간크리깅 모델과 비교한다. 또한 수치오차가 존재하는 동역학 응답에 대해 일반크리깅 모델의 효용성을 살펴 본 후, 공학적 설계에 있어서 일반크리깅 모델의 활용 방안을 제안한다.

2. 크리깅 모델

모델링할 실제 함수는 평균에 해당하는 전역모델 $\mathbf{f}(\mathbf{x})^T \boldsymbol{\beta}$, 편차항(deviation) $z(\mathbf{x})$, 그리고 랜덤오차의 합으로 표현할 수 있다고 가정한다.

$$Y(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} + z(\mathbf{x}) + \varepsilon_r(\mathbf{x}) \quad (1)$$

여기서 $\mathbf{f}(\mathbf{x}) = \{f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_p(\mathbf{x})\}^T$ 는 입력변수 $\mathbf{x} \in R^n$ 로 정의된 벡터이고, $\boldsymbol{\beta} = \{\beta_1, \beta_2, \dots, \beta_p\}^T$ 는 미지의 계수벡터이다. $z(\mathbf{x})$ 와 $\varepsilon_r(\mathbf{x})$ 는 각각 $N(0, \sigma_z^2 \mathbf{V})$, $N(0, \sigma_\varepsilon^2)$ 의 서로 독립적인 정규분포를 갖는다고 가정한다.

n 개의 샘플링 점 $(\mathbf{x}^i, i=1, 2, \dots, n)$ 에서 구한 응답벡터 $\mathbf{Y} = \{Y(\mathbf{x}^1), Y(\mathbf{x}^2), \dots, Y(\mathbf{x}^n)\}^T$ 에 대해서 식 (1)을 적용하면 다음과 같은 식이 성립한다.

$$\mathbf{Y} = \mathbf{F}\boldsymbol{\beta} + \mathbf{Z} + \boldsymbol{\varepsilon}, \quad (2.1)$$

$$\mathbf{F} \equiv \begin{bmatrix} f_1(\mathbf{x}^1) & f_2(\mathbf{x}^1) & \dots & f_p(\mathbf{x}^1) \\ f_1(\mathbf{x}^2) & f_2(\mathbf{x}^2) & \dots & f_p(\mathbf{x}^2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(\mathbf{x}^n) & f_2(\mathbf{x}^n) & \dots & f_p(\mathbf{x}^n) \end{bmatrix} \quad (2.2)$$

$$\mathbf{Z} \equiv \begin{bmatrix} z(\mathbf{x}^1) \\ z(\mathbf{x}^2) \\ \vdots \\ z(\mathbf{x}^n) \end{bmatrix}, \quad \boldsymbol{\varepsilon}_r \equiv \begin{bmatrix} \varepsilon_r(\mathbf{x}^1) \\ \varepsilon_r(\mathbf{x}^2) \\ \vdots \\ \varepsilon_r(\mathbf{x}^n) \end{bmatrix} \quad (2.3)$$

편차와 랜덤오차가 독립적이라는 가정 하에서 두 응답 $Y(\mathbf{x}^i)$ 와 $Y(\mathbf{x}^j)$ 의 공분산행렬은 다음과 같이 표현할 수 있다.

$$\begin{aligned} \text{cov}[Y(\mathbf{x}^i), Y(\mathbf{x}^j)] \\ = \text{cov}[z(\mathbf{x}^i), z(\mathbf{x}^j)] + \text{cov}[\varepsilon_r(\mathbf{x}^i), \varepsilon_r(\mathbf{x}^j)] \end{aligned} \quad (3)$$

$z(\mathbf{x})$ 의 공분산은 다음과 같이 분산 σ_z^2 과 상관행렬(correlation matrix) $\mathbf{R} \in R^{n \times n}$ 의 곱으로 표현할 수 있다.⁽⁹⁾

$$\text{cov}[z(\mathbf{x}^i), z(\mathbf{x}^j)] \equiv \sigma_z^2 \mathbf{R} \quad (4)$$

상관행렬은 데이터들의 공간적인 관계(spatial correlation)를 표현하는 상관함수 R 로 정의된다.

$$\mathbf{R} = \begin{bmatrix} R(\mathbf{x}^1, \mathbf{x}^1, \boldsymbol{\theta}, \boldsymbol{\alpha}) & R(\mathbf{x}^1, \mathbf{x}^2, \boldsymbol{\theta}, \boldsymbol{\alpha}) & \dots & R(\mathbf{x}^1, \mathbf{x}^n, \boldsymbol{\theta}, \boldsymbol{\alpha}) \\ R(\mathbf{x}^2, \mathbf{x}^1, \boldsymbol{\theta}, \boldsymbol{\alpha}) & R(\mathbf{x}^2, \mathbf{x}^2, \boldsymbol{\theta}, \boldsymbol{\alpha}) & \dots & R(\mathbf{x}^2, \mathbf{x}^n, \boldsymbol{\theta}, \boldsymbol{\alpha}) \\ \vdots & \vdots & \ddots & \vdots \\ R(\mathbf{x}^n, \mathbf{x}^1, \boldsymbol{\theta}, \boldsymbol{\alpha}) & R(\mathbf{x}^n, \mathbf{x}^2, \boldsymbol{\theta}, \boldsymbol{\alpha}) & \dots & R(\mathbf{x}^n, \mathbf{x}^n, \boldsymbol{\theta}, \boldsymbol{\alpha}) \end{bmatrix} \quad (5)$$

여기서 상관함수는 두 샘플링 점 \mathbf{x}^i 와 \mathbf{x}^j 의 일차원 상관도의 곱(products of one-dimensional correlations)으로 나타낼 수 있는 상관관계로 한정한다. 일차원 상관도의 곱은 각 차원의 상관관계를 그 차원의 거리의 함수로만 표현하고 이들의 곱으로 상관관계를 표현한다.⁽⁷⁾

$$\begin{aligned} \mathbf{R} &= [R(\mathbf{x}^i, \mathbf{x}^j, \boldsymbol{\theta}, \boldsymbol{\alpha})] \\ &= \left[\exp \left(- \sum_{k=1}^{n_d} \theta_k (x_k^i - x_k^j)^{\alpha_k} \right) \right] \end{aligned} \quad (6)$$

여기서 $\boldsymbol{\theta}, \mathbf{x}, \boldsymbol{\alpha} \in R^{n_d}$ 이고, 상관행렬은 대각요소가 1 을 가지는 양정행렬 (positive definite matrix)이다. 상관계수 θ_k 는 x_k 방향에 대한 응답값들의 상관관계를 나타내는 파라미터로써 크리깅 모델의 x_k 방향 곡률을 결정한다. 즉, θ_k 가 크면 상관함수의 정의에 따라 상관도는 작고 이 입력변수 방향에 대한 크리깅 모델의 곡률은 비선형적 특성을 보인다. 반면 θ_k 값이 작으면 상관도는 커지며 이 입력변수 방향에 대해서 크리깅 모델의 예측값은 선형적인 특성을 보인다. 또한 상관지수 α_k 는 예측할 응답의 특성을 반영하는 파라미터로써 모델링할 응답의 연속성이 좋으면 $\alpha_k = 2$, 미분 가능하지

않거나 첨도(peakness)가 심한 함수는 $\alpha_k = 1$ 로 하는 것이 크리깅의 예측 성능을 좋게 하는 것으로 알려져 있다. 그러나 본 연구의 일반크리깅 모델에서는 이 값을 모수로 채택하고 있기 때문에 표현할 수 있는 함수 특성의 범위가 다양하게 되는 장점이 있다.

서로 다른 두 샘플링 점에서 랜덤오차가 상관관계가 존재하지 않는 상수라고 가정하면 다음과 같이 표현할 수 있다.

$$\text{cov}[\varepsilon_r(\mathbf{x}^i), \varepsilon_r(\mathbf{x}^j)] = \sigma_\varepsilon^2 \mathbf{I}, \quad \mathbf{I} \in R^{n \times n} \quad (7)$$

식 (4), (6), (7)을 식 (3)에 대입하면 아래와 같이 표현할 수 있다.

$$\begin{aligned} & \text{cov}[Y(\mathbf{x}^i), Y(\mathbf{x}^j)] \\ &= \sigma_\varepsilon^2 \left[\exp\left(-\sum_{k=1}^{n_k} \theta_k (x_k^i - x_k^j)^{\alpha_k}\right) + \gamma \right] \quad (8) \\ & \equiv \sigma_\varepsilon^2 \mathbf{V}, \quad \mathbf{V} \in R^{n \times n} \end{aligned}$$

여기서 $\gamma = \sigma_\varepsilon^2 / \sigma_x^2$ 는 응답에 대한 랜덤오차의 분산 비를 나타낸다. 일반적으로 전산실험에서는 시뮬레이션으로부터 얻어지는 응답값이 결정적이기 때문에 $\gamma = 0$ 이 되며, 오차가 존재하는 물리적인 실험에서는 $\gamma > 0$ 인 형태를 사용할 수 있다. 하지만 전산실험이라 하더라도 오차가 포함된 응답이나 수치계산 과정에서 수치오차가 존재하는 응답은 σ_ε^2 항을 포함한 형태가 사용되어야 한다. 이때 γ 의 크기에 따라 크리깅 모델의 회귀특성이 결정되는데 이 값이 커질수록 크리깅 모델이 평활화(smooth)되는 특징이 있다. 한편 평활화 크리깅을 연구한 Sasena 등은 지질통계학에서 사용되는 (1-nugget)에 해당하는 스칼라 인자가 $\sigma_\varepsilon^2 / (\sigma_\varepsilon^2 + \sigma_x^2)$ 와 일치한다는 것을 밝힌 바 있다.⁽⁸⁾

샘플링 점과 예측점에서의 응답값들의 상관관계는 상관벡터(correlation vector) \mathbf{v} 로 정의된다.

$$\begin{aligned} & \text{cov}[Y(\mathbf{x}^i), Y(\mathbf{x})] \\ &= \sigma_\varepsilon^2 [R(\mathbf{x}^1, \mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\alpha}) + \gamma, \dots, R(\mathbf{x}^n, \mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\alpha}) + \gamma]^T \quad (9) \\ & \equiv \sigma_\varepsilon^2 \mathbf{v}(\mathbf{x}) \end{aligned}$$

크리깅 모델은 식 (2.1)에서 정의된 응답값의 선형 합으로 표현되는 선형예측량(linear predictor)를 구하는 것으로 다음과 같이 표현된다.

$$\hat{Y}(\mathbf{x}) = \mathbf{c}(\mathbf{x})^T \mathbf{Y} \quad (10)$$

크리깅 모델은 선형 예측량 중에서 비편향성(unbiasedness)을 만족하면서 평균제곱오차(mean squared error: MSE)가 최소가 되는 최량비편향 예측량(best linear unbiased predictor: BLUP)을 구하는 것이다.⁽¹⁰⁾

따라서 크리깅 모델은 다음과 같은 최적화 문제로 정식화 할 수 있다.^(부록참조)

$$\begin{aligned} & \text{Minimize } \sigma_\varepsilon^2 \left([1 + \gamma + \mathbf{c}(\mathbf{x})^T \mathbf{V} \mathbf{c}(\mathbf{x}) - 2\mathbf{c}(\mathbf{x})^T \mathbf{v}(\mathbf{x})] \right) \quad (11) \\ & \text{subject to } \mathbf{F}^T \mathbf{c}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) \end{aligned}$$

카리쉬-쿤-터커 (Karush-Kuhn-Tucker: KKT) 필요조건을 구하기 위해 라그랑지 함수를 정의하면 다음과 같다.

$$\begin{aligned} & A(\mathbf{c}(\mathbf{x}), \boldsymbol{\lambda}) \\ &= \sigma_\varepsilon^2 [1 + \gamma + \mathbf{c}(\mathbf{x})^T \mathbf{V} \mathbf{c}(\mathbf{x}) - 2\mathbf{c}(\mathbf{x})^T \mathbf{v}(\mathbf{x})] \quad (12) \\ & \quad + \boldsymbol{\lambda}^T (\mathbf{F}^T \mathbf{c}(\mathbf{x}) - \mathbf{f}(\mathbf{x})) \end{aligned}$$

따라서 KKT 필요조건은 다음과 같다.

$$\begin{aligned} & \sigma_\varepsilon^2 [-2\mathbf{v}(\mathbf{x}) + 2\mathbf{V} \mathbf{c}(\mathbf{x})] + \mathbf{F} \boldsymbol{\lambda} = \mathbf{0} \quad (13) \\ & \mathbf{F}^T \mathbf{c}(\mathbf{x}) - \mathbf{f}(\mathbf{x}) = \mathbf{0} \end{aligned}$$

식 (13)을 행렬형태로 정리하면 다음과 같다.

$$\begin{bmatrix} \mathbf{0} & \mathbf{F}^T \\ \mathbf{F} & \mathbf{V} \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\lambda}} \\ \mathbf{c}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\mathbf{x}) \\ \mathbf{v}(\mathbf{x}) \end{bmatrix} \quad (14)$$

여기서 $\boldsymbol{\lambda} = 2\sigma_\varepsilon^2 \tilde{\boldsymbol{\lambda}}$ 이고, 분해행렬의 역행렬을 이용하면 $\mathbf{c}(\mathbf{x})$ 를 다음과 같이 구할 수 있다.

$$\mathbf{c}(\mathbf{x}) = \mathbf{V}^{-1} [\mathbf{v}(\mathbf{x}) - \mathbf{F}(\mathbf{F}^T \mathbf{V}^{-1} \mathbf{F})^{-1} (\mathbf{F}^T \mathbf{V}^{-1} \mathbf{v}(\mathbf{x}) - \mathbf{f}(\mathbf{x}))] \quad (15)$$

식 (15)를 식 (10)에 대입하여 정리하면 크리깅 모델을 다음과 같이 유도할 수 있다.

$$\hat{Y}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \hat{\boldsymbol{\beta}} + \mathbf{v}(\mathbf{x})^T \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{F} \hat{\boldsymbol{\beta}}) \quad (16)$$

여기서 $\hat{\boldsymbol{\beta}}$ 는 일반 최소제곱추정량(generalized least square estimator)으로 $\hat{\boldsymbol{\beta}} = (\mathbf{F}^T \mathbf{V}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{V}^{-1} \mathbf{Y}$ 로 표현된다.

식 (16)의 크리깅 모델은 식 (8)에서 정의된 모수 $\theta_k, \alpha_k, \gamma$ 와 편차의 분산 σ_ε^2 를 구해야만 구현이 가능하다. 편차의 분산은 일반최소제곱법을 이용하여 n 개의 데이터들로부터 추정된 분산으로 대체할 수 있다.

$$\hat{\sigma}_\varepsilon^2 = \frac{(\mathbf{Y} - \mathbf{F} \hat{\boldsymbol{\beta}})^T \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{F} \hat{\boldsymbol{\beta}})}{n} \quad (17)$$

식 (16)의 크리깅 모델에서 남은 미지수는 $\theta_k, \alpha_k, \gamma$ 이고, 이것은 최우량추정법(maximum likelihood estimation: MLE)에 의해서 구한다.⁽¹¹⁾

크리깅 모델의 모수를 결정하는 우량함수(likelihood function)는 다음과 같이 정의된다.

$$L = \frac{1}{(2\pi)^{n/2} \sqrt{|\hat{\sigma}_\varepsilon^2 \mathbf{V}|}} \exp\left(-\frac{(\mathbf{Y} - \mathbf{F} \hat{\boldsymbol{\beta}})^T \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{F} \hat{\boldsymbol{\beta}})}{2\hat{\sigma}_\varepsilon^2}\right) \quad (18)$$

우량함수를 최대로 만드는 $\theta_k, \alpha_k, \gamma$ 을 구한다는 것은 확률적으로 가능성이 높은 $\theta_k, \alpha_k, \gamma$ 값을 추정한다는 의미이다.

일반적으로 우량함수는 다중극부점을 갖고 해석적인 방법으로 구할 수 없는 것으로 알려져 있기 때문에 최우량추정 과정에서는 수치적인 최적화가 요구된다. 식 (18)의 우량함수에 로그를 취하고 상수항을 제외하고 정리하면 다음과 같다.

$$\text{maximize } -\left(\frac{n \ln \hat{\sigma}_z^2 + \ln |\mathbf{V}|}{2}\right) \quad (19)$$

그런데 최적화 과정에서 비정칙행렬(singular matrix)이 될 경우, $\ln |\mathbf{V}|$ 의 수치계산이 불가능해지기 때문에 식 (19)와 등가의 최우량 추정과정을 다음과 같이 표현할 수 있다.

$$\begin{aligned} &\text{minimize}_{\theta, \alpha, \gamma} \hat{\sigma}_z^2 |\mathbf{V}|^{1/n} \\ &\text{subject to } \theta_k > 0, \quad k=1,2,\dots,n_d \\ &\quad 0 < \alpha_k \leq 2, \quad k=1,2,\dots,n_d \\ &\quad 0 \leq \gamma \leq 1 \end{aligned} \quad (20)$$

3. 예 제

일반크리깅 모델의 유용성을 검증하기 위해 정규분포와 삼각함수 형태의 오차를 갖는 수학함수, 그리고 부싱 조인트(bushing joint)로 구성된 동역학 응답에 대해 크리깅 모델을 구성한다. 그리고 일반크리깅 모델을 생성하는 과정에서 구해진 최우량추정 결과를 보간크리깅 모델의 추정 결과와 비교하여 근사화에 미치는 영향을 살펴본다.

최우량추정 과정에서 보간크리깅 모델은 $\alpha_k = 2$, $\gamma = 0$ 으로 고정하고 θ_k 만을 최적화 변수로 하였으며, 일반크리깅 모델은 $\theta_k, \alpha_k, \gamma$ 모두를 최적화 변수로 설정하였다. 최우량추정 과정의 최적화 범위는 $\theta_1 \in [0.0001, 100]$, $\alpha_1 \in [0.0001, 2]$, $\gamma_1 \in [0.0001, 1]$, 로 설정하였다.

3.1 정규분포 형태의 오차를 갖는 수학함수
정규분포를 따르는 오차를 포함한 다음의 수학함수를 고려한다.

$$y(x) = 10 \sin(3\pi x + 10) + \varepsilon_r, \quad x \in [1, 10], \quad \varepsilon_r \sim N(0,3) \quad (21)$$

Fig. 1에서 근사화할 함수가 전역적인 비선형성과 함께 랜덤오차가 존재할 경우, 보간크리깅 모델이 실제함수를 근사화하는데 효과적이지 못함을 볼 수 있다. 반면, 일반크리깅 모델은 랜덤오차의 효과를 완화하면서도 전역최대값의 위치와 최소값의 위치를 정확하게 예측하고 있음을 알 수 있다.

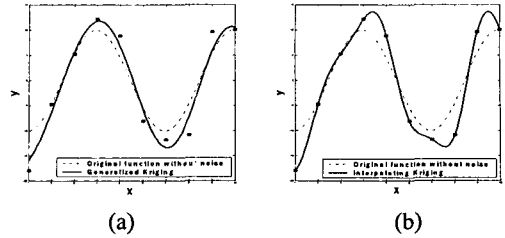


Fig. 1 Metamodels for function with noise of normal distribution type: (a) generalized Kriging model and (b) interpolating Kriging model

Table 1 Parameters obtained from MLE process

Estimation	Generalized Kriging	Interpolating Kriging
$\hat{\beta}$	-0.149	-0.092
$\hat{\sigma}_z^2$	153.697	115.979
$\hat{\theta}_1$	1.970	5.415
$\hat{\alpha}_1$	2.000	2.000
$\hat{\gamma}$	0.081	0.0
MLE	0.571	0.608

Table 1은 크리깅 모델을 생성하는 과정에서 최우량 추정법에 의해 추정된 모수들을 나타낸다. Table 1에서 두 가지의 크리깅 모델은 식 (21)의 평균값을 각각 $\hat{\beta} = -0.149, -0.092$ 으로, 분산을 $\hat{\sigma}_z^2 = 153.697, 115.979$ 로 추정하였다. 상관 파라미터 $\hat{\theta}_1$ 은 일반크리깅 모델보다 보간 크리깅이 조금 더 비선형적이라는 것을 나타낸다. 또한 일반크리깅 모델은 상관지수를 $\hat{\alpha}_1 = 2$ 로 추정하였으므로 공간적인 관계를 가우시안 상관함수로 예측하였음을 알 수 있다. $\hat{\gamma}$ 값은 $\hat{\sigma}_z^2$ 에 대한 상대적인 값이므로 문제에 따라 정량적인 값이 중요한 의미를 갖는다고는 볼 수 없지만, 랜덤오차의 분산이 비교적 크지 않기 때문에 회귀정도를 결정하는 $\hat{\gamma}$ 값이 상대적으로 작게 나왔다고 볼 수 있고, 그 결과는 Fig. 1에서도 확인할 수 있다.

3.2 삼각함수 형태의 오차를 갖는 수학함수
삼각함수 형태의 오차를 갖는 다음의 수학함수를 고려한다.

$$\begin{aligned} y(x) &= 10 \sin\left(\frac{\pi}{3} + 10\right) \sin\left(\frac{\pi}{3}x\right) - 10 \cos\left(\frac{\pi}{3}x\right) \\ &\quad + 5 \sin\left(\frac{\pi}{8}x\right) + \varepsilon_r, \quad (22) \\ x &\in [1, 9], \quad \varepsilon_r = 5 \cos(10\pi x) \end{aligned}$$

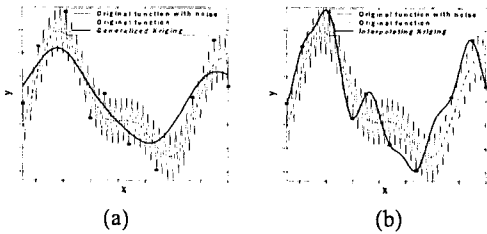


Fig. 2 Metamodels of function with noise of trigonometric type: (a) generalized Kriging model and (b) interpolating Kriging model

Table 2 Parameters obtained from MLE process

Estimation	Generalized Kriging	Interpolating Kriging
$\hat{\beta}$	-0.126	-0.073
$\hat{\sigma}_z^2$	99.751	102.012
$\hat{\theta}_1$	5.485	18.434
$\hat{\alpha}_1$	2.00	2.000
$\hat{\gamma}$	0.418	0.0
MLE	0.841	0.893

Fig. 2 에서 일반크리깅 모델의 예측 성능은 변곡점과 극부 최소값에 대해서는 정확성이 다소 떨어지는 경향을 보이지만 전역적인 함수의 경향성과 전역최대값의 위치를 효과적으로 근사화하고 있음을 볼 수 있다. 이에 반해 보간크리깅 모델은 오차의 영향으로 전역적인 비선형성을 효과적으로 표현하지 못하고 있을 뿐 아니라 극부점들을 갖는 형태로 근사화가 되어 보간특성이 근사화에 부정적인 영향을 끼치고 있음을 알 수 있다.

식 (22)의 함수는 랜덤오차의 진폭이 다소 크기 때문에 Table 1 의 결과와 비교할 때 응답에 대한 랜덤오차의 분산의 비가 $\hat{\gamma} = 0.418$ 로 다소 크게 나온 것을 볼 수 있다. 또한 보간크리깅 모델보다 일반크리깅 모델의 상관계수가 작다는 것은 일반크리깅 모델의 평활성(smoothness)이 보간크리깅 모델보다 더 크다는 것을 나타낸다. 또한 Table 1 에서와 마찬가지로 Table 2 에서 보간크리깅 모델의 우량함수값이 일반 크리깅 모델보다 작게 나왔다. 이것은 보간크리깅 모델보다 일반크리깅 모델이 랜덤오차가 존재할 때 데이터의 변동을 예측하는데 있어 가능성(likelihood)이 더 높은 모델이라는 것을 나타낸다.

3.3 부싱을 갖는 강체 동역학 모델의 응답

부싱은 동역학 모델에서 물체(body)들의 연결부에 충격을 완화하거나 강체 조인트를 대신해 조인트의 간격(clearance)효과를 주기 위해서 사용된다. 일반적으로 부싱은 스프링 요소로 모델링되며

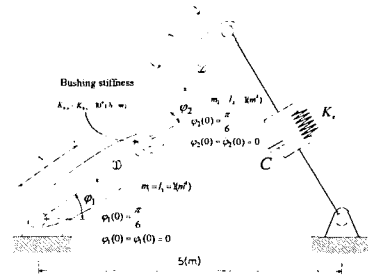


Fig. 3 Two-pendulum model jointed with bushing element

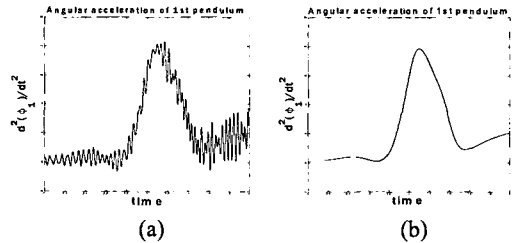


Fig. 4 Angular acceleration response of pendulum ① connected with (a) bushing and (b) rigid joint

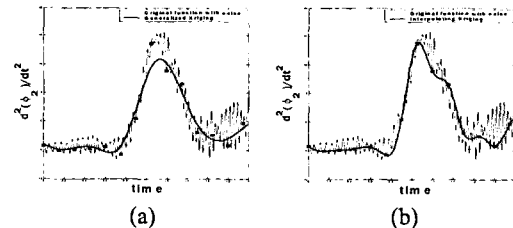


Fig. 5 Metamodels for angular acceleration with numerical errors of pendulum ①: (a) generalized Kriging model and (b) interpolating Kriging model

스프링의 강성을 크게 할수록 강체 조인트와 유사한 응답을 얻을 수 있다. 하지만 스프링 강성이 클수록 시간에 대한 적분과정에서 수치적인 오차가 포함된 응답을 보이는 특징이 있다.

Fig. 3 은 부싱조인트로 연결된 진자(pendulum) 모델을 나타내고 있다.

Fig. 4(a)는 스프링과 댐핑 상수가 각각 $K_s = 1(N/m), C = 1(N \cdot s/m)$ 일 때 시간에 따른 진자 ①의 각가속도를 나타낸 것이고 Fig. 4(b)는 오차가 제거된 실제 각가속도를 도시한 것이다. Fig. 4(a)에서 진자 ①의 각가속도는 비선형적인 응답과 함께 수치적인 적분오차를 포함하고 있다.

Fig. 5 는 진자 ①의 각가속도에 대한 일반크리깅 모델과 보간크리깅 모델을 도시한 것이다. 일반크리깅 모델은 Fig. 4 에 도시한 각가속도의 경향성을 매우 정확하게 근사화하고 있음을 확인할 수 있다.

Table 3 Parameters obtained from MLE process for angular acceleration of pendulum ①

Estimation	Generalized Kriging	Interpolating Kriging
$\hat{\beta}$	-0.008	0.002
$\hat{\sigma}_z^2$	28.996	28.836
$\hat{\theta}_1$	3.894	12.340
$\hat{\alpha}_1$	2.000	2.000
$\hat{\gamma}$	0.167	0.0
MLE	0.493	0.573

Table 4 Parameters obtained from MLE process for angular acceleration of pendulum ②

Estimation	Generalized Kriging	Interpolating Kriging
$\hat{\beta}$	0.028	-0.007
$\hat{\sigma}_z^2$	83.737	92.113
$\hat{\theta}_1$	2.698	15.150
$\hat{\alpha}_1$	2.000	2.000
$\hat{\gamma}$	0.270	0.0
MLE	0.520	0.650

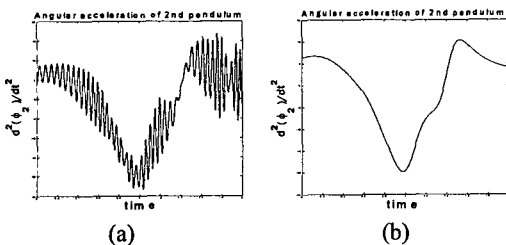


Fig. 6 Angular acceleration response of pendulum ② connected with (a) bushing and (b) rigid joint

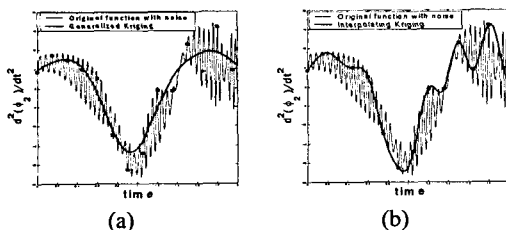


Fig. 7 Metamodels for angular acceleration with numerical errors of pendulum ②: (a) generalized Kriging model and (b) interpolating Kriging model

Table 3 에서는 분산 $\hat{\sigma}_z^2$ 가 비슷하게 나온 것이 특징적이라고 할 수 있다. 일반크리깅 모델과 보간크리깅 모델의 상관계수 θ_1 은 각각 $\theta_1 = 3.894, 12.340$ 으로 추정되었다. 또한 랜덤오차의 영향이 크지 않아 $\hat{\gamma} = 0.167$ 로 추정되었으며 일반크리깅 모델에서 상관함수의 형태는 가우시안으로 추정되었다.

Fig. 6 은 각각 부싱과 강체조인트를 가지는 진자모델에서 진자 ②의 시간에 따른 각가속도 그래프를 도시한 것이다.

Fig. 7 에서 보간크리깅 모델이 수치오차의 영향으로 실제 각가속도의 경향성을 왜곡하거나 국부적인 응답을 가진 형태로 근사화된 반면 일반크리깅 모델은 실제 각가속도의 경향성을 정확하게 근사화하고 있음을 확인할 수 있다.

Table 4 에서 보간크리깅 모델의 상관계수는 $\theta_1 = 15.150$ 으로 추정되었고, 이것은 보간크리깅 모델의 응답 특성이 강한 비선형성을 띄고 있음을 나타낸다. 반면, 일반크리깅 모델은 상대적으로 작은 값인 $\theta_1 = 2.698$ 으로 추정되었다. $\hat{\gamma}$ 는 0.27 로 추정되어 일반크리깅 모델이 보간크리깅 모델에 비해 매끈한 형태로 근사화된 것을 알 수 있다.

4. 결론

응답 특성에 따라서 회귀나 보간을 할 수 있는 일반크리깅 모델을 통계적 개념에 기초하여 제시하였다. 수치 오차를 포함한 응답에 대해서 일반크리깅 모델의 최우량 추정 결과를 보간크리깅 모델의 결과와 비교하여 근사화 결과에 미치는 영향을 파악하였다. 그 결과, θ_k 가 작아지고 γ_k 가 커질수록 크리깅 모델의 곡률이 평활화되는 것을 알 수 있었다. 보간크리깅 모델의 곡률을 결정하는 유일한 파라미터가 θ_k 인 반면, 일반크리깅 모델은 θ_k 와 γ_k 가 동시에 영향을 미치고 있음을 알 수 있었다. 그러나 α_k 는 항상 2 로 추정되어 α_k 의 영향은 판단할 수가 없었다.

그리고 일반크리깅 모델의 예측성능을 살펴본 결과, 일반크리깅 모델이 보간크리깅 모델보다 수치오차의 영향에 강건하면서도 전역적인 비선형성을 효과적으로 표현할 수 있는 방법이라는 결론을 얻었다. 따라서 수치적 노이즈를 포함한 전산실험에서는 일반크리깅 모델이 효과적인 근사화 기법으로 폭넓게 활용될 수 있을 것으로 판단된다. 또한 랜덤오차가 포함된 비선형적인 응답을 보이는 물리적 실험에서는 일반크리깅 모델이 반응표면모델의 한계를 극복할 수 있는 대안으로써 제안될 수 있을 것이다.

본 연구에서는 수치오차를 효과적으로 완화시킬 수 있는 적절한 샘플링 개수에 대한 연구가 포함되지 못했다. 향후 샘플링 방법 및 샘플링 개수에 따른 일반크리깅 모델의 예측성능에 관한 연구가 수행되어야 할 것이다.

후 기

이 연구는 한국과학재단 지정 최적설계신기술 연구센터의 지원과 해양수산 특정 연구과제인 ‘심해저 집광시스템 및 채광운용기술 개발’ 과제의 일부분으로 수행되었습니다.

참고문헌

(1) Matheron, G., 1963, "Principles of Geostatistics," *Economic Geology*, Vol. 58, pp. 1246~1266.

(2) Sacks, J., Welch, W.J., Mitchell, T.J. and Wynn, H.P., 1989, "Design and Analysis of Computer Experiments," *Statistical Science*, Vol. 4, No. 4, pp. 409~435.

(3) Simpson, T.W., Mauery, T.M., Korte, J.J. and Mistree, F., 2001, "Kriging Models for Global Approximation in Simulation-Based Multidisciplinary Design Optimization," *AIAA Journal*, Vol. 39, No. 12, pp. 2234~2241.

(4) Lee, T.H., Lee, C.J. and Lee, K.K., 2003, "Shape Optimization of a CRT based on Response Surface and Kriging Metamodels," *Trans. of KSME (A)*, Vol. 27, No. 30, pp. 381~386.

(5) Lee, T.H., Lee, C.S., Jung, J.J., Kim, H.W., Hong, S., and Choi, J.S., 2003, "Prediction of the Motion of Tracked Vehicle on Soft Soil Using Kriging Metamodel," *Proceedings of The Fifth Ocean Mining Symposium*, ISOPE, pp. 144~149.

(6) Sacks, J., Schiller, S.B. and Welch, W.J., 1989, "Designs for Computer Experiment," *Technometrics*, Vol. 31, No.1, pp. 41~47.

(7) Mitchell, T.J. and Morris, M.D., 1992, "The Spatial Correlation Function Approach to Response Surface Estimation," *Proceedings of the 1992 Winter Simulation Conference*, pp. 565~571.

(8) Sasena, M., Parkinson, M., Goovaerts, P., Papalambros and P., Reed, M., 2002, "Adaptive Experimental Design Applied to An Ergonomics Testing Procedure," *Proceedings of DETC'02 ASME 2002 Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Montreal.

(9) Koehler, J.R. and Owen, A.B., 1996, "Compter Experiments," in Ghosh, S. and Rao, C.R., eds, *Handbook of Statistics*, 13, pp. 261~308, Elsevier Science, New York.

(10) Štulajter, F., 1997, "Predictions in Nonlinear Regression Models," *Acta Mathematica Universitatis Comenianae*, Vol. LXVI, No.1, pp. 71~81.

(11) Cox, D.D., Park, J.J. and Singer, C.E., 2001, "A Statistical Method for Tuning a Computer Code to a Data Base," *Computational Statistics & Data Analysis*, Vol. 37, pp. 77~92.

부 록

크리깅 모델의 평균제곱오차는 다음과 같이 표현된다.

$$\begin{aligned} MSE[\hat{Y}(\mathbf{x})] &= E[(\hat{Y}(\mathbf{x}) - Y(\mathbf{x}))^2] \\ &= \text{var}[Y(\mathbf{x})] + \text{var}[\hat{Y}(\mathbf{x})] \\ &\quad - 2 \text{cov}[\hat{Y}(\mathbf{x}), Y(\mathbf{x})] + (E[\hat{Y}(\mathbf{x})] - E[Y(\mathbf{x})])^2 \end{aligned} \tag{A.1}$$

이때 추정량의 기대값이 모수(parameter)와 같은 성질인 비편향조건은 다음과 같다.

$$\begin{aligned} E[\hat{Y}(\mathbf{x})] &= E[Y(\mathbf{x})] \\ \Leftrightarrow \mathbf{c}(\mathbf{x})^T \mathbf{F} \boldsymbol{\beta} &= \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} \quad \forall \boldsymbol{\beta} \\ \Leftrightarrow \mathbf{F}^T \mathbf{c}(\mathbf{x}) &= \mathbf{f}(\mathbf{x}) \end{aligned} \tag{A.2}$$

식 (A.1)에서 식 (A.2)을 이용하면 평균제곱오차는 다음과 같다.

$$\begin{aligned} MSE[\hat{Y}(\mathbf{x})] &= \text{var}[Y(\mathbf{x})] + \text{var}[\hat{Y}(\mathbf{x})] - 2 \text{cov}[\hat{Y}(\mathbf{x}), Y(\mathbf{x})] \end{aligned} \tag{A.3}$$

식 (A.3)은 평균제곱오차가 응답함수의 분산, 예측량의 분산 그리고 예측량과 응답함수의 공분산으로 표현된다는 것을 나타낸다. 식 (A.3)의 각 항들은 식 (2.1)~(2.3), 식 (8)~(10)을 이용하면 다음과 같이 표현할 수 있다.

$$\begin{aligned} \text{var}[Y(\mathbf{x})] &= E[(z(\mathbf{x}) + \varepsilon_r(\mathbf{x}))^2] \\ &= \sigma_z^2 + \sigma_\varepsilon^2 \end{aligned} \tag{A.4}$$

$$\begin{aligned} \text{var}[\hat{Y}(\mathbf{x})] &= E[(\hat{Y}(\mathbf{x}) - E[\hat{Y}(\mathbf{x})])^2] \\ &= \sigma_z^2 \mathbf{c}(\mathbf{x})^T \mathbf{V} \mathbf{c}(\mathbf{x}) \end{aligned} \tag{A.5}$$

$$\begin{aligned} \text{cov}[\hat{Y}(\mathbf{x}), Y(\mathbf{x})] &= E[(\hat{Y}(\mathbf{x}) - E[\hat{Y}(\mathbf{x})])(Y(\mathbf{x}) - E[Y(\mathbf{x})])] \\ &= \sigma_z^2 \mathbf{c}(\mathbf{x})^T \mathbf{v} \end{aligned} \tag{A.6}$$

식 (A.4)~(A.6)의 결과를 이용하여 평균제곱오차를 표현하면 아래와 같다.

$$\begin{aligned} MSE[\hat{Y}(\mathbf{x})] &= \sigma_z^2 (1 + \gamma + \mathbf{c}(\mathbf{x})^T \mathbf{V} \mathbf{c}(\mathbf{x}) - 2 \mathbf{c}(\mathbf{x})^T \mathbf{v}(\mathbf{x})) \end{aligned} \tag{A.7}$$