Journal for History of Mathematics
Vol. 35 No. 1 (Feb. 2022), 1–18
http://dx.doi.org/10.14477/jhm.2022.35.1.001

Gaussian Process Regression and Its Application to Mathematical Finance

가우시언 과정의 회귀분석과 금융수학의 응용

Lim Hyuncheul 임현철

This paper presents a statistical machine learning method that generates the implied volatility surface under the rareness of the market data. We apply the practitioner's Black-Scholes model and Gaussian process regression method to construct a Bayesian inference system with observed volatilities as a prior information and estimate the posterior distribution of the unobserved volatilities. The variance instead of the volatility is the target of the estimation, and the radial basis function is applied to the mean and kernel function of the Gaussian process regression. We present two types of Gaussian process regression methods and empirically analyze them.

Keywords: practitioner's Black-Scholes model, Gaussian process regression, Bayesian, volatility, radial basis function; 실무자의 블랙-숄즈 모델, 베이지언, 가우시언 프로세스 회귀, 방사형 기저함수, 내재변동성.

MSC: 60G15. 62C10, 91G80

1 서론

실무자의 블랙-숄즈 모델 고전적인 블랙-숄즈 모델 [3]은 고정된 변동성을 가정한다. 그러나 실무적으로는 만기 및 행사가격에 따라 다른 내재 변동성이 시장에서 관측되면서 만기와 행사가격에 따라 다른 변동성을 적용한다. 이렇게 조정된 방법

Lim supported by NRF-2019R1I1A3A03059382, and BK21 FOUR (Fostering Outstanding Universities for Research, NO.5120200913674) funded by the Ministry of Education(MOE, Korea) and NRF, 이 논문은 전남대학교 학술연구비(과제번호: 2021-2529) 지원에 의하여 연구되었음.

Lim Hyuncheul: Dept. of Math. Chonnam National Univ. E-mail: limhc@jnu.ac.kr
Received on Jan. 21, 2022, revised on Feb. 6, 2022, accepted on Feb. 13, 2022.

을 실무자의 Black–Scholes(PBS) [4]방법이라 한다. PBS는 거래소에서 바닐라 옵션이 평가되는 표준적인 방법이다.

$$\frac{dS(t)}{S(t)} = (r-q)dt + \sigma_{K,T}dW(t), \tag{1}$$

여기서 $W(t)$는 위너-과정, $\sigma_{K,T}$는 만기시점 T와 행사가격 K인 바닐라 옵션의 변동성이 다름에 따라 이를 구분하여 표기하였다. 여기서 r, q는 만기 T까지 적용되는 T에 따라 다른 값을 갖는 이자율 및 배당율이다[1].

미래 시점의 주가를 만기와 행사가격으로 구성한 2차원 공간에 표시할 수 있다. 이를 상태공간 $\mathcal{S} = \{(K,T) \mid K > 0, T > 0\}$라 한다. 각 점 (K,T)의 값을 행사가격 K와 만기 T로 갖는 바닐라 옵션에 대응시킬수 있다. PBS방법은 서로 구별되는 블랙-숄즈 모델들로 이루어진 $\mathcal{S}$위의 확률 미분방정식계로 간주할 수 있는데, 각 점에 대응되는 블랙-숄즈 모델이 서로 다르다는 가정을 내포하고있다.

실무자의 블랙-숄즈 모델계 블랙의 선물에 대한 옵션가격 모형 [2]의 전개와 동일하게, 만기 T인 선물가격 $F(T)$가 기초자산인 선물옵션에 대해서도 PBS방법을 적용시킬 수 있다. 실무적인 근거는 통화옵션시장에서 통화선도가격에 대한 콜,풋 옵션이 거래되는 방식에서 찾을 수 있다. 선물가격의 프로세스를 $F(t,T)$라 하자. PBS방법에 의한 선물가격의 동역학과 로그 선물가격의 분포는 다음 식과 같다.[2]

$$\frac{dF(t,T)}{F(t,T)} = \sigma_{K,T}dW(t), \tag{2}$$

$$\ln\frac{F(T,T)}{F(0,T)} \sim \mathcal{N}(-\frac{1}{2}\sigma_{K,T}^2 T, \sigma_{K,T}^2 T) \tag{3}$$

이 식에서 선물가격 프로세스 $F(t,T)$는 행사가격 K, 만기 T에 따라 다른 변동성 $\sigma_{K,T}$을 적용하는 선물에 대한 옵션의 기초자산이 갖는 동역학이다. 따라서 매개변수 (K,T)가 다르면 위험의 원천인 위너-측도 $W(t)$도 다르다. 식 (3)에서 $F(t,T)$는 행사가격 K와 무관하게 시장에서 관찰된 만기 T인 선물가격 $F(T) = F(0,T)$이 기댓값임을 의미한다[3]. 그러나 식 (2)가 내재한 분포함수는 T는 물론 행사가격 K마다 다르다. 그 결과 선도 위험 중립 조건은 성립하지 않는다. 동일한 만기의 서로 다른 행사가를 갖는 옵션들에 대하여 선도 위험 중립 측도 $W^T(t)$는 일치하여야 하지만 행사가 K에 따라 로그(단위)선물가격의 분포가 다르다. 따라서 PBS를 적용한다는 것은 "만기와

1) $r = -\frac{1}{T}\ln P(T), q = r - \frac{1}{T}\ln\frac{F(T)}{S(0)}$, $P(T), F(T)$는 만기 T인 단위 할인채권 가격 및 주식의 선도가격

2) 측도론의 관점에서 선물과 선도가격은 구별을 하여야 하지만 [2]에서 선물이라 표기하였다.

3) 정규분포에 관한 적률생성함수 결과를 적용하면, $X = \ln\frac{F(T,T)}{F(0,T)} \Rightarrow \mathbb{E}[e^X] = e^{\mathbb{E}[X]+\frac{1}{2}\mathbb{V}[X]} = 1$

행사가격마다 다른 블랙-숄즈 모델을 사용함" 으로 해석한다.[4)]

K, T의 식인 (3)을 PBS시스템이라 지칭하자. 이 시스템에서 로그선물가격은 정규분포를 따르는 확률변수이면서 $\mathcal{S}$에서의 위치 (K, T)를 매개변수로 하는 함수가 된다. 이는 상태공간 $\mathcal{S}$의 모든 위치를 변수집합으로 갖는 다변량 정규분포로 확장할 수 있다. 그러나 PBS시스템을 다변량 정규분포로 가정할 경우 두 가지 문제가 발생한다.

- $\mathcal{S}$의 위치에 기반한 분포들 사이의 공분산 구조가 필요함 : PBS가 $\mathcal{S}$의 각 점에서 개별적으로 평가하는 방식을 의미하기 때문에 서로 다른 (K, T)에 기초한 로그선도가격 분포들 사이의 공분산은 주어지지 않는다.
- 일부 시장을 제외한 주가지수 파생상품시장의 유동성의 문제 : 우리나라 시장은 최근월물과 차근월물을 제외한 영역에서 변동성 데이터가 관측이 되지 않거나 매우 드물게 관측된다.

통계학적 머신러닝 기법중의 하나인 가우시언 프로세스 회귀(GPR)모형 [9]은 PBS시스템과 같이 매개변수로 연결된 대상들이 다변량 정규분포를 따름을 가정한다. 관측치를 조건부로 비관측치에 대한 베이지언 추론을 적용하는 회귀분석 방법이다. 확률분포의 특징을 갖는 GPR의 관측 대상들은 매개변수를 갖는 함수(GPR함수)이다. PBS시스템의 로그선물가격(3)은 행사가격과 만기시점을 매개변수로 갖는 GPR함수이다. 지수시장의 희박한 변동성 문제를 해결하기 위하여 이 논문에서는 PBS시스템의 가정에 더하여 통계학적 머신러닝 기법중의 하나인 가우시언 프로세스 회귀(GPR)모형 [9]을 적용하는 새로운 방법론을 제시한다.

이 논문의 기여 이 논문은 지수시장의 희박한 변동성 데이터 문제를 해결하는 설명가능한 머신러닝 방법을 제시한다. 저자의 기존 논문 [6]에서 방사기저함수(Radial Basis Function: RBF)에 의한 함수적 곡면 근사와 교차검증법(Leave One Out Cross Validation, LOOCV)을 적용하여 분산곡면을 구성하고 여러 RBF함수에 의한 곡면추정 결과를 보였다. RBF함수들중 박판스플라인함수(Thin Plate Spline: TPS)를 적용한 분산곡면의 근사는 매우 실용적이었지만 비관측 영역의 추정이 RBF근사의 특징에 의존하는 단점이 있었다. 그리고 TPS가 아닌 RBF함수를 이용한 곡면 추정은 적합한 형상모수(shape parameter)를 기존의 방식으로 찾기가 힘들었다. 논문은 실무적 방법인 PBS와 머신러닝방법인 GPR의 유사함을 고찰하였다. 또한 제시하는

4) 즉 PBS방법의 모든 블랙-숄즈 모델들은 K, T에 따라 다른 측도를 갖는 위너-과정이 적용된다.

문제에 대하여 GPR 방법과 RBF근사는 본질적으로 동일함을 발견하였다. PBS시스템은 블랙-숄즈 모델을 따르는 로그선물가격들의 모임이면서 가격들이 다변량 정규분포를 따름을 자연스럽게 가정할 수 있다. 이를 바탕으로 하는 첫번째 GPR접근법을 제시한다. PBS방법을 GPR의 틀안에서 설명하는 PBS시스템으로 재구성하였다.

두 번째 접근 방법은 GPR이 현실 문제에 적용되는 최근의 머신러닝에서 사용되는 방식에 충실한 접근법이다. 식 (3)이 의미하는 선도가격들의 분포에 대한 가정을 피하고 블랙-숄즈 모델의 분산들을 직접 관측대상으로 한다. 여기서 분산이란 변동성의 제곱과 파생상품의 만기를 곱한 값인 식의 (3) $\sigma^2_{K,T}T$이다. 이들이 갖는 공분산 구조(커널)에 파라메터를 갖는 RBF함수를 설정하는 방법이다. 커널함수의 최적화를 통하여 관측된 변동성에 대한 적합도를 높인 설명력을 얻고 이를 바탕으로 비관측 변동성을 추정한다. 두 방식 모두 두 점 사이의 거리를 변수로 갖는 방사기저함수(Radial Basis Function: RBF)를 사용한다. 이럴 경우 RBF함수의 뛰어난 산포된 데이터 적합 능력에 의하여 GPR함수의 평균값인 PBS분산이 효과적으로 근사된다. 여기에 더하여 다변량 정규분포의 해석적 이론이 기본이 되는 GPR의 베이지언 추정은 관측된 변동성이 희박한 경우에도 매우 효과적으로 관측되지 않은 변동성의 추정을 가능하게 한다.

논문의 구성 (1)은 PBS를 설명하고 이를 확장한 PBS시스템을 구성한다. PBS시스템에 GPR방법을 도입하는 이유에 대하여 설명한다. (2)는 GPR의 이론적 기반인 다변량 정규분포의 결합분포와 이를 응용한 조건부 분포들을 기술한다. 그리고 이들 분포에 의하여 유도되는 베이지안 추론 방법을 설명한다. PBS시스템을 GPR추정문제로 생각할 경우 PBS분산과 GPR의 평균함수 그리고 PBS시스템의 공분산과 GPR 커널이 대응됨을 보인다(14). 이를 이용하여 관측되지 않은 PBS분산의 추정문제를 GPR 방법을 적용하여 계산하는 과정을 기술한다. 두번째 방식의 설명에서는 GPR의 설명력을 높이는 RBF커널함수의 하이퍼-파라메타 최적화를 설명한다. 커널의 하이퍼-파라메타를 관측된 변동성을 기준으로 적합시키는 알고리즘을 기술한다. 대표적인 RBF커널함수인 제곱지수함수(가우시언)(Squared Exponential)와 마턴커널(Matérn 3/2 Kernel)(20)을 이용하여 분산곡면을 추정하는 식을 기술한다. (5)는 구현된 결과를 적용한 실제 예제를 보인다. 방법 1,2의 차이점을 분석한다.

2 가우시언 확률과정 회귀

GP와 GPR 가우시언 프로세스 (Gaussian Process: GP)는 랜덤과정(Random Process)의 한 종류이다. 일반적인 랜덤과정의 의미는 시간의 진행에 따라 나열된 확률변수 $X_t, t \in \mathbb{R}_+^1$의 모임을 의미한다. 가우시언은 각 확률변수가 정규분포를 따름을 의미한다. 블랙-숄즈 모델의 위너-과정 W_t을 예로 들면, 시간에 따른 매개화에 의하여 서로 다른 시점의 확률변수 $W_s, W_t, 0 \leq s < t$들이 독립증분과 증분의 정규분포를 갖는 강한 조건이 부여되는 주가의 경로가 생성된다. 그러나 GP에서 확률변수들 사이의 매개변수는 꼭 시간일 필요가 없다. 프로세스의 실현(realization)인 샘플이 관측된 데이터 사이트(data site, 위치)의 의미를 갖는다. 대신 위너-과정과 달리 확률변수들간의 공분산구조인 데이터 사이트들간의 이항함수를 명시적으로 부여하고 이를 커널이라 한다. 일반적으로 커널은 데이터 사이트들 사이의 거리에만 의존하는 방사형 기저함수로 구성한다.

GP는 프로세스이지만 매개변수를 회귀모델의 데이터 사이트(변수)로 갖는 함수로 생각할 수 있다. 매개변수가 연속된 값이면 무한차원(infinite dimensional) 다변량 정규분포를 따르게 된다. 매개변수를 $\boldsymbol{X}$라 두고 GPR함수를 $g(\boldsymbol{X})$로 표기하자. 두 점 $x, x' \in \boldsymbol{X}$에 대하여 정의되는 $g(\boldsymbol{X})$의 공분산 구조를 커널함수라 한다. 이를 $k(x, x')$으로 표기하자. $g(\boldsymbol{X})$는 아래와 같은 분포를 따른다.

$$g(\boldsymbol{X}) \sim \mathcal{GP}(m(\boldsymbol{X}), k(\boldsymbol{X}, \boldsymbol{X})) \tag{4}$$

여기서 $\mathcal{GP}$는 다변량 정규분포를 의미한다. 데이터 사이트 $\boldsymbol{X}$에 대하여 $m(\boldsymbol{X})$는 $g(\boldsymbol{X})$의 평균함수, 커널이라 불리우는 $k(\boldsymbol{X}, \boldsymbol{X})$는 공분산 행렬을 만드는 함수 이다. 여기서 평균 m 및 커널 k도 모두 매개변수 $\boldsymbol{X}$의 함수 또는 함숫값임을 주의하여야 한다. m을 GPR평균함수, 커널 k를 GPR커널함수라 지칭하자. 만일 $\boldsymbol{X}$의 크기가 N이면 $m(\boldsymbol{X})$는 N차원 벡터, $k(\boldsymbol{X}, \boldsymbol{X})$는 크기 $N \times N$인 공분산 행렬이다.

실제 관측에는 노이즈가 포함되는 경우가 많기 때문에 위치 $\boldsymbol{X}$에서 확률변수 $g(\boldsymbol{X})$가 관측된 값을 $\boldsymbol{y}$라 두고 이 값이 노이즈 $\sigma_n, (n : \text{noise})$를 포함하는 선형회귀모델을 추가한다. $g(\boldsymbol{X})$의 관측값 $\boldsymbol{y}$의 분포는 다음과 같다.

$$\begin{aligned} \boldsymbol{y} &= g(\boldsymbol{X}) + \epsilon,\ \epsilon = (\epsilon_1, \cdots, \epsilon_N)^\top,\ \epsilon_i \sim \mathcal{N}(0, \sigma_n^2) \\ \boldsymbol{y} &\sim \mathcal{GP}(m(\boldsymbol{X}), k(\boldsymbol{X}, \boldsymbol{X}) + \sigma_n^2 \boldsymbol{I}) \end{aligned} \tag{5}$$

회귀모형으로 확장한 GP를 가우시언 프로세스 회귀(Gaussian Process Regression: GPR) 이라 한다. GPR함수 $g(\boldsymbol{X})$는 평균함수 $m(\boldsymbol{X})$, 커널함수 $k(\boldsymbol{X}, \boldsymbol{X})$를 갖는 가우시언 분포이므로 GP는 함수들의 분포라고 할 수 있다. GPR은 GPR함수를

$\boldsymbol{X}$ 위의 관측데이터를 적합시키는 회귀분석을 추가한 정의이다. GPR은 러시아의 수학자 안드레이 콜모고로프(1903 $\sim$ 1987)에 의하여 소개된 가우시언 프로세스를 기원으로 두고 있지만 시간에 따른 매개변수를 일반적인 변수(위치, 만기)로 확장시킨 비선형 회귀분석이다. 다변량 정규분포를 따름을 가정하기 때문에 공분산에 해당되는 커널은 매개변수와 측정 값의 관계를 최대한 단순화 시킨다. 이때 대표적인 커널함수로 위치(매개변수)의 거리에만 의존하는 방사기저함수가 매우 자연스럽게 도입되었다. 한편 방사기저함수는 지리적 정보를 다루는 산림 [7], 광산, 지질에 종사하는 학자들에 의하여 1900년대 중반부터 실무적 필요에 의하여 연구되었다. 이후 수학적으로는 메쉬프리방법 [5] 그리고 통계학에서 누락데이터(Missing Data Problem)를 추정하는 한 방법 [10]으로 적용이 되었으며 현재 컴퓨터 과학자들에 의하여 베이지언 방법론이 머신러닝에 적용되면서 많이 알려졌다 [9].

GPR의 베이지안 추론구조 GPR은 베이지안 추론을 이용하여 누락(비관측)영역의 함수를 추정한다. 과정을 이해하기 위하여 다변량 정규분포 이론에서 유도되는 다음 정리들을 기술한다.

정리 2.1 (다변량 결합 정규 분포의 조건부 분포 [1, (2.4), (2.5)]): 두 정규확률벡터 $\boldsymbol{X}_1, \boldsymbol{X}_2$의 평균이 각각 m_1, m_2, 공분산 행렬이 A, B 그리고 $\boldsymbol{X}_1, \boldsymbol{X}_2$사이의 공분산이 C일 경우, 결합확률분포와 $\boldsymbol{X}_1$이 주어졌을 때 $\boldsymbol{X}_2$의 조건부 확률 분포는 다음과 같다.

$$\begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \begin{pmatrix} A & C \\ C^\top & B \end{pmatrix} \right) \tag{6}$$

$$\boldsymbol{X}_2 \,|\, \boldsymbol{X}_1 (= X_1) \sim \mathcal{N}\left(m_2 + C^\top A^{-1}(x_1 - m_1),\ B - C^\top A^{-1} C \right) \tag{7}$$

이제부터 표기를 단순화하기 위하여, 조건부를 나타내는 기호 $\cdot \,|\, \cdot\, (=)$에서 오른쪽 변수가 의미하는 확률변수(볼드체)와 괄호안의 실현된 값은 볼드체가 아닌 괄호안의 실현된 값 기호만 표기한다. 윗 식의 예를 들면 $\boldsymbol{X}_2 \,|\, \boldsymbol{X}_1 (= X_1)$를 $\rightarrow \boldsymbol{X}_2 \,|\, x_1$로 표기한다.

따름정리 2.2 ([9, Ch. 2, 식 2.19]): 머신러닝의 용어에서 관측치의 위치를 데이터 사이트*(datasite)*라 지칭한다. 관측 가능한 데이터 사이트를 두 부분으로 분리하여 관측된 데이터 사이트를 $\boldsymbol{X}_1$, 비관측된 데이터 사이트를 $\boldsymbol{X}_2$, 해당하는 *GPR*함수를 각각 $\boldsymbol{g}_1 = g(\boldsymbol{X}_1),\ \boldsymbol{g}_2 = g(\boldsymbol{X}_2)$라 하자. $\boldsymbol{g}_1$의 실현 값 g_1을 사전정보*(Prior Information)*로 갖는 $\boldsymbol{g}_2$의 사후분포*(Posterior Distribution)*는 다음과 같다.

$$\boldsymbol{g}_2 \,|\, g_1 \sim \mathcal{N}\left(m_2 + k_{12} k_{11}^{-1}(g_1 - m_1),\ k_{22} - k_{12} k_{11}^{-1} k_{12} \right) \tag{8}$$

여기서 m_1, m_2는 각각 $\boldsymbol{g}_1, \boldsymbol{g}_2$의 평균, $k_{pq} = k(\boldsymbol{X}_p, \boldsymbol{X}_q)$는 $\boldsymbol{g}_p, \boldsymbol{g}_q$, $p, q \in \{1, 2\}$의 공분산 행렬이다.

따름정리 2.3 ([9, Ch. 2, 식 2.22]): $g(\boldsymbol{X}_1)$의 관측오차 σ_n^2가 존재하는 경우, 관측모형 $\boldsymbol{y}_1 = g(\boldsymbol{X}_1) + \epsilon,\ \epsilon \sim \mathcal{N}(0, \sigma_n^2 \boldsymbol{I})$을 가정한다. $\boldsymbol{X}_1$의 개수는 n_1, $\boldsymbol{I}$는 $n_1 \times n_1$차원 단위행렬이다. 이때 $\boldsymbol{g}_1$대신 $\boldsymbol{y}_1$의 실현 값 y_1을 사전정보로 갖는 $\boldsymbol{g}_2$의 조건부 분포는 다음과 같다.

$$\boldsymbol{g}_2 \,|\, y_1 \sim \mathcal{N}\left(m_2 + k_{12}(k_{11} + \sigma_n^2 I)^{-1}(y_1 - m_1),\ k_{22} - k_{12}\left(k_{11} + \sigma_n^2 I\right)^{-1} k_{12}\right) \tag{9}$$

관측오차를 포함할 경우 실제 관측치 $\boldsymbol{y}$가 $\boldsymbol{X}$의 각 점에 대하여 $\sigma_n^2 > 0$만큼의 불확실성(분산)을 갖는 정규분포를 따른다고 가정한다. 그러면 GPR 문제는 아래와 같이 표현된다.

$$\begin{aligned} \boldsymbol{y} &= g(\boldsymbol{X}) + \epsilon,\ \epsilon \sim \mathcal{N}(\boldsymbol{0}, \sigma_n^2 \boldsymbol{I}), \\ &\sim \mathcal{N}(m(\boldsymbol{X}), k(\boldsymbol{X}, \boldsymbol{X}) + \sigma_n^2 \boldsymbol{I}) \end{aligned} \tag{10}$$

여기서 $\boldsymbol{I}$는 $nm \times nm$단위행렬이다.

비관측 위치 $\boldsymbol{X}_2$의 분산의 조건부 분포 $\boldsymbol{g}_2$와 조건부 추정치 $\widehat{\boldsymbol{g}_2}$ 따름정리 2.2, 2.3와 같은 의미로 $\boldsymbol{g}_1, \boldsymbol{g}_2$를 사용하자. 관측오차가 존재하는 훈련(관측된)데이터 y_1를 사용한 조건부 추정함수 $\boldsymbol{g}_2 \,|\, y_1$과 그 평균 $\widehat{\boldsymbol{g}_2} \,|\, y_1$을 (2.1), (2.3)의 절차에 따라 계산할 수 있다.

$$\begin{pmatrix} \boldsymbol{g}_2 \\ \boldsymbol{y}_1 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} m_2 \\ m_1 \end{pmatrix}, \begin{bmatrix} k_{22}, & k_{21} \\ k_{12}, & k_{11} + \sigma_n^2 \boldsymbol{I} \end{bmatrix}\right) \tag{11a}$$

$$\boldsymbol{g}_2 \,|\, \boldsymbol{y}_1 (= y_1) \sim \mathcal{N}\big(\underline{\widehat{\boldsymbol{g}_2} \,|\, y_1},\ k_{22} - k_{21}(k_{11} + \sigma_n^2 \boldsymbol{I})^{-1} k_{12}\big) \tag{11b}$$

$$\underline{\widehat{\boldsymbol{g}_2} \,|\, y_1} = m_2 + k_{21}(k_{11} + \sigma_n^2 \boldsymbol{I})^{-1}(y_1 - m_1) \tag{11c}$$

여기서 $m_p = m(\boldsymbol{X}_p)$는 $\boldsymbol{g}_p$ (단, $p \in \{1, 2\}$)의 평균, $k_{pq} = k(\boldsymbol{X}_p, \boldsymbol{X}_q)$는 $\boldsymbol{g}_p$와 $\boldsymbol{g}_q$의 공분산 행렬이다. 훈련데이터 쌍 $\boldsymbol{X}_1, y_1$에 의하여 커널함수의 하이퍼-파라메터를 계산한 다음 테스트 점 $\boldsymbol{X}_2$의 사전평균 m_2을 구할 수 있음을 의미한다.

3 GPR을 이용한 PBS분산의 추정

3.1 사용하는 표기법 및 기호의 정리

앞에서 기술한 정리와 따름정리의 결과를 이용하여 관측되지 않은 변동성을 관측된 변동성을 이용하여 추정하는 문제를 다룬다. GPR의 매개변수 집합 $\boldsymbol{X}$은 GPR함수 $g(\cdot)$의 가능한 관측 위치(데이터 사이트)를 의미한다. $\boldsymbol{X}$를 상태공간 $\mathcal{S}$에서 변동성이 관측 가능한 위치의 집합이라 두자. 거래소 시장은 고정된 행사가격, 만기가 있는 2차원 격자가 만들어지므로 $\boldsymbol{X} = \{x_{i,j} = (K_i, T_j) \,|\, i = 1, \cdots, n, j = 1, \cdots, m\}$ 형식으로 표현된다. 따라서 이제부터 상태공간은 행사가격과 만기로 이루어진 관측할 수 있는 이산공간에 국한시키고 기호로 $\boldsymbol{X}$를 사용한다. 이 중 관측된 위치의 집합을 $\boldsymbol{X}_1$, 비관측 집합을 $\boldsymbol{X}_2 = \boldsymbol{X} - \boldsymbol{X}_1$으로 둔다. GPR은 베이지언 확률을 적용한다. 따라서 매개변수, 확률변수, 매개변수를 갖는 확률변수의 집합인 확률함수(GPR함수), 이들의 조건부 분포 및 그 실현값을 나타내는 기호를 주의하여야 한다. 논문에서 사용하는 표기를 아래에 정리한다.

(1) $\boldsymbol{x} = (x, t) \in \boldsymbol{X}$ 매개변수의 (가능한) 임의의 위치를 나타낸다.

(2) f는 로그선도가격, g는 PBS분산을 추정하기위한 GPR함수이다.

(3) $\boldsymbol{y}, \boldsymbol{y}_1$은 GPR함수 $g(\boldsymbol{X}), g(\boldsymbol{X}_1)$에 관측오차를 내재한 분포이다.

(4) $\boldsymbol{g}_1, \boldsymbol{g}_2$는 $\boldsymbol{X}_1, \boldsymbol{X}_2$가 매개변수인 투사된 GPR함수이다. $g \,|\, \boldsymbol{X}_1, g \,|\, \boldsymbol{X}_2$와 같다.

(5) g_1, g_2, y_1은 위의 $\boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{y}_1$가 실현된 값이다.

(6) $\boldsymbol{g}_2 \,|\, y_1$은 $\boldsymbol{y}_1$의 실현값 y_1을 사전정보로 갖는 $\boldsymbol{g}_2$의 사후분포이다. $\boldsymbol{g} \,|\, \boldsymbol{X}_2, \boldsymbol{y}_1 (= y_1)$와 같다.

(7) m, k는 GPR함수 g의 평균, 공분산을 나타내는 GPR평균함수, GPR커널함수이다.

(8) M, K는 방법1의 GPR함수 g의 평균, 공분산을 나타내는 함수이다.

(9) m_1, k_{12}, M_1, K_{12}에 쓰인 아랫첨자는 위치 $\boldsymbol{X}_1, \boldsymbol{X}_2$가 대입된 함숫값이다. (주의: n_1은 $\boldsymbol{X}_1$의 갯수)

다음의 두 소절에서는 기술한 두 가지 방식으로 나누어 분산의 GPR추정을 진행한다.

3.2 PBS시스템을 적용한 GPR 추정 - 방법1

로그 단위 선도가격의 분포 로그 단위 선도가격들의 다변량 정규분포를 나타내는 GPR 함수를 $f(\boldsymbol{X})$라 두자. $f(\boldsymbol{X})$의 평균을 $m(\boldsymbol{X})$, 공분산 구조를 커널함수 $k(\boldsymbol{X},\boldsymbol{X})$의 형태로 표현하자.

$$f(\boldsymbol{X}) \sim \mathcal{GP}(m(\boldsymbol{X}), k(\boldsymbol{X},\boldsymbol{X})) \tag{12}$$

평균함수(PBS분산)의 사전분포 $g(\boldsymbol{X}) = -2f(\boldsymbol{X})$라 두자. $x = (K,T) \in \boldsymbol{X}$에 대하여 $g(x)$의 평균 $-2m(x) = \sigma^2_{K,T}T$이므로 PBS시스템의 x위치의 PBS분산과 일치한다. 즉 $g(\boldsymbol{X})$의 공분산 행렬 $K(\boldsymbol{X},\boldsymbol{X}) = 4k(\boldsymbol{X},\boldsymbol{X})$이다.

$$\begin{aligned} g(\boldsymbol{X}) &\sim \mathcal{GP}(-2m(\boldsymbol{X}), 4k(\boldsymbol{X},\boldsymbol{X})), \\ &\sim \mathcal{GP}(M(\boldsymbol{X}), K(\boldsymbol{X},\boldsymbol{X})) \end{aligned} \tag{13}$$

즉 $g(\boldsymbol{X})$를 PBS분산을 추정하기 위한 GPR구조로 만들기 위하여 평균벡터를 $M(\boldsymbol{X}) = -2m(\boldsymbol{X})$, 공분산행렬을 $K(\boldsymbol{X},\boldsymbol{X}) = 4k(\boldsymbol{X},\boldsymbol{X})$으로 재정의하였다. 로그 단위 선도가격에 대한 GPR 함수 $f(\boldsymbol{X})$에서 스케일만 변화시킨 함수 $g(\boldsymbol{X})$는 $\boldsymbol{X}$ 위의 PBS분산에 대한 GPR함수이다. PBS시스템의 식 3에 적용시키면 $f(\boldsymbol{X})$의 평균은 커널함수와 다음과 같은 관계를 갖음을 알 수 있다.

$$m(\boldsymbol{X}) = -\frac{1}{2}\text{diag}\left(k(\boldsymbol{X},\boldsymbol{X})\right)$$

여기서 기호 diag는 대각원소를 의미한다. $g(\boldsymbol{X})$의 사전분포식 (13)의 커널함수 $K(\boldsymbol{X},\boldsymbol{X})$와 사전평균함수 $M(\boldsymbol{X})$는 다음 관계를 갖는다.

$$M(\boldsymbol{X}) = \frac{1}{4}\text{diag}\left(K(\boldsymbol{X},\boldsymbol{X})\right) \tag{14}$$

일반적인 GPR의 응용에서 평균함수를 0으로 가정한다. 그 이유는 GPR함수에 대한 특별한 가정을 하기 힘든 경우가 많기 때문이다. 그러나 방법1은 사전분포의 평균함수 $M(\boldsymbol{X})$와 커널의 관계가 명시적으로 주어진다. 이 경우 비관측 위치를 포함한 $M(\boldsymbol{X})$를 RBF근사를 이용하여 계산하여야 한다. 따라서 일반적인 GPR 방법과 차별점이 있다. 그러나 RBF를 커널함수로 사용하는 GPR방법은 RBF근사의 입장에서는 본질적으로 동일함을 보였다. 일반적인 GPR 방법이 갖는 장점은 베이지언에 기반하여 커널의 파라메터를 최대우도 추정법(Maximum Likelihood Estimation: MLE)을 이용하여 최적화를 할 수 있다는 점이다. 그러나 방법1은 커널이 $M(\boldsymbol{X})$의 결과로 주어지는 식 (14)을 적용한다. 따라서 GPR함수인 PBS분산의 평균은 0이 아니다. 그러므로 GPR함수(PBS분산)의 사전분포를 따로 추정하여야 하며 이 사전분포는 커널함수인 공분산행렬과 조건 (14)을 만족하여야 한다. 따라서 커널을 임의로 가정할

수 없다. GPR함수인 PBS분산의 사전분포를 구하기 위하여 RBF에 의한 곡면근사 자체가 사전분포를 구하는 방법이 된다. 방법 1의 연구는 곡면근사를 위하여 RBF함수의 일종인 TPS를 이용하였다. 언급한 변동성 곡면 근사과정은 다음 사항들을 포함하여 완성할 수 있다.

(1) 실무자의 블랙-숄즈 모델과 GPR방법을 결합하는 PBS시스템의 도입.

(2) 변동성을 분산으로 바꾸어 분산곡면을 근사하는 이유.

(3) 칼랜다 차익거래 방지를 제약조건하의 분산곡면 생성을 위한 최적화 문제로 설정하고 누락데이터 대치법(Missing Data Imputation)을 적용한 순차적 부스트래핑 전략

(4) RBF 근사의 의 성질을 이용한 최적복원정리를 이용한 분산곡면의 유일해 및 존재 정리.

이 논문에서는 제시하는 GPR근사방법의 목적에 해당하는 항목 (1), (2)를 중심으로 기술한다. 수학적 증명과 엄밀성이 요구되는 나머지 사항은 추후의 연구로 계속 진행할 계획이다.

(1): PBS와 GPR방법을 적용하여 베이지언 추정에 의한 이론적 근거를 추가하였다. (2): 확장된 TPS 기저를 사용하고 제약조건을 부여한 이차식 최적화를 적용하여 사전분포를 계산한다. "변동성 데이터는 특정 영역에서 매우 희박하게 관측되며 관측된 데이터로 구성한 곡면의 곡률이 크기 때문에 시간의 증가에 따른 선형성이 큰 분산으로 변경하여 근사시키는 것"이 TPS를 적용하는 주된 이유이다. 같은 이유로 선형예측자인 베이지언 추정에도 적합하다. 라스무쎈(Rasmussen) [9, Ch. 2.2, Ch. 5.4]에 의하면, 부록에 기술한 GPR함수의 평균을 근사하는 방법인 재생커널 복원법(RKHS)와 GPR의 유사성을 알 수 있다. (3)은 논문 [6]의 결과를 포함하는 개선된 방식이다. (4)는 최소한의 데이터를 사용하여 분산곡면의 구성이 가능함을 증명하는 정리이다.

GPR함수의 사전분포인 분산곡면이 생성된 다음, 커널을 생성한다. GPR함수가 따르는 다변량분포에서, 위치 $\boldsymbol{x}_i$의 표준편차를 v_i, $\boldsymbol{x}_i$와 $\boldsymbol{x}_j$의 상관계수를 ρ_{ij}로 가정하면 $K_{ij} = v_i v_j \rho_{ij}$ 형식이다. 식 (14)에서 $v_i = 2\sqrt{M_i}$이므로 적절한 방식[5]으로 상관계수 ρ_{ij}를 추정한다면 커널 행렬의 값 $K_{ij} = 4\sqrt{M_i M_j}\rho_{ij}$ 를 유추할 수 있다.

5) 데이터가 희박하므로 여러 일자의 관측된 변동성들을 수집하여 샘플 상관계수를 측정

마지막 단계로 커널의 정보를 이용한 비관측 분산에 대한 베이지언 룰에 근거한 수정 작업을 한다. 곡면근사에서 각 점 $\boldsymbol{x}_i \in \boldsymbol{X}, i = 1, \cdots, nm$의 분산 $M_i = M(\boldsymbol{x}_i) = \frac{1}{4}\text{diag}(K(\boldsymbol{x}_i, \boldsymbol{x}_i))$를 계산한다. $\boldsymbol{y}_1$의 관측 값 y_1이 주어진, 비관측 위치의 PBS분산 $\boldsymbol{g}_2 \,|\, y_1$과 평균벡터 $\widehat{\boldsymbol{g}_2} \,|\, y_1$은 아래 식과 같다.

$$\boldsymbol{g}_2 \,|\, y_1 \sim \mathcal{N}\big(\widehat{\boldsymbol{g}_2} \,|\, y_1,\, K_{22} - K_{21}(K_{11} + \sigma_n^2 \boldsymbol{I})^{-1} K_{12}\big) \tag{15a}$$

$$\widehat{\boldsymbol{g}_2} \,|\, y_1 = M(\boldsymbol{X}_2) + K_{21}(K_{11} + \sigma_n^2 \boldsymbol{I})^{-1}(y_1 - M(\boldsymbol{X}_1)) \tag{15b}$$

여기서 $K_{11} = K(\boldsymbol{X}_1, \boldsymbol{X}_1), K_{21} = K(\boldsymbol{X}_2, \boldsymbol{X}_1), K_{22} = K(\boldsymbol{X}_2, \boldsymbol{X}_2), M(\boldsymbol{X}_1) = \frac{1}{4}\text{diag}(K_{11}), M(\boldsymbol{X}_2) = \frac{1}{4}\text{diag}(K_{22})$이다. $y_1 - M(\boldsymbol{X}_1)$은 관측된 분산과 곡면근사에 의한 분산값의 차이이다. 비관측영역 $\boldsymbol{X}_2$의 곡면근사에 의한 분산값 $M(\boldsymbol{X}_2)$에 이 차이만큼이 베이지언 룰에 의하여 반영됨을 의미한다.

3.3 일반적인 GPR 추정 - 방법2

PBS시스템을 그대로 적용하는 방법1과 달리 평균과 커널함수의 관계를 주지 않는다. GPR함수는 사전평균을 **0**으로 두고 공분산은 동일한 파라메터를 갖는 커널함수로 둔다.

평균함수(PBS분산)의 사전분포와 비관측 평균함수의 사후분포

$$\begin{aligned} & g(\boldsymbol{X}) \sim \mathcal{GP}(\boldsymbol{0}, k(\boldsymbol{X}, \boldsymbol{X})) \\ & \boldsymbol{g}_2 \,|\, y_1 \sim \mathcal{N}\big(\widehat{\boldsymbol{g}_2} \,|\, y_1,\, k_{22} - k_{21}(k_{11} + \sigma_n^2 \boldsymbol{I})^{-1} k_{12}\big) \\ & \widehat{\boldsymbol{g}_2} \,|\, y_1 = k_{21}(k_{11} + \sigma_n^2 \boldsymbol{I})^{-1} y_1 \end{aligned} \tag{16}$$

식 (16)의 마지막 식인 기댓값 $\widehat{\boldsymbol{g}_2} \,|\, y_1$은 2가지 방식의 해석이 가능하다.

- 첫째: 관측치 벡터 y_1의 선형결합으로 보는 방식.
  변동성 추정 문제에서 매개변수공간 $\boldsymbol{X}$의 전체 크기는 nm, 관측치 $\boldsymbol{X}_1$의 사이즈를 n_1이라 두자. 비관측치 벡터 $g_2 = \widehat{\boldsymbol{g}_2} \,|\, y_1$의 각 원소의 값들은 오차를 갖는 관측치 $y_{1,i}$들의 선형 결합으로 표현된다:
  $$g_{2,j} = \sum_{i=1}^{n_1} w_{j,i} y_{1,i},\; j = 1, \cdots, nm - n_1$$
  즉 GPR은 베이지언 선형 예측자이다.[6] RBF커널의 특징에 의하여 두 위치의 거리가 가까울수록 j, i사이의 공분산이 더 크다.[7] 따라서 커널이 관측치들에

6) 베이지언 선형 예측자인 GPR이 변동성 곡면대비 선형성이 강조된 분산곡면을 잘 추정하는 직관적인 이유를 준다.

7) 선형예측자: 존재하는 관측치 $y_{1,i}$들의 선형결합으로 예측값을 만드는 함수.

균질하게 적용될 경우 비관측치에 미치는 관측치의 영향은 상호간의 거리에 의존한다. 식 (16)은 촐레스키 분해를 사용하여 계산할 수 있다.

- 두번째: 커널함수의 선형 합으로 표현되는 RBF함수 근사방식.
  식 (16)를 풀어서 쓰면 $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_{n_1})^\top = (k_{11} + \sigma_n^2 \boldsymbol{I})^{-1} y_1$, $k_{j,i} = k_{i,j} = k(\boldsymbol{x}_j, \boldsymbol{x}_i)$.

$$g_{2,j} = \sum_{i=1}^{n_1} \alpha_i k_{j,i},\ j = 1, \cdots, nm - n_1 \tag{17}$$

  윗 식에서 GPR의 공분산에 RBF커널을 이용한 GPR의 평균함수(PBS분산)가 본질적으로 1차기저를 사용하지 않은 RBF함수근사식과 일치한다.[8] σ_n^2은 평활화를 위한 벌점항 상수의 역할을 한다. GPR모델은 베이지언 방법의 일반적인 최적화 툴인 우도함수 최대화를 적용하여 RBF함수의 형상모수(shape parameter)인 하이퍼-파라메터를 계산한다. 이 방식을 적용한 방법2의 결과는 제곱지수(가우시언)커널, 마턴(Matérn)커널 [9, Ch. 4]이 유용하였다. 이들은 GPR에서 자주 사용되는 커널의 종류이다. 그 중 마턴(Matérn)3/2커널이 매우 효과적임을 발견하였다. 전통적인 RBF함수근사 또한 형상모수를 GCV, LOOCV등의 방식으로 추정한다. 계산속도의 문제가 있으며, 변동성 곡면 근사에 대한 결과가 만족스럽지 않았다 [6].

커널함수의 선택과 하이퍼-파라메터의 추정 GPR의 커널함수는 RBF함수를 적용하는 것이 일반적이다.[9] RBF커널들 중에서 가우시언 커널, 마턴 커널은 이를 결합시켜 만든 곡면의 적합도를 높이기 위한 형상모수가 존재한다[10]. GPR은 이를 하이퍼-파라메터라 부른다. 이 연구에서는 가우시언 커널, 마턴 커널을 사용하여 실증분석을 한다.

관측오차를 내재한 GPR함수 $\boldsymbol{y}$에 관한 주변부 우도 $p(\boldsymbol{y}; \Theta)$는 아래의 적분 식이다.

$$p(\boldsymbol{y}; \Theta) = \int \underbrace{p(\boldsymbol{y} \mid g(\boldsymbol{X}; \Theta))}_{\text{likelihood}} \underbrace{p(g(\boldsymbol{X}; \Theta))}_{\text{prior}}\, dg \tag{18}$$

GPR함수 $\boldsymbol{y}$는 다변량 가우시언 분포를 따른다 (2). 다행히 이 적분은 가우시언함수의 특징에 의하여 해석적으로 계산이 된다. $\boldsymbol{y}$의 우도함수를 최대로 하는 커널의 하이퍼-

8) 1차기저를 포함하여 GPR모델을 만들 수 있다. 3.3를 참조

9) 와바(Wahba) [10], 마턴(Matérn) [7]등에 의하여 초기에 연구된 산포된 데이터 적합문제에 사용되었다. 응용수학에서는 다차원 곡면근사를 위한 메쉬프리(Mesh Free)방법 또는 재생커널이론(Reproducing Kernel Hilbert Space: RKHS) [5]에서 RBF함수가 사용된다. 이들 연구에서 함수적 구조가 알려져 있지 않은 데이터로만 이루어진 곡면을 근사하는 문제에 매우 효과적임이 검증되었다.

10) Thin Plate Spline등의 일부는 제외

파라메터를 최적화 알고리즘을 사용하여 찾을 수 있다. 주어진 관측 값 $\boldsymbol{y}_1 = y_1$에 대하여 계산한 파라메터 값은 분포함수 $p(\boldsymbol{y} \mid \theta)$에 대한 로그를 취하여도 결과는 동일하다.

우도함수의 로그 식과 미분의 계산은 아래와 같다. [9, Ch. 2.2, Ch. 5.4] 참조.

$$\log p(y_1 \mid \Theta) = -\frac{1}{2}(y_1 - m_1)^\top \left(k_{11} + \sigma_n^2 \boldsymbol{I}\right)^{-1} (y_1 - m_1) - \frac{1}{2}\log\left(k_{11} + \sigma_n^2 \boldsymbol{I}\right) \tag{19a}$$

$$- \frac{n_1}{2}\log 2\pi \frac{\partial}{\partial \Theta} \log p(y_1 \mid \Theta) \tag{19b}$$

$$= \frac{1}{2}(y_1 - m_1)^\top k_{11}^{-1} \frac{\partial k_{11}}{\partial \Theta} k_{11}^{-1}(y_1 - m_1) - \frac{1}{2}\mathrm{tr}\left(k_{11}^{-1}\frac{\partial k_{11}}{\partial \Theta}\right) \tag{19c}$$

$$= \frac{1}{2}\mathrm{tr}\left((\boldsymbol{\alpha}\boldsymbol{\alpha}^\top - k_{11})^{-1}\frac{\partial k_{11}}{\partial \Theta}\right) \tag{19d}$$

여기서 $k_{11} = k(\boldsymbol{X}_1, \boldsymbol{X}_1; \Theta)$, $\boldsymbol{\alpha} = (k_{11} + \sigma_n^2 \boldsymbol{I})^{-1}(y_1 - m_1)$, N은 $\boldsymbol{X}_1$의 갯수이다. 방법2일 경우 $m_1 = 0$이다. 식 (19b)는 Θ를 변수로 갖는 비선형 스칼라 함수이다. 비선형 스칼라 함수의 최소 또는 최댓값은 파라메터의 갯수가 식의 갯수보다 적거나 같으며 변수에 대한 1, 2계 미분이 계산될 경우 뉴튼 방식의 최적화 방법이 적당하다. 특히 파라메터의 개수가 상대적으로 적으며 유계인 구간이면 쉽게 해를 구할 수 있다. 일반적인 GPR추정에서 모든 점에 서로 다른 파라메터를 갖는 방식은 관측 위치의 개수 대비 많은 자유도를 주기 때문에 해를 찾기 힘들다. 따라서 $2 \sim 3$개의 파라메터를 갖는 RBF함수를 모든 위치의 공분산에 동일하게 주는 것이 일반적이다. RBF 커널을 가우시언 또는 마턴(Matérn)타입으로 줄 경우 대칭행렬 $k_{11}, k_{11} + \sigma_n^2 \boldsymbol{I}$이 양정치가 된다. 따라서 촐레스키 분해를 적용하면 경제적으로 역행렬 $k_{11}^{-1}, (k_{11} + \sigma_n^2 \boldsymbol{I})^{-1}$을 구할 수 있다 [9, Algorithm 2.1, p. 19]. 따라서 최적화를 위한 과정에서, 파라메터 Θ가 변화함에 따른 로그우도함수 $L = \log p(y_1 \mid \Theta)$, 그래디언트 함수 $\nabla_\Theta L = \frac{\partial \log p(y_1 \mid \Theta)}{\partial \Theta}$와 의사 헤시언 함수 $(\nabla_\Theta L)^\top \cdot \nabla_\Theta L$를 어렵지 않게 계산할 수 있다. 관측가능한 변동성 데이터는 일반적으로 $30 \sim 200$개 정도이다. 신뢰영역 반사법(Trust Region Reflective) [8, Ch. 4]을 사용하여 효율적으로 해를 구할 수 있다.

RBF 커널함수 테스트에 적용한 두 가지 타입의 RBF 커널함수는 다음과 같다.

$$\text{가우시언 커널:} \qquad k(\boldsymbol{x}, \boldsymbol{x}' \mid \sigma, l) = \sigma^2 \exp(-\frac{r^2}{2l^2}) \tag{20}$$

$$\text{마턴(Matérn)3/2 커널:} \qquad k(\boldsymbol{x}, \boldsymbol{x}' \mid \sigma, l) = \sigma^2(1 + \sqrt{3}r)\exp(-\sqrt{3}r), \tag{21}$$

여기서 $r = ||\boldsymbol{x} - \boldsymbol{x}'||_2 = \sqrt{(x - x')^2 + (t - t')^2}$이다.

선형기저를 추가한 확장 RBF근사에서 RBF기저와 더불어 일차식 기저를 추가하여 확장된 형태가 주로 쓰인다. 우리의 2차원 문제에서 위치변수 $\boldsymbol{x} = (x, t)$에 대한 일차식 기저는 $\text{span}\{x, t, 1\} = b_1 x + b_2 t + b_3, b_i \in \mathbb{R}^1$ 형태이다. 그 이유는 RBF기저만 사용할 경우 RBF함수가 갖는 비선형성 때문에 평면과 같은 편평한 모양을 근사하기 힘들기 때문이다. 결국 식 (17)에 의하면 GPR함수의 기댓값은 커널의 선형결합으로 표현이 되기 때문에 일차식 기저를 추가되면 더 좋은 근사가 될 것이라 예상할 수 있다. 라스무쎈 [9, Ch 2.7]에 의하면 일차식 기저를 포함하는 GPR함수 모델을 구축할 수 있다. 다변량 정규분포를 가정한 새로운 확률변수 $\boldsymbol{\beta}$를 커널함수의 새로운 파라메터로 추가하여 기존의 GPR모델에 있는 RBF의 하이퍼-파라메터와 결합시켜 만들 수 있다.

$$\boldsymbol{\beta} \sim \mathcal{N}(\boldsymbol{b}, \boldsymbol{B}),\ \boldsymbol{b} = (b_1, b_2, b_3)^\top$$
$$g(\boldsymbol{x}) \sim \mathcal{GP}\left(m(\boldsymbol{x}) + \bar{\boldsymbol{x}}^\top \boldsymbol{b},\ k(\boldsymbol{x}, \boldsymbol{X}) + \bar{\boldsymbol{x}}^\top \boldsymbol{B} \bar{\boldsymbol{x}}\right)$$

이때 $\bar{\boldsymbol{x}} = (x, t, 1)$, $k(\boldsymbol{x}, \boldsymbol{X})$는 $k(\boldsymbol{X}, \boldsymbol{X})$의 $\boldsymbol{x}$가 위치한 행. 확률변수 $\boldsymbol{\beta}$의 공분산 행렬 $\boldsymbol{B}$의 원소를 $b_{i,j}, i, j = 1, 2, 3$이라 두자. $\bar{\boldsymbol{x}}^\top \boldsymbol{B} \bar{\boldsymbol{x}}$가 아래 식으로 계산되므로

$$\begin{pmatrix} x, t, 1 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{12} & b_{22} & b_{23} \\ b_{13} & b_{23} & b_{33} \end{pmatrix} \begin{pmatrix} x \\ t \\ 1 \end{pmatrix} = b_{11}x^2 + b_{22}t^2 + 2b_{12}xt + 2b_{13}x + 2b_{23}t + b_{33} \tag{22}$$

방법1의 경우 커널을 RBF함수로 가정할 필요는 없지만, 평균과 공분산은 관계식 (14)를 만족시켜야 하므로 다음이 성립한다.

$$b_1 x + b_2 t + b_3 = \frac{1}{4}\left(b_{11}x^2 + b_{22}t^2 + 2b_{12}xt + 2b_{13}x + 2b_{23}t + b_{33}\right) \tag{23}$$

따라서 $b_{11} \to \epsilon, b_{22} \to \epsilon, |b_{12}| \to \epsilon, \epsilon > 0$, 그리고 곡면근사의 결과에서 얻은 b_1, b_2, b_3를 각각 $\frac{1}{2}b_{13}, \frac{1}{2}b_{23}, \frac{1}{4}b_{33}$으로 두면 커널행렬과 분산의 추정 값이 오차 $\epsilon > 0$ 한도내에서 일치한다. 이제 확장된 기저를 사용한 GPR평균함수의 곡면근사의 결과와 이 GPR이 적용된 PBS시스템과 일치성을 갖게할 수 있다.

4 결론

이 연구는 바닐라 콜,풋 옵션의 평가를 위하여 사용하는 실무자의 블랙숄즈 모델이 행사가, 만기에 따라 서로 다른 변동성을 적용하는 방식을 다변량 정규분포로 해석하는 문제를 연구하면서 착안되었다. PBS시스템3의 특징은 시장에서 블랙-숄즈 모델로 평가되는 상품이 갖는 불확실성의 대상을 분산으로 표현한 점이다. 연구과정에서, 추정하려는 PBS분산 또는 로그선물가격 그리고 PBS시스템의 다변량 정규분포는 머

신러닝에서 연구되는 가우시언 프로세스 회귀방법과 자연스럽게 연결될 수 있음을 인지하였다. GPR은 머신러닝의 범주에 속하지만 블랙박스 머신러닝이 아니다. 다변량 정규분포이론과 베이지언 추론이 결합된 정교한 구조를 갖는 해석적 방법이다. 먼저 블랙숄즈 모델 자체의 관점을 유지하여 로그단위선물가격이 따르는 만기의 분포들을 다변량 정규분포로 해석한다. 이 관점에 따르면 평균에 해당하는 PBS분산들과 커널인 공분산 행렬간의 함수관계가 존재한다. GPR함수의 사전분포를 구하기 위하여 커널의 상관계수 행렬을 과거 데이터 혹은 다른 방식으로 준비한다. GPR함수의 평균인 PBS 분산들을 박판스플라인 확장기저로 근사하여 평균과 공분산을 동시에 구할 수 있다. GPR을 적용하는 일반적인 방식과는 다르지만, 실무자의 블랙숄즈 모델과 GPR을 연결하는 흥미로운 구조를 내포하기 때문에 강한 이론적 틀을 갖는다. 두번째 방식은 행사가, 만기로 이루어진 상태공간위의 분산들을 직접 GPR의 대상으로 한다. 다양한 커널함수들을 적용할 수 있다. 하이퍼-파라메터를 계산하여 최적 모델을 도출한다. 또한 머신러닝이 연구되면서 개발된 여러 기술들을 접목시킬 수 있는 장점이 있다. 두 방식 모두 수집한 데이터가 무위험 차익거래를 허용하지 않는 조건에 부합하면 수집 데이터가 적은 경우일지라도 베이지언 추정 이론에 근거하여 비관측 변동성을 효과적으로 추정한다.

5 실증분석

테스트를 위해 변동성 데이터의 많은 부분이 누락된 S&P500 (2019년 7월 11일)를 사용하였다. 방법1은 사전분포를 계산하는 TPS곡면근사만 하였으며 방법2는 마틴(Matérn)3/2 커널을 적용한 GPR 추정을 하였다. 방법2에 의한 최적 하이퍼-파라메터는 $\sigma^2 = 0.1739, l = 9.604, \sigma_n = 4.425$e-16이다. 두 방법 모두 주어진 사전 데이터는 유의한 오차없이 복원한다. 방법1은 베이지언 수정을 적용하기전 단계(15b)의 TPS에 의한 분산곡면 근사이다. 따라서 TPS근사의 특징에 의하여 편평함이 강조되고 방법2와 비교하여 거친 모양을 보인다. 특히 데이터가 부족한 영역에서 더욱 그렇다. 반면 방법2는 매우 매끈하게 연결된 분산곡면을 생성하여 준다. 두 방식 모두 RBF함수의 일차결합으로 분산곡면을 생성하지만 방법2에 의한 베이지언 방식은 하이퍼-파라메터의 최적 추정에 의한 결과 관측치의 정보에 비 관측치가 매끄럽게 결합되었음을 볼 수 있다. 또한 풋옵션의 최외가격 영역 변동성이 방법1과 비교하여 높음을 관찰할 수 있다. 미답 영역의 추론능력이 단순 곡면 추정 대비 상대적으로 우월함을 의미한다. 그러나 방법 1은 사전분포를 추정하는 RBF곡면근사를 보여준 결과이다. 새로운 다양한

방법을 개발할 수 있을 것으로 판단한다.

References

1. T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, 2003.
2. F. Black, The Pricing of Commodity Contracts, *Journal of Financial Economics*, 3 (1976), 167–179.
3. F. Black, Myron Schoels, The Pricing of Options and Corporate Liabilities, *Journal of Polytical Economy* 81(3) (1973), 637–654.
4. Bernard Dumas, Jeff Fleming and Robert E. Whaley, Implied volatility functions: Empirical tests, *The Journal of Finance*, LIII0 (6): 2059–2106, 1998.
5. Gregory E. Fasshauer, *Meshfree approximation methods with MATLAB*, World Scientific, 2007.
6. Lim Hyuncheul, Construction of the Implied Volatility Surface by Thin Plate Spline Function, 금융공학연구 18(4) (2019), 1–36.
7. Bertil Matérn, Wiley StatsRef: Statistics Reference Online 2018.
8. Jorge Nocedal, *Numerical Optimization, Volume 2, 2nd Edition*, volume 1, Springer, 1999.
9. Carl E. Rasmussen and Christopher K. I. Williams, *Gaussian processes for machine learning*, MIT press, Cambridge, MA, 2006.
10. Grace Wahba, *Spline models for observational data*, Siam, 1990.

Table 1. S&P500 VOLATILITY 07/11/2019, Spot: 3085.18

Maturity(days)	43	253	407	561	771	1107	1499	1870
ZeroRate	0.01708	0.01795	0.01727	0.01685	0.01633	0.01604	0.01592	0.01598
DividendRate	0.01510	0.01648	0.01668	0.01645	0.01626	0.01627	0.01633	0.01649
0.50	-	-	-	0.30484	-	-	-	-
0.55	0.41464	0.33913	0.30897	-	0.27159	-	-	-
0.70	-	0.2768	-	-	-	-	-	-
0.75	0.31032	-	-	-	-	-	-	-
0.80	0.27512	-	-	-	-	-	-	-
0.85	0.23864	0.2159	-	-	0.20524	-	-	-
0.90	-	-	0.19483	-	-	-	-	-
0.95	0.1601	0.17415	-	0.1828	-	-	-	0.19261
1.00	0.11289	-	-	-	-	-	-	-
1.05	0.09399	0.13216	-	-	-	0.17563	0.18099	0.18424
1.10	0.10868	-	0.13537	-	-	-	-	-
1.15	0.11499	0.10484	-	-	-	0.16327	-	-
1.20	0.12247	0.10142	-	-	-	-	-	-
1.25	0.12996	-	-	-	-	-	-	-
1.30	-	-	-	-	0.13027	-	-	-
1.45	0.13	-	-	-	-	-	-	-
1.50	0.13	0.1	0.105	0.112	0.11482	-	-	-

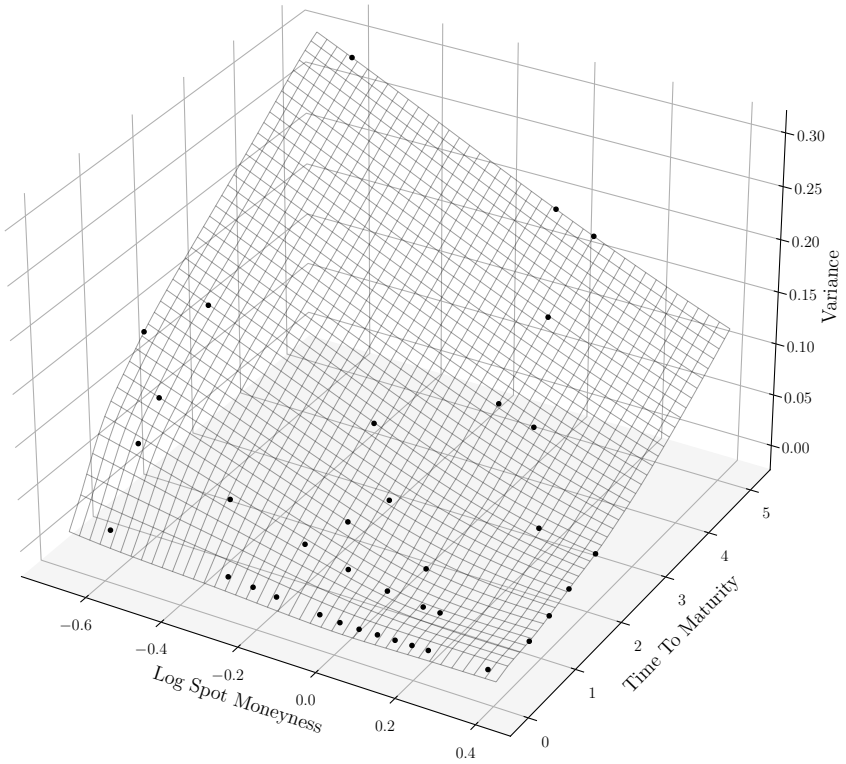

Figure 1. GPR 추정(방법2)에 의한 만기에 따른 추정된 분산곡면과 관측변동성(분산으로 변환). 행사가격은 관측된 스팟머니니스의 최소값 0.5, 최댓값 1.5를 40등분함. 시간축은 최소만기 43일과 최대 만기 5.1년을 40등분함. 격자점위의 값들은 GPR 추정에 의함. 관측 데이터는 S&P500 2019년 7월 11일, EG자산평가(주) 제공

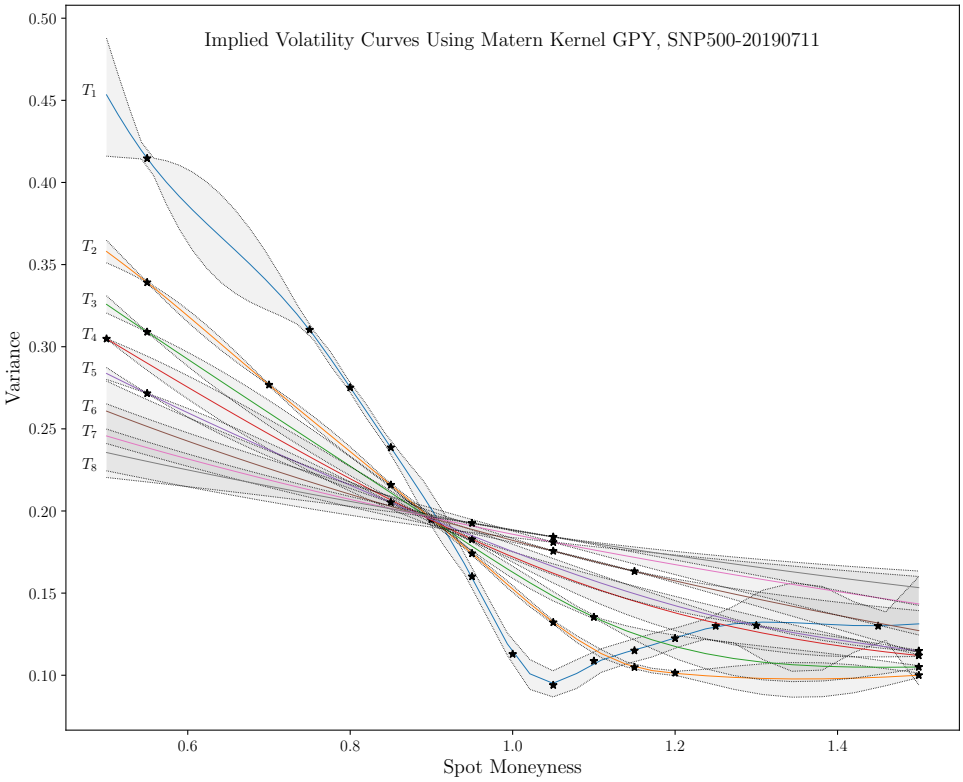

Figure 2. GPR 추정(방법2)에 의한 만기 $T_1, \cdots, T_8$에 따른 추정된 변동성 곡선과 이들의 관측치$\star$ 및 95% 신뢰구간. 관측 데이터는 S&P500 2019년 7월 11일, EG자산평가(주)제공