Journal of The Korea Society of Computer and Information
Vol. 18, No. 3, March 2013
www.ksci.re.kr
http://dx.doi.org/10.9708/jksci.2013.18.3.091

r-fold Wiener process에 대한 유한근사함수의 특성에 관한 연구

최성희*, 황석형**

A study on the properties of the finite-dimensional approximation of an r-fold Wiener Process

Sung-Hee Choi*, Suk-Hyung Hwang**

요 약

$r-$fold Wiener process는 실질적으로 infinite dimension이고, 컴퓨터는 finitely dimensional subspace만 취급할 수 있기 때문에 $r-$fold Wiener process는 컴퓨터로 구현될 수 없다. 따라서 본 논문에서는 $r-$fold Wiener process의 m-dimensional approximation 함수의 특성에 대해 연구한다.

▸ Keywords : Wiener process, average case 오차, Probability measure

Abstract

Because the r-fold Wiener process is truly infinitely dimensional and a computer can only handle finitely dimensional subspaces, we study in this paper the basic properties of the m-dimensional approximation function of the r-fold Wiener process.

▸ Keywords : Wiener process, average case error, Probability measure

• 제1저자 : 최성희
• 투고일 : 2012. 12. 27, 심사일 : 2013. 1. 29, 게재확정일 : 2013. 2. 23.
* 제1저자, 선문대학교 컴퓨터공학과
** 선문대학교 컴퓨터공학과

I. 서 론

Informaion-based complexity를 연구하는 분야에서 다루는 문제 중에서는 complexity를 분석하기가 아주 어렵거나, 아예 불가능한 문제들이 있다. 예를 들어 n개의 데이터를 사용하는 수치적분문제에서 minimal 오차가 $\Theta(n^{-p})$인 것이 알려져 있는데 여기서 비록 정수 p값을 알고 있다하더라도 Θ-notation에서 사용되는 상수의 값을 우리는 알지 못한다. 따라서 프로그램할 때 언제 계산을 끝내야할지 모르기 때문에 optimal 알고리즘을 프로그램하는데에는 어려움이 있다. 또 다른 예로 어떤 함수 $f \in C^r$(여기서 r은 0이상의 정수이며 집합 C^r은 r차 도함수가 연속인 함수의 집합을 말한다)의 global maximum의 근사값을 찾는 문제인 global optimization 문제에서는 어떤 정수 r이 $r \geq 1$일 때 $\max_x f(x)$의 분포(distribution)가 알려져 있지 않아서 가장 간단한 알고리즘에서조차 그 average case 오차를 알 수가 없다. 이러한 문제들을 다룰 때, 만약 우리가 r-fold Wiener measure ω_r에 의해 분포되는 random 함수들을 만들어낼 수 있다면 random method를 사용하여 위 문제들에 대해 complexity를 보다 쉽게 계산을 할 수 있을 것이다. 따라서 본 논문에서는 r-fold Wiener random 함수의 특성([2], [5] 참조)에 대해 연구한다.

Information-based complexity 분야에서 계산하는 average case 오차는 L_2-norm을 사용하여 측정되는 function approximation 문제에서, 어떤 양수 r에 대해서 r-fold Wiener measure ω_r과 양수 m (여기서 m은 주어진 information의 개수를 나타냄)과 information

$$M_m(f) \;=\; [L_1(f), L_2(f), \cdots, L_m(f)]$$
$$(L_i \;:\; \text{some linear functional})$$

이 주어지면, algorithm ϕ를 사용하여 이 문제를 푸는 경우 average case 오차

$$e^a(\phi, M_m, L_2; \omega_r) \;=\; \sqrt{E_{\omega_r}(\|f - \phi(M_m(f))\|_2^2)}$$

는 $\omega_r(\cdot \mid M_m(f) = y)$의 conditional mean인 algorithm $\phi(y) \;=\; \psi_m^*(y)$ 에 의해서 최소화된다. 따라서 average case 오차는 일차적으로는 문제에 주어지는 information M_m에 따라 변화하고, 이차적으로는 사용하는 알고리즘 ϕ에 따라 그 값이 변화한다. 본 논문에서는 같은 알고리즘을 사용한다고 가정할 때, average case 오차를 최소화하는 몇 가지의 information에 대한 특성에 대해 연구한다.

II. 용어정의

본 논문에서는 함수의 집합에 probability measure인 r-fold Wiener measure ω_r이 주어진다고 가정한다. 여기서 ω_r은 평균이 0이고 correlation function이

$$E_{\omega_r}(f(x)\,f(y)) = \int_F f(x) f(y)\,\omega_r(df)$$
$$= \int_0^1 \frac{(x-t)_+^r}{r!}\,\frac{(y-t)_+^r}{r!}\,dt$$

인 Gaussian measure이다. 여기서

$$(z-t)_+ = \max 0, (z-t)$$

를 의미한다. 다시 말해서 ω_r에 의해 분포되는 함수 f는 평균이 0이고 autocorrelation이 위와 같이 주어진 Gaussian stochastic process이다. 또한 ω_r에 의해 분포되어 있기 때문에 boundary condition 인

$$f(0) = f'(0) = f''(0) = \;\cdots\; = f^{(r)}(0) = 0$$

을 만족한다.

r-fold Wiener process는 사실상 infinite dimension이고 컴퓨터에서는 finitely dimensional subspaces만 취급할 수 있기 때문에 진짜 r-fold Wiener process f 는 컴퓨터로 구현될 수 없다. 따라서 process f 의 근사함수에 대하여

$$\dim(lin\, f_m) \;=\; m \;,\;\; f_m \rightarrow f \;\; as \;\; m \rightarrow \infty$$

를 만족하는 approximate random process f_m의 성질을 본 논문에서 연구한다.

III. r-fold Wiener Process의 유한근사함수

오차가 L_2-norm을 사용하여 측정되는 function approximation 문제에서, 어떤 양수 r에 대해서 r-fold Wiener measure ω_r과 양수 m(여기서 m은 주어진 information의 개수를 나타냄)과 information

$$M_m(f) = [L_1(f), L_2(f), \cdots, L_m(f)]$$
$(L_i$: some linear functional)

이 주어지면, algorithm ϕ를 사용하여 이 문제를 푸는 경우 average case 오차

$$e^a(\phi, M_m, L_2; \omega_r) = \sqrt{E_{\omega_r}(\|f-\phi(M_m(f))\|_2^2)}$$

는 $\omega_r(\cdot \mid M_m(f)=y)$의 conditional mean 인 algorithm $\phi(y) = \psi_m^*(y)$ 에 의해서 최소화된다. 이러한 최소 오차는 information M_m의 radius 라 불리며 $e^a(M_m, L_2 ; \omega_r)$로 표기된다. 따라서 어떤 함수 $g_i \in L_2$에 대해서,

$$\psi_m^*(y)(x) = \sum_{i=1}^{m} y_i \cdot g_i(x)$$

이다. 사실상 ψ_m^*는 다른 어떠한 norm을 사용한다 하더라도 function approximation 문제에서의 average case 오차를 최소화시킨다. 특히, 임의의 $q>0$ 와 $p=\infty$ 를 포함하는 임의의 $p>1$ 에 대해서

$$E_{\omega_r}(\|f - \psi_m^*(M_m(f))\|_p^q) = \min_\phi \mathrm{E}_{\omega_r}(\|\mathrm{f} - \phi(\mathrm{M_m(f)})\|_\mathrm{p}^\mathrm{q}) .$$

또한 임의의 linear operator S에 대해서 $S(\psi_m^*(M_m(f)))$ 는 $S(f)$를 approximating 하는 문제의 average case 오차를 최소화시킨다. 즉,

$$e^a(M_m, S ; \omega_r) = e^a(S \circ \psi_m^*, M_m ; \omega_r) = \min_\phi \sqrt{E_{\omega_r}(\|S(f) - \phi(M_m(f))\|^2)} .$$

여기서 random function f_m의 기본적인 성질에 대해 알아본다. 주어진 문제에 대한 average case 오차는 일차적으로는 information M_m에 따라 변화하고, 이차적으로는 사용하는 algorithm ϕ에 따라 그 값이 변한다.

Theorem1. 어떤 algorithm ϕ에 대해서 average case 오차를 최소화시키는 information은

$$M_m^*(f) = \left[\frac{\langle f, \eta_1 \rangle_2}{\sqrt{\lambda_1}}, \cdots, \frac{\langle f, \eta_m \rangle_2}{\sqrt{\lambda_m}} \right]$$

이며 이때의 오차는

$$\sqrt{E_{\omega_r}(\|f-\phi(M_m^*(f))\|_2^2)} = \Theta\left(\left(\frac{1}{m}\right)^{r+\frac{1}{2}} \right)$$

이다. 여기서 λ_i 와 η_i는 T_r의 eigenpairs이며

$$T_r \eta_i = \lambda_i \eta_i, \quad \lambda_1 \geq \lambda_2 \geq \cdots, \quad \langle \eta_i, \eta_j \rangle_2 = \delta_{ij}$$

를 만족한다. $\langle \cdot , \cdot \rangle_2$는 L_2 inner product를 나타내며 T_r는 integral operator

$$(T_r f)(x) = \int_0^1 K_r(x,t) f(t) dt$$

로써 $K_r(x,t)$는 ω_r의 covariance function 이다.

Proof: 증명은 [5]를 참조하기 바란다.

■

Theorem2. m개의 등간격점(equally spaced points)에서의 함수값으로 구성된 information

$$M_m^{st}(f) = \left[f\left(\frac{1}{m}\right), \ldots, f\left(\frac{m}{m}\right) \right] .$$

을 사용할 경우 average case 오차는

$$\sqrt{E_{\omega_r}(\|f-\psi_m^*(M_m^{st}(f))\|_2^2)} = \Theta\left(\left(\frac{1}{m}\right)^{r+\frac{1}{2}}\right)$$

이 된다.

Proof: 이 information에 대응되는 optimal algorithm ψ_m^* 는 m개의 등간격점을 사용하여 함수 f 를 보간하는 $2r+1$차의 natural spline 함수에 의해서 주어진다. 이 증명 또한 [5]를 참조하기 바란다.

■

따라서 Theorem1과 Theorem2에 의해 M_m^{st}도 M_m^* 만큼 오차를 최소화시키는 information이 된다. 본 논문에서 사용할 또 다른 information M_m^r 은 등간격점에서의 함수 $f^{(r)}$의 함수 값으로 구성된다.

$$M_m^r(f) = \left[f^{(r)}\left(\frac{1}{m}\right), \cdots, f^{(r)}\left(\frac{m}{m}\right)\right] .$$

Theorem3. Information $M_m^r(f)$을 사용하여 $e^a(M_m^r, L_2 ; \omega_r)$과 $e^a(M_m^r, \int ; \omega_r)$을 L_2-approximation 문제와 적분문제에서의 average radius라고 하자. 그러면 $r \geqq 1$ 일 때

$$e^a(M_m^r, L_2 ; \omega_r) = \Theta\left(\frac{1}{m}\right) = e^a(M_m^r, \int ; \omega_r) .$$

Proof: 상한은 [3]과 이미 알려져 있는 사실 $e^a(M_n^0, \int ; \omega_r) = \Theta(1/m)$ 에 의해

$$e^a(M_m^r, \int ; \omega_r) \leqq e^a(M_m^r, L_2 ; \omega_r) \leqq e^a(M_m^0, \int ; \omega_r)$$

이다. 따라서 본 증명에서는 적분문제에서의 하한만 증명하면 된다.

g를 ω_r에 의해 분포되어 있고, f를 w_0에 의해 분포되어 있다고 하자. $\omega_r = \omega_0 \circ D^r$이므로 $f(x) = g^{(r)}(x)$이다. 즉

$$g(x) = \int_0^1 \frac{(x-t)_+^{r-1}}{(r-1)!} f(t)\, dt .$$

따라서 적분문제에서의 M_m^r의 average radius는

$$(e^a(M_m^r, \int ; \omega_r))^2$$
$$= E_{\omega_r}\left[\left(\int_0^1 g(x)dx - \psi_r^*(M_m^r(g))\right)^2\right]$$
$$= E_{\omega_0}\left[\left(\int_0^1 \int_0^1 \frac{(x-t)_+^r}{(r-1)!} f(t)\, dt\, dx - \psi_0^*(M_m^0(f))\right)^2\right]$$

을 만족한다. 따라서

$$S(f) = \int_0^1 \int_0^x \frac{(x-t)^r}{(r-1)!} f(t)\, dt\, dx$$

라 할 때 이 문제에서의 하한은

$$e(m) = \sqrt{E_{\omega_0}\left[\left(S(f) - \psi_0^*(M_m^0(f))\right)^2\right]}$$

이 된다. 함수 f가 ω_0에 의해 분포되고 $S(f)$의 approximation을 구하는 문제에서는 ψ_0^*이 optimal 알고리즘이다.

S가 functional이기 때문에 average case 오차 $e(m)$은, ω_0에 의해 생성되는 reproducing kernel Hilbert space상에서 unit ball에 대한 worst case setting일 경우 ψ_0^*를 사용하여 S의 근사값과 같게 된다. 이것은 다음과 같은 Sobolev space인 것이 알려져 있다.([4] 참조) Norm이 $\|f\| = \|f'\|_2$이고

$$W_2^1([0,1]) = \mathrm{f} \;:\; \mathrm{f}(0)=0,\;\; \mathrm{f}\ \text{는연속이고}\ \ \mathrm{f}' \in \mathrm{L}_2$$

다시 말해,

$e(m)$

$= \{\infty_\phi$

$= \sup\{|S(f)| \;: f \in W_2^1,\ |f'|_2 \leqq 1,$

and $f(x_i) = 0,\ \ \forall i\}$.

여기서 $m_i = (x_i + x_{i+1})/2$ 이고

$$f(x) = \begin{cases} x - x_{i-1}, & \text{if } x \in [x_{i-1}, m_i], \\ x_i - x, & \text{if } x \in [m_i, x_i], \end{cases}$$

라 하면 $f \in W_2^1$이고 모든 i에 대해서 $f(x_i) = 0$ 이 되며 또한 $\int (f')^2 = 1$이 된다. 그러면

$$\begin{aligned} S(f) &= \int_0^1 \int_0^1 \frac{(x-t)_+^{r-1}}{(r-1)!} f(t)\,dt\,dx \\ &= \int_0^1 f(t) \int_t^1 \frac{(x-t)^{r-1}}{(r-1)!}\,dx\,dt \\ &= \int_0^1 f(t)\,\frac{(1-t)^r}{r!}\,dt \\ &= \sum_{i=1}^m \int_{x_{i-1}}^{x_i} f(t)\,\frac{(1-t)^r}{r!}\,dt. \end{aligned}$$

구간 $[x_{i-1}, m_i]$에서의 적분값은

$$\begin{aligned} &\int_{x_{i-1}}^{m_i} (t - x_{i-1})\,\frac{(1-t)^r}{r!}\,dt = \\ &\quad -\frac{(1-m_i)^{r+1}}{(r+1)!}\,\frac{x_i - x_{i-1}}{2} \\ &\qquad + \frac{(1-x_{i-1})^{r+2}}{(r+2)!} - \frac{(1-m_i)^{r+2}}{(r+2)!} \end{aligned}$$

이고, 구간 $[m_i, x_i]$에서의 적분값은

$$\begin{aligned} &\int_{m_i}^{x_i} (x_i - t)\,\frac{(1-t)^r}{r!}\,dt = \\ &\quad -\frac{(1-m_i)^{r+1}}{(r+1)!}\,\frac{x_i - x_{i-1}}{2} \\ &\qquad - \frac{(1-m_i)^{r+2}}{(r+2)!} + \frac{(1-x_i)^{r+2}}{(r+2)!} \end{aligned}$$

이다. 따라서 $\zeta(x) = (1-x)^{r+2}/(r+2)!$ 이고 $h = 1/m$ 이라 할 때

$$S(f) = \sum_{i=1}^m \left(\zeta(x_{i-1}) + \zeta(x_{i-1}+h) - 2\zeta\left(x_{i-1} + \frac{h}{2}\right) \right)$$

이다. 여기서 $\widetilde{x_{i-1}} \in [x_{i-1}, x_i]$, $\widetilde{\widetilde{x_{i-1}}} \in [x_{i-1}, x_i]$ 이라 할 때

$$\zeta(x_{i-1}+h) = \zeta(x_{i-1}) + h\zeta'(x_{i-1}) + \frac{h^2}{2}\zeta''(\widetilde{x_{i-1}}),$$

$$\zeta\left(x_{i-1} + \frac{h}{2}\right) = \zeta(x_{i-1}) + \frac{h}{2}\zeta'(x_{i-1}) + \frac{h^2}{8}\zeta''(\widetilde{\widetilde{x_{i-1}}})$$

이 됨을 알 수 있다. $\zeta''(x) = (1-x)^r/r!$ 이므로

$$\begin{aligned} &\zeta(x_{i-1}) + \zeta(x_{i-1}+h) - 2\zeta\left(x_{i-1} + \frac{h}{2}\right) \\ &= \frac{h^2}{4r!}\left(2(1-\widetilde{x_{i-1}})^r - (1-\widetilde{\widetilde{x_{i-1}}})^r\right) \\ &\geq \frac{h^2}{4r!}\left(2(1-x_i)^r - (1-x_{i-1})^r\right) \end{aligned}$$

이 되고, 따라서

$$\begin{aligned} S(f) &= \frac{h^2}{4r!}\left(2\sum_{i=1}^m (1-x_i)^r - \sum_{i=1}^{m-1} (1-x_i)^r\right) \\ &= \frac{h^2}{4r!}\left(\sum_{i=1}^m (1-x_i)^r - 1\right) \quad = \frac{h^2}{4r!}\left(\sum_{j=0}^{m-1} x_j^r - 1\right) \\ &\geq \frac{h}{4r!}\left(\int_0^{1-h} x^r\,dx - h\right) = \frac{h}{4r!}\left(\frac{(1-h)^{r+1}}{r+1} - h\right) \\ &= \frac{h}{4(r+1)!}\,(1 + \Theta(h)) \end{aligned}$$

이다.

■

IV. 결론

Infinite dimension인 r-fold Wiener process를 finitely dimensional subspaces만 취급할 수 있는 컴퓨터에서 구현하기 위해서는 process f 의 approximate random process f_m 를 생성하여 이 근사함수를 이용하여 문제를 풀어야 한다. 함수를 취급하는 문제에 대한 average case 오차는 일차적으로는 information M_m 에 따라 변화하고, 이차적으로는 사용하는 algorithm ϕ 에 따라 그 값이 변한다. 본 논문에서는 m 개의 함수값을 information으로 사용한다고 가정할 때, 두 가지의 information $M_m^*(f)$ 와 $M_m^{st}(f)$ 를 사용하면 Theorem1과 Theorem2에서 average case 오차가 최소화된다는 것을 보였다. 또한 Theorem3에서는 경우에 따라서는 $M_m^*(f)$ 와 $M_m^{st}(f)$ 대신에 $M_m^{(r)}(f)$ 를 사용하여 average case 오차를 최소화할 수 있다는 것을 보였다.

앞으로의 과제는 위와 같은 특성을 가진 information을 생성하는 프로그램을 작성하여 L_2-approximation 문제나 적분문제에 대한 average case 오차를 실험적으로 계산하는 것이다.

참고문헌

〔1〕 S. Choi and B. Hong, "An Effor of Simpson's Quadrature in the Average Case Setting," J. of Korean Math. Soc. 33 (1996), 235-247.

〔2〕 H. H. Kuo, "Gaussian Measures in Banach Spaces," Lecture Notes in Mathematics 463, Springer-Verlag, Berlin, 1975.

〔3〕 L. Plaskota, "Average Approximation and Integration in the Wiener Space with Noisy Data," J. of Complexity 8 (1992), 301-323.

〔4〕 K. Ritter, G. W. Wasilkowski, and H. Wozniakowski, "On Multivariage Integration for Stochastic Processes," International Series of Numerical Mathematics 112 (1993), 331-347.

〔5〕 J. F. Traub, G. W. Wasilkowski, and H. Wozniakowski, "Information-Based Complexity," Academic Press, New York, 1988.

저 자 소 개

최 성 희

1977: 서강대학교 수학과 이학사.
1988: 미국 Pennsylvania State Univ. 전자계산학과 이학석사.
1994: 미국 Univ. of Kentucky 전자계산학과 이학박사
현 재: 선문대학교 컴퓨터공학과 교수
관심분야: 수치해석, 알고리즘
Email : shchoi@sunmoon.ac.kr

황 석 형

1991: 강원대학교 전자계산학과 공학사.
1994: 일본 오사카대학교 정보공학과 공학석사.
1997: 일본 오사카대학교 정보공학과 공학박사
현 재: 선문대학교 컴퓨터공학과 교수
관심분야: 소프트웨어공학, 온톨로지공학, 시멘틱 웹, 빅데이터마이닝
Email : shwang@sunmoon.ac.kr