

유전 알고리즘의 확률 미분방정식에 의한 동역학 분석에 대한 연구

A Study on the Dynamics of Genetic Algorithm Based on Stochastic Differential Equation

석 진 육, 조 성 원

Jinwuk Seok, Seongwon Cho

홍익대학교 전자·전기 공학부

School of Electronic and Electric Engineering Hong-Ik University

Abstract

Recently, the genetic algorithm has been applied to the various types of optimization problems and these attempts have been very successfully. However, in most cases in these approaches, there is not given by investigator about to the theoretical analysis. The reason that the analysis of the dynamics for genetic algorithm is not clear, is the probabilistic aspect of genetic algorithm. In this paper, we investigate the analysis of the internal dynamics for genetic algorithm using stochastic differential method. In addition, we provide a new genetic algorithm, based on the study of the convergence property for the genetic algorithm.

I. 서 론

최적화 기법중의 하나로 많은 관심을 모으고 있는 유전 알고리즘은 현재까지 알고리즘의 내부 동역학에 대한 명확한 해석이 주어지지 않았다. 유전알고리즘의 내부 동역학이 명확하게 해석되지 않은 가장 큰 이유는 알고리즘의 확률적 성격 때문으로, 기존의 동역학 해석 방법론으로는 확률적 성격이 강한 유전 알고리즘의 해석이 어렵기 때문이다. 본 논문에서는 확률미분방정식에 의한 유전 알고리즘의 동역학을 해석하여 알고리즘의 수렴성과 안정성 조건을 밝히며, 이에 기초한 새로운 유전알고리즘을 제안한다.

II. 유전알고리즘

유전알고리즘은 1970년대 미국의 John Holland 교수에 의해 정립된 이론으로서 자연의 유전학과 자연선택의 원리에 근거한 최적해 탐색 방법이다[1]. 기존의 최적해 탐색기법이 국부탐색을 하는데 반해 유전 알고리즘은 여러해를 동시에 찾는 탐색방법을 채택함으로서 전역적인 해를 찾을 확률이 기존의 최적화 탐색방법들에 비하여 큰 것이 특징이다. 일반적인 유전 알고리즘은 이진 부호화 기법에 의해 생물과 같이 복제(Reproduction), 교배(Crossover), 돌연변이(Mutation)를 거쳐 다음세대의 자손을 만들어 낸다.

2.1 유전 알고리즘의 미분 방정식 표현

유전 알고리즘을 통해 각 개체(Individual)가 나타내는 실 공간상의 어떤 값을 x 라 하면 이 값은 유전 알고리즘의 특성상 다음과 같은 확률 미분 방정식의 형태로 나타낼 수 있다.

$$dX_{r,i}^{x,\sigma}(w) = a(S(t), X_{r,i}^{x,\sigma}(w), t)dt + b(t, \gamma)\sigma(t)dW_t \quad (1)$$

식에서 $X_{r,i}^{x,\sigma}(w)$ 시간 Index t에서 단위 스트링 혹은 개별 개체(Individual)가 표현하는 스칼라 혹은 벡터로서 표현되는 대상 공간이 R^n 일 때, $X_{r,i}^{x,\sigma}(w) \in R^n$ 이다. $X_{r,i}^{x,\sigma}(w)$ 는 랜덤 변수 w 에 의해 선택되며 평균

$$EX_{r,i}^{x,\sigma}(w) = \int_{\mathbb{R}[0,1]} X_{r,i}^{x,\sigma}(w) dP(w) \text{ 주변에 분포하}$$

며 Differential Wiener Process dW_t 의 계수가 $\sigma(t)$ 임을 나타낸다. r은 개별 개체를 나타내는 첨자(Index)이다. $a(S(t), X_{r,i}^{x,\sigma}(w), t)$ 는 유전 알고리즘의 교배연산자에 의해 나타나는 $X_{r,i}^{x,\sigma}(w)$ 의 변화값을 나타내는 함수로서 $X_{r,i}^{x,\sigma}(w)$ 와 Population $S(t)$ 중의 하나의 Individual과 교배연산을 하게된다. 랜덤 변수 w 는 Evaluation에 의해 나타나는 값으로 $X_{r,i}^{x,\sigma}(w)$ 가 다음세대로 복제 혹은 사상(Dead)을 결정한다. 즉, w 값이 클수록 복제 과정에서 $X_{r,i}^{x,\sigma}(w)$ 가 생존 혹은 복제될 가능성이 크며 낮을수록 사상될 가능성이 크다. 이를 수식으로 표현하면 다음과 같다.

$$\begin{cases} X_{r,i+1}^{x,\sigma}(w) = X_{r,i}^{x,\sigma}(w) & \text{for some } w > w_0 \\ X_{r,i+1}^{x,\sigma}(w) = X_{k,i}^{x,\sigma}(w) & \text{for some } w < w_0 \end{cases} \quad (2)$$

where $r \neq k$

$b(t, \gamma)\sigma(t)dW_t$ 는 돌연변이(Mutation)의 영향을 나타내기 위한 항으로 $b(t, \gamma)$ 는 돌연변이된 Bit가

$X_{r,t}^{x,\sigma}(w)$ 의 크기에 미치는 영향을 나타내는 함수이며 시간 Index t와 Bit에 대한 Index γ 의 함수로 나타난다. $\sigma(t)$ 는 돌연변이율과 Annealing Scheduling에 해당하는 항이다. 마지막으로 dW_t 는 Wiener Process의 Differential이다.

따라서 최적화 하고자 하는 목적함수가 $U(x, t) \in C^2$ 이면 Ito-Calculus에 의해 $X_{r,t}^{x,\sigma}(w)$ 과 t에 대하여 다음의 미분 방정식을 만족한다[2].

$$\begin{aligned} dU(X_{r,t}^{x,\sigma}(w), t) = & \left\{ \frac{\partial U}{\partial t}(X_{r,t}^{x,\sigma}(w), s) \right. \\ & + a(S(t), X_{r,t}^{x,\sigma}(w), t) \frac{\partial U}{\partial x}(X_{r,t}^{x,\sigma}(w), s) \quad (3) \\ & + \frac{1}{2} [b(t, \gamma)\sigma(t)]^2 \frac{\partial^2 U}{\partial x^2}(X_{r,t}^{x,\sigma}(w), s) \Big\} dt \\ & + b(t, \gamma)\sigma(t) \frac{\partial U}{\partial x}(X_{r,t}^{x,\sigma}(w), s) dW_t \end{aligned}$$

2.2 Schema Theorem에 의한 Steady State Genetic Algorithm 해석

Steady State Genetic Algorithm (SSGA)는 Elite 선택법을 사용하는 유전 알고리즘으로서 국소 최소점에 다른 유전 알고리즈다 빠르게 수렴하여 구현하기 편리한 장점이 있다. SSGA의 동역학을 해석하기 위해 String의 길이가 k인 다음의 Schema를 가정하자.

$$S_1^{(0)} = \{ * * * * \cdot \cdot \cdot * 1 \}$$

이때 Reproduction과 Recombine과정을 거쳐 나올 수 있는 다음 세대의 Schema를 또한 다음과 같이 놓자.

$$S_1^{(1)} = \{ * * * * \cdot \cdot \cdot 0 1 \}$$

이렇게 하면 $S_1^{(0)}$ 와 $S_1^{(1)}$ 의 차이는 $S_1^{(0)}$ 중 하나의 비트가 유전 알고리즘 연산 후 결정된 것을 의미한다. 이러한 방식으로 k개의 Schema를 다음과 같이 가정한다.

$$\begin{aligned} S_1^{(0)} &= \{ * * * \cdot \cdot \cdot * * 1 \} \\ S_1^{(1)} &= \{ * * * \cdot \cdot \cdot * 0 1 \} \\ S_1^{(2)} &= \{ * * * \cdot \cdot \cdot 1 0 1 \} \\ &\vdots \\ S_1^{(k-1)} &= \{ 1 1 0 \cdot \cdot \cdot 1 0 1 \} \end{aligned}$$

Schema Theorem[4][5]에 의해 SSGA에서 $S_1^{(t)}$ 의 다음 세대에서의 개수는 식 (4)와 같다.

$$\begin{aligned} n(S_1^{(t)}, t+1) \geq & n(S_1^{(t)}, t) \\ & \cdot \{1 + u(\mathcal{f}(S_1^{(t)})/\mathcal{f}(P) - \max_s \mathcal{f}(S_1))/\mathcal{f}(P)\} \\ & \cdot (1 - P_c[I(S_1^{(t)}, t)/(k-1)])(1 - \delta(S_1^{(t)}, S)) \\ & - o(S_1^{(t)})p_m \end{aligned} \quad (4)$$

식 (4)에서 $n(S_1^{(t)}, t)$ 은 $S_1^{(t)}$ 의 t 세대 때의 개수를 의미하고 $\mathcal{f}(P)$ 는 전체 개체에 대한 평균 적합도 (Fitness)를 의미하며 $\mathcal{f}(S_1^{(t)})$ 는 개체 $S_1^{(t)}$ 에 대한 평균 평균 적합도를 의미한다. $u(x)$ 는 단위 Step

Function을 의미하며 $x \geq 0$ 이면 $u(x) = 1$ 이며 $x < 0$ 이면 $u(x) = 0$ 이다. P_c 는 교배 확률이며

$I(S_1^{(t)}, t)$ 은 개체 $S_1^{(t)}$ 의 Defining 길이이다.

$\delta(S_1^{(t)}, S)$ 는 Delta-Dirac 함수로서 $S_1^{(t)} = S$ 이면 1, 그렇지 않으면 0이다. $o(S_1^{(t)})$ 는 개체 $S_1^{(t)}$ 의 0 혹은 1의 개수이며 p_m 은 돌연변이 확률이다.

식 (4)에서 $S_1^{(0)}$ 에서 $S_1^{(1)}$ 로의 전이(Transition)를 살펴보면 $I(S_1^{(0)}, t) = 0$ 이고 $o(S_1^{(0)})$ 이므로 식 (4)는 다음과 같이 된다.

$$\begin{aligned} n(S_1^{(0)}, t+1) \geq & n(S_1^{(0)}, t) \\ & \cdot \{1 + u(\mathcal{f}(S_1^{(0)})/\mathcal{f}(P) - \max_s \mathcal{f}(S_1)/\mathcal{f}(P))\} \\ & \cdot (1 - p_m) \end{aligned} \quad (5)$$

식 (5)에서 $n(S_1^{(0)}, t)$ 을 살펴보면, SSGA의 특성상 $\mathcal{f}(S_1^{(0)})/\mathcal{f}(P) - \max_s \mathcal{f}(S_1)/\mathcal{f}(P)) = \delta(S_1^{(0)}, S_1)$ 이다. 따라서 $S_1^{(0)}$ 이 Elite로서 선택될 확률을 $P(S_1^{(0)})$ 라 놓으면 식 (5)는 다음과 같이 놓을 수 있다.

$$\begin{aligned} E_{S^{(0)}} n(S_1^{(0)}, t+1) \geq & \int_{S^{(0)}} (1 + \delta(S_1^{(0)}, S_1))(1 - p_m) \\ & \cdot n(S_1^{(0)}, t) dP(S^{(0)}) \\ = & (1 - p_m) \int_{S^{(0)}} (1 + \delta(S_1^{(0)}, S_1)) n(S_1^{(0)}, t) dP(S^{(0)}) \\ = & (1 - p_m) (E_{S^{(0)}} n(S_1^{(0)}, t) + P(S_1^{(0)}) n(S_1^{(0)}, t)) \end{aligned} \quad (6)$$

식 (6)에서 $P(S_1^{(0)})n(S_1^{(0)}, t) \geq \epsilon \cdot E_{S^{(0)}} n(S_1^{(0)}, t)$ 을 만족하는 임의의 양수 ϵ 이 존재한다고 가정하면 식 (6)은 다음과 같다.

$$E_{S^{(0)}} n(S_1^{(0)}, t+1) \geq (1 - p_m)(1 + \epsilon) E_{S^{(0)}} n(S_1^{(0)}, t) \quad (7)$$

따라서, 전체 개체수를 m이라 했을 때, $E_{S^{(0)}} n(S_1^{(0)}, t)$ 이 임의의 $t > t_0$ 에 대하여 $E_{S^{(0)}} n(S_1^{(0)}, t) = m$ 이기 위해서는 $(1 - p_m)(1 + \epsilon) > 1$ 이어야 하므로 돌연변이 확률 p_m 은 $S_1^{(0)}$ 에 대하여 최소한 다음의 상계(Upper Bound)를 가져야 한다.

$$p_m < \frac{\epsilon}{1 + \epsilon} \quad (8)$$

이와 같은 개념을 l Bit의 Definite Bit를 가지는 Schema $S_1^{(t)}$ 에 대하여 확대하면 식 (9)와 같다.

$$\begin{aligned} \int_S n(S_1^{(t)}, t+1) dP(S_1^{(t)}) \geq & \int_{S^{(t)}} n(S_1^{(t)}, t) \{1 + u(\mathcal{f}(S_1^{(t)})/\mathcal{f}(P) - \max_s \mathcal{f}(S_1)/\mathcal{f}(P))\} \\ & \cdot (1 - P_c[I(S_1^{(t)}, t)/(k-1)])(1 - \delta(S_1^{(t)}, S)) \\ & - o(S_1^{(t)})p_m dP(S_1^{(t)}) \end{aligned}$$

$$\begin{aligned}
&= \int_{S_1^{(0)}} n(S_1^{(0)}, t) dP(S^{(0)}) \\
&+ \int_{S_1^{(0)}} n(S_1^{(0)}, t) u(f(S_1^{(0)})/f(P) - \max s(S_1)) f(P) dP(S^{(0)}) \\
&- \int_{S_1^{(0)}} n(S_1^{(0)}, t) P_c [l/(k-1)] (1 - \delta(S_1^{(0)}, S)) \\
&\quad + o(S_1^{(0)}) p_m dP(S^{(0)}) \\
&- \int_{S_1^{(0)}} n(S_1^{(0)}, t) u(f(S_1^{(0)})/f(P) - \max s(S_1)) f(P) \\
&\cdot P_c [l/(k-1)] (1 - \delta(S_1^{(0)}, S)) + o(S_1^{(0)}) p_m dP(S^{(0)})
\end{aligned} \tag{9}$$

식 (9)를 정리하면,

$$\begin{aligned}
E_{S^{(0)}} n(S_1^{(0)}, t+1) &\geq E_{S^{(0)}} n(S_1^{(0)}, t) \\
&+ \int_{S_1^{(0)}} n(S_1^{(0)}, t) \delta(S_1^{(0)}, S) dP(S^{(0)}) \\
&- \int_{S_1^{(0)}} n(S_1^{(0)}, t) P_c [l/(k-1)] (1 - \delta(S_1^{(0)}, S)) \\
&\quad + o(S_1^{(0)}) p_m dP(S^{(0)}) \\
&- \int_{S_1^{(0)}} n(S_1^{(0)}, t) \delta(S_1^{(0)}, S) \cdot (P_c [l/(k-1)] (1 - \delta(S_1^{(0)}, S)) \\
&\quad + o(S_1^{(0)}) p_m dP(S^{(0)})
\end{aligned} \tag{10}$$

(10)

그러므로

$$\begin{aligned}
E_{S^{(0)}} n(S_1^{(0)}, t+1) &\geq E_{S^{(0)}} n(S_1^{(0)}, t) + n(S_1^{(0)}, t) P(S_1^{(0)}) \\
&- P_c [l/(k-1)] E_{S^{(0)}} n(S_1^{(0)}, t) + P_c [l/(k-1)] \\
&\cdot n(S_1^{(0)}, t) P(S_1^{(0)}) \\
&- (l+1) p_m E_{S^{(0)}} n(S_1^{(0)}, t) - (l+1) p_m n(S_1^{(0)}, t) P(S_1^{(0)}) \\
\\
&= E_{S^{(0)}} n(S_1^{(0)}, t) (1 - (l+1) p_m) - P_c [l/(k-1)] \\
&+ n(S_1^{(0)}, t) P(S_1^{(0)}) (1 - (l+1) p_m + P_c [l/(k-1)]) \\
&= (1 - (l+1) p_m) (E_{S^{(0)}} n(S_1^{(0)}, t) (1 - P_c [l/(k-1)]) \\
&+ n(S_1^{(0)}, t) P(S_1^{(0)}) (1 + P_c [l/(k-1)]))
\end{aligned} \tag{11}$$

$S_1^{(0)}$ 의 경우와 마찬가지로

$P(S_1^{(0)}) n(S_1^{(0)}, t) \geq \epsilon \cdot E_{S^{(0)}} n(S_1^{(0)}, t)$ 을 만족하는 임의의 양수 ϵ 이 존재한다고 가정하면 식 (11)은 다음과 같이 정리된다.

$$E_{S^{(0)}} n(S_1^{(0)}, t+1) \geq (1 - (l+1) p_m) ((1 - P_c [l/(k-1)]) + \epsilon \cdot (1 + P_c [l/(k-1)])) E_{S^{(0)}} n(S_1^{(0)}, t) \tag{12}$$

따라서, 임의의 l Bit의 Definite Bit를 가지는 Schema $S_1^{(0)}$ 에 대하여 Schema $S_1^{(0)}$ 이 수렴하기 위해서는 $E n(S_1^{(0)}, t) = m$ 이어야 하므로 식 (12)의 첫 번째 항에서 다음의 상계를 구할 수 있다.

$$p_m < \frac{1}{l+1} \tag{13}$$

또한,

$$(1 - (l+1) p_m) ((1 - P_c [l/(k-1)]) + \epsilon \cdot (1 + P_c [l/(k-1)])) > 1$$

을 만족하여야 하므로 이를 통해 Schema $S_1^{(0)}$ 의 수렴조건을 유도하면 식 (14)와 같다.

$$p_m < \frac{1}{l+1} \cdot \frac{\epsilon \cdot (1 + P_c [l/(k-1)]) - P_c [l/(k-1)]}{1 + \epsilon \cdot (1 + P_c [l/(k-1)]) - P_c [l/(k-1)]} \tag{14}$$

돌연변이 확률 p_m 은 양수인 제한조건이 있으므로 $\epsilon > \frac{P_c [l/(k-1)]}{1 + P_c [l/(k-1)]}$ 이 만족되도록 선택한다.

III. Weak Convergence와 확률 미분방정식에 의한 유전 알고리즘의 수렴성 해석

식 (1)에서 $b(t, y) = 1$ 로 놓으면, 유전 알고리즘을 기술하는 동역학 방정식은 다음과 같이 단순화 된다.

$$dX_{r,t}^{x,\sigma}(w) = a(S(t), X_{r,t}^{x,\sigma}(w), t) dt + \sigma(t) dW_t \tag{15}$$

식 (15)와 같이 확률 미분방정식이 주어졌을 때 시간 s 일 때 상태 x 에서 시간 t 일 때 상태 y 로의 상태전이 확률 $p(s, x, t, y)$ 는 식 (16)과 같은 Fokker-Plank 방정식을 만족한다[6].

$$\begin{aligned}
\frac{\partial p}{\partial t} &= -a(S(t), X_{r,t}) \frac{\partial p}{\partial x} + \frac{1}{2} \sigma^2(t) \frac{\partial^2 p}{\partial x^2} \\
\therefore p(s, x, t, y) &:= p
\end{aligned} \tag{16}$$

식 (16)에서 $a(S(t), X_{r,t}) \in L^2$ 이고 최적화 하고자 하는 목적함수의 Gradient $\nabla_x U$ 와 Involutive Distribution $\Delta = \text{span}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ 의 벡터필드 λ_k 와 식 (17)과 같이 선형 결합에 의해 구성될 수 있다고 가정한다.

$$a(S(t), X_{r,t}) = \sum_{i=1}^n \alpha_i \lambda_i^T \nabla U \tag{17}$$

따라서 확률미분 방정식의 1 차 미분항의 계수 $a(S(t), X_{r,t})$ 는 목적함수의 Gradient $\nabla_x U$ 의 함수로서 목적함수 U 의 어떤 변화율의 함수로서 볼 수 있다. 이 함수를 ϕ 라 놓고 다음이 만족되도록 $\nabla_x \phi$ 를 놓는다[7].

$$\nabla_x \phi = \sum_{i=1}^n \alpha_i \lambda_i^T \nabla U = a(S(t), X_{r,t}) \tag{18}$$

식 (18)을 식 (16)에 대입하면 다음과 같은 Fokker-Plank 방정식을 얻게된다.

$$\frac{\partial p}{\partial t} = -\nabla_x \phi(x) \frac{\partial p}{\partial x} + \frac{1}{2} \sigma^2(t) \frac{\partial^2 p}{\partial x^2} \tag{19}$$

식 (19)에서 초기시간 t_0 에 대해 모든 상태가 0상태에 있다는 극한 가정을 두면, 식 (19)의 해는 $t \rightarrow \infty$ 에 대하여 다음과 같이 구해진다[3][8].

$$\pi_r(x) = \exp(-\frac{\phi(x)}{\sigma^2(t)}) \quad \text{as } t \rightarrow \infty \tag{20}$$

따라서 유전 알고리즘의 일반적인 수렴성 해석을 위해서는 SSGA알고리즘의 해석과는 달리, 상태전이 확률의 분포가 식 (20)에서 나타나는 분포함수로 수렴한다는 것을 증명해야 한다[9][10].

Definition 1. Fokker-Plank 방정식에 대한 Infinitesimal Operator L_t^σ 를 다음과 같이 정의한다[6].

$$L_t^\sigma = -\nabla_x \phi(x) \frac{\partial}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2} \quad (21)$$

Definition 2. Infinitesimal Operator L_t^σ 는 $p(\cdot) \in L[0,1]$ 인 임의의 함수 $p(\cdot)$ 에 대하여 다음과 같이 Quasi-Norm을 정의한다[6].

$$L_t = \|L_t^\sigma\| := \sup_{t_0 \leq t \leq t_0 + \epsilon} L_t^\sigma p$$

Infinitesimal Operator L_t^σ 를 사용하여 Fokker-Plank 방정식을 다시쓰면 다음과 같다.

$$\frac{\partial p}{\partial t} = L_t^\sigma p \quad (22)$$

Definition 3. Infinitesimal Operator L_t^σ 는 $p(\cdot) \in L[0,1]$ 인 임의의 함수 $p(\cdot)$ 에 대하여 다음과 같이 Quasi-Norm을 정의한다[6]

$$L_t = \sup_{t_0 - \epsilon \leq s \leq t + \epsilon} L_s^\sigma p \quad \forall p \in L[0,1]$$

Definition 1, 2, 3에 의해 식 (22)는 다음의 부등식을 만족한다.

$$\frac{\partial p}{\partial t} = L_t^\sigma p \leq L_t p \quad (23)$$

양변에 $\exp(-\int_{t_0}^t L_s ds)$ 를 곱하여 정리한다.

$$\frac{\partial p}{\partial t} \exp(-\int_{t_0}^t L_s ds) - L_t^\sigma p \exp(-\int_{t_0}^t L_s ds) \leq 0 \quad (24)$$

$$\frac{\partial}{\partial t} \left\{ p \exp(-\int_{t_0}^t L_s ds) \right\} \leq 0$$

식 (24)를 Lebesgue 적분하면

$$\begin{aligned} & \int_{t_0}^{t_0 + \epsilon} \partial p \exp(-\int_{t_0}^t L_s ds) \\ & \Leftrightarrow p_t \exp(-\int_{t_0}^t L_s ds) \leq p_{t_0} \Leftrightarrow p_t \leq p_{t_0} \exp(\int_{t_0}^t L_s ds) \end{aligned} \quad (25)$$

식 (25)의 양변에 최종 극한 상태 f 에 대한 극한 전이확률 $\pi_t^{\sigma(t)}(f)$ 를 빼고 Norm을 취하면 다음의 관계식이 성립한다.

$$\begin{aligned} & p_t - \pi_t^{\sigma(t)}(f) \leq p_{t_0} \exp(\int_{t_0}^t L_s ds) - \pi_t^{\sigma(t)}(f) \\ & \Leftrightarrow p_t - \pi_t^{\sigma(t)}(f) \leq p_{t_0} \exp(\int_{t_0}^t L_s ds) \\ & \Leftrightarrow \|p_t - \pi_t^{\sigma(t)}(f)\| \leq p_{t_0} \exp(\int_{t_0}^t L_s ds) \end{aligned} \quad (26)$$

L_s 의 Eigenvalue를 $\lambda_{L_s}(\sigma(t))$ 라 하고 t_0 를 0으로 놓으면 식 (26)은 다음과 같이 정리된다.

$$\|p_t - \pi_t^{\sigma(t)}(f)\| \leq p_{t_0} \exp(t \lambda_{L_s}(\sigma(t))) \quad (27)$$

만일 $\lim_{\sigma(t) \rightarrow 0} \sigma^2(t) \log(-\lambda_{L_s}(\sigma(t))) = -c_1$ 이면, $c > c_1$

예 대하여 $c > c_1 + a$ 로서 $\lambda_{L_s}(\sigma(t)) \geq \exp(-(c_1 + a)/\sigma^2(t))$ 를 만족하는 c_1 을 매우 작은 $\sigma^2(t)$ 에 대하여 취한다. 특히 $\sigma^2(t) = c/\log t$ 로 잡으면 식 (27)은 다음과 같이 약수렴 함을 알 수 있다.

$$\|p_t - \pi_t^{\sigma(t)}(f)\| \leq p_{t_0} \exp(-t^{1-(c_1+a)/c}) \rightarrow 0 \quad (28)$$

따라서 유전 알고리즘의 돌연변이 확률은 $\sigma^2(t) = c/\log t$ 의 형태로 시간에 대한 단조감소 함수로 높아야 함을 알 수 있다.

IV. 시간에 대해 단조감소하는 돌연변이 확률과 의사 Elite선택법을 사용하는 유전 알고리즘

일반적인 유전 알고리즘의 수렴성 해석을 통해 유전 알고리즘은 SSGA 알고리즘을 제외하고 전이확률의 분포함수의 수렴성만을 보장할 수 있음을 보였다. 이와같은 유전 알고리즘의 특성과 Elite선택법의 특성을 사용할 경우 유전 알고리즘의 수렴성과 탐색 성능을 향상시킬 수 있음을 알 수 있다.

유전 알고리즘에서 사용되는 개체군(Population)을 $P(t)$ 라 할 경우 본 논문에서 제안하는 알고리즘은 다음과 같다.

```

begin
  t=0
  Initialize P(t)
  Evaluate P(t)
  while(not termination condition) do
    P(t) → P(t+1)
    begin contractive mapping
      t=t+1
      select P(t) from P(t-1)
      recombine P(t)
      evaluate P(t)
      if rand() > exp(-eval P(t-1) - eval P(t)) / σ²(t)
        then t=t-1
      end
end

```

이 때 Mutation 확률은 다음과 같다.

$$p_m = \frac{c}{K \cdot \log t + 1} \quad (29)$$

Mutation 확률은 개체군 선택을 위한 함수 $\exp(-(eval P(t) - eval P(t-1))/\sigma^2(t))$ 의 $\sigma^2(t)$ 와 같은 값을 사용한다. 또한 각 개체들의 수렴성을 높여 SSGA의 특성들을 알고리즘내에 가지고도록 하기 위하여 적합도(Fitness)가 가장 높은 두 개의 개체에 대한 Fitness값을 2배씩 증가시킨다. 이로서, Schema Theorem에서 교배와 돌연변이를 제외했을 경우 적합도가 높은 개체가 재생산(Reproduction)될 확률을 높임으로서 각 개체들의 수렴이 보다 높은 확률로 이루어질 수 있도록 한다.

V. 실험결과

실험은 식 (30)으로 나타난 Scalar Filed 함수의 최대값을 찾는 문제에 대하여 행하였다.

$$f(x_1, x_2) = 21.5 + x_1 \cdot \sin(4\pi x_1) + x_2 \cdot \sin(20\pi x_2) \quad (30)$$

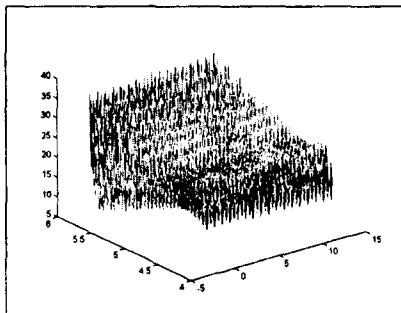


그림 1. 함수 $f(x_1, x_2)$ 의 그래프

각 개체의 길이는 10Bit로서 x_1, x_2 에 대하여 각각 5Bit씩 할당 하였다. 고배확률은 0.6이었으며 돌연변이 확률은 단순 유전 알고리즘(GA)과 SSGA, Contractive Mapping GA(CMGA) 그리고 Simulated Annealing(SAGA)이 결합된 GA의 경우 0.25였으며 제안된 GA의 경우에는 다음과 같이 높았다.

$$p_m = \frac{0.5}{\log t + 1}$$

개체수는 20개였으며 실험환경은 IBM-PC Pentium 100에서 행하였으며 각각 독립적으로 100~1000세대 까지 행하였다.

	제안된 GA	GA	SAGA	SSGA	CMGA
100	38.4430	37.4558	37.6442	38.4430	36.0857
200	38.3388	38.3388	38.4430	38.3388	38.4430
300	38.4430	36.1686	38.4430	38.4430	38.3388
400	38.4430	36.2529	38.3388	38.4430	38.4430
500	38.4430	36.3320	37.6442	38.3388	37.6442
600	38.4430	38.4430	38.3388	38.3388	37.0266
700	38.4430	35.8894	37.6442	37.6442	38.4430
800	38.4430	38.4430	38.3388	38.4430	38.4430
900	38.4430	37.6442	37.6442	38.4430	38.3388
1000	38.4430	38.4430	38.4430	38.4430	37.6442
평균	38.4326	37.3411	38.0922	38.3318	37.8850

표 1. 각 GA 알고리즘에 대한 목적함수 최대값 탐색 결과

실험결과 제안된 GA의 경우 안정적으로 가능한 최대값을 다른 알고리즘에 비해 잘 찾아줄 수 있음을 알 수 있다.

VI. 결론

본 논문에서는 유전 알고리즘을 확률 미분 방정식의 해석 방법을 사용하여 돌연변이 확률에 의한 수렴조건을 해석하였다. 또한 SSGA 알고리즘의 수렴조건을 Schema Theory를 사용하여 분석하였다. 수렴조건의 분석을 통해 나타난 결과와 SSGA 알고리즘이 가지는 특성을 결합하여 새로운 유전 알고리즘을 제안하였고 실험결과 다른 유전 알고리즘들에

비하여 안정적으로 목적함수의 최대값을 잘 찾아줄 수 있음을 알 수 있었다.

참고문헌

- [1] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, MA : Addison-Wesley, 1989.
- [2] P. E. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equation*, Springer-Verlag : Berlin, 1992.
- [3] I. I. Gihman A. V. Skorohod, *The Theory of Stochastic Process III*, Springer-Verlag : Berlin, 1975
- [4] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, 3rd-ED. 1992.
- [5] P. D. Wasserman, *Advanced Methods in Neural Computing*, Van Norstand Reinhold : New York, 1993.
- [6] 석 진욱, 조 성원, "Fokker-Plank 방정식의 해석을 통한 Langevine 경쟁학습의 동역학 분석", 대한전자공학회 논문지, 제 34권, C편, 제 7 호, 1997년 7월, pp. 82-91.
- [7] F. W. Warner, *Foundations of Differential Manifold and Lie Groups*, Scott, Freshman and Company : London, 1970.
- [8] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland Publishing Company : Amsterdam, 1981.
- [9] T. S. Chiang, C. R. Hwang, S. J. Sheu, "Diffusion for Global Optomization in R^n ", *SIAM J. Control and Optimization*, vol 25, No. 3, May 1987.
- [10] H. Kushner, A. Shwartz, "Weak Convergence and Asymptotic Properties of Adaptive Filters with Constant Gain", *IEEE trans. on Information Theory*, vol. IT-30, No. 2, March 1984.