

# 유전자 알고리즘을 이용한 Moving Average 의 최적 Period 예측 시스템 구현

김소영, 한치근  
경희대학교 전자계산학과  
e-mail : [cozycafe@algorithms.khu.ac.kr](mailto:cozycafe@algorithms.khu.ac.kr)

## A Genetic Algorithm for Optimal Period Forecasting Of Moving Average

So-Young Kim, Chi-Geun Han  
Dept. of Computer Science, Kyung-Hee University

### 요 약

주가지수선물시장은 주식투자에 따르는 위험을 효과적으로 관리할 수 있는 제도적 장치로서 오늘날 불안한 주식시장 현황에 있어서 더욱더 중요한 위치를 갖고 있다. 현재 이러한 주가지수선물거래에 있어서 Moving Average 를 예측하고자 하는 여러 트레이딩 시스템을 선보이고 있다. 이 논문에서는 과거의 데이터를 토대로 한 Moving Average Line 분석에 있어서 일반적으로 기존방법보다 효과적이라고 알려진 유전자 알고리즘을 이용하여 Moving Average 의 최적 Period 예측 시스템을 구현한다.

### 1. 서론

선물 트레이딩의 효율성 제고와 포트폴리오 관리의 개선을 위한 여러 진보된 기법들이 연구되어 오고 있다. 이에는 신경망, 유전자 알고리즘, 퍼지 로직, 카오스 이론 등이 포함된다.

이 중 유전자 알고리즘은 전통적으로 정형화된 최적화 방법으로 해를 찾는 것이 아니라 해를 진화시킴으로써 답을 찾는 기법이다: 유전자 알고리즘은 매우 광범위한 문제들의 해를 찾을 수 있는 최적화 알고리즘이며, 다른 최적화 알고리즘들과 혼합하여 사용되기도 한다. 유전자 알고리즘은 트레이딩 시스템의 input 최적화,

트레이딩 규칙들의 추론 트레이딩에 사용할 규칙들의 구성, 주어진 모형 집합들로부터의 모형 조합, 포트폴리오 선택 문제, 인덱스 펀드의 작성등에 사용된다. 본 논문에서는 유전자 알고리즘을 이용하여 최적의 Moving Average Line 을 구축하여 효율적인 예측 시스템을 구현한다.

### 2. 문제정의

포트폴리오 선택 문제는 본래 기대 수익의 최대화와 위험의 최소화라는 다목적 최적화 문제이고 그 해는 파레토 최적해로서 특징 지워진다. 포트폴리오 선택 문제에서는 파레토 최적해 집합을

효율적 투자선(efficient frontier)이라고 부르고 있다. 유전적 알고리즘에서는 현재 세대의 집합 중에서 비파레토 최적의 집합을 도태시키는 전략에 따라 세대 교체를 행하여 효율적 투자선을 용이하게 구할 수 있다.

종래의 많은 다목적 최적화는 목적함수의 1 차원화와 단목적 최적화의 반복에 의하여 해를 구하였지만 유전적 알고리즘은 복수의 해 후보 집단을 구성하여 파레토 최적해 집단을 직접 구한다는 점이 다르다.

소극적 투자전략의 일환으로 투자자들은 보유 포트폴리오의 시장 연동성 추구를 위해 인덱스 펀드의 구축이 필요하다는 것이 주지의 사실이다. 그러나 지금까지 인덱스 펀드 구축은 구성 종목 선정시 계산상의 방대함이나 빈번한 포트폴리오 재조정 문제로 인하여 대상 지수에 대한 연동 효율성 측면에서 많은 어려움이 있었다. 그러나 유전자 알고리즘은 이러한 문제에 대한 해결 가능성을 제시할 수 있다.

### 3. 유전자 알고리즘 적용

#### 3.1 유전자 알고리즘 기본 개념

유전자 알고리즘은 생물의 진화 법칙을 컴퓨터 알고리즘에 적용하여 해를 찾고자 하는 알고리즘이다. 여기서는 적자 생존의 법칙이 적용된다. 즉 해들 중에서 좀 더 나은 해는 다음 세대에 살아남고 그렇지 않은 해는 사라진다. 그리고 각각의 자손(다음 세대의 해)는 부모해로부터 재생성(reproduction), 교차(cross-over), 변이(mutation)의 과정을 거쳐 생성된다. 이러한 과정을 반복하여 의미 있는 값을 지니는 해를 찾거나 최선의 해를 찾는다.

유전자 알고리즘은 이론적으로나 경험적으로 복잡한 공간에서의 견고한 탐색을 제공하는 것으로 증명되었고, 복잡한 수식을 요구하지 않으며

탐색공간과 같은 한정적인 요소에 의해 제한 받지 않는다.

```

procedure genetic algorithm
begin
    t := 0;
    initialize P(t);
    evaluate P(t);
    while(not termination-condition) do
        begin
            t := t + 1;
            select P(t) from P(t-1);
            alter P(t) by crossover;
            alter P(t) by mutation;
            evaluate P(t);
        end
    return the best solutions;
end.
    
```

[유전자 알고리즘의 일반적인 구조]

#### 3.2 염색체 구성

유전자 알고리즘으로 시스템을 구현하기 위해서는 해를 표현하는 적절한 염색체(chromosome) 구성이 먼저 이루어져야 한다.

```

struct chromosome
{
    Decoding 정보    string'
    Moving average value;
    Profit_rate value;
    Loss_rate value;
    Fitness          value;
};
    
```

[염색체의 구성]

염색체는 유전자 알고리즘에서의 하나의 세대를

구성하는 많은 개체들 중의 하나의 개체를 나타내며, 본 논문에서는 염색체를 그림에서처럼 Decoding 정보, Moving average, Profit\_rate, Loss\_rate, Benefit, Fitness 로 구성하였다.

### 3.3 평가, 적합 함수 및 선택

염색체 구성이 끝나면 다음으로는 연산자를 적용하기 전 어떻게 부모를 선택할 것인가가 가장 중요한 문제가 된다. 이에 대해서는 대표적인 선택 방법인 품질 비례 룰렛휠 선택법을 사용한다. 이 방법은 각 해들은 각각 품질을 평가한 다음 가장 좋은 해의 적합도가 가장 나쁜 해의 적합도보다  $k$  배가 되도록 조절한다. 해집단 내의 해  $i$ 의 적합도는 다음과 같이 계산된다.

$$f_i = (C_w - C_i) + (C_w - C_b)/(k - 1), k > 1$$

- $C_w$ : 해집단 내에서 가장 나쁜 해의 비용
- $C_b$ : 해집단 내에서 가장 좋은 해의 비용
- $C_i$ : 해  $i$ 의 비용

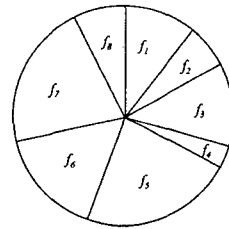
여기서  $k$  값을 높이면 선택압이 높아진다. 일반적으로 가장 흔히 쓰는  $k$  값은 3~4이다. 이 적합도 값을 기준으로 룰렛휠 선택을 하게 된다.

위와 같이 계산한 각 염색체의 적합도를 모두 합한 값 만큼의 크기를 가진 룰렛휠을 가정한다. 각 염색체는 이 룰렛휠 상에 자신의 적합도 만큼의 공간을 배정 받는다. 그림은 각 염색체가 적합도에 비례해서 룰렛휠 상의 공간을 배정 받은 모양의 예를 보인다. 여기에 활을 쏘면 각 염색체의 선택 확률은 배정된 공간의 크기에 비례하게 된다. 룰렛휠 선택은 아래와 같이 간단히 구현할 수 있다.

```
point = random() % SumOfFitnesses ;
sum = 0 ;
for i = 0 to N-1 {
```

```
sum = sum + fi ;
if (point < sum) return i ;
}
[룰렛휠 공간 배정]
```

여기서  $f_i$ 는 염색체  $i$ 의 적합도이고, SumOfFitnesses는 모든 염색체들의 적합도 값을 더한 수치이다.



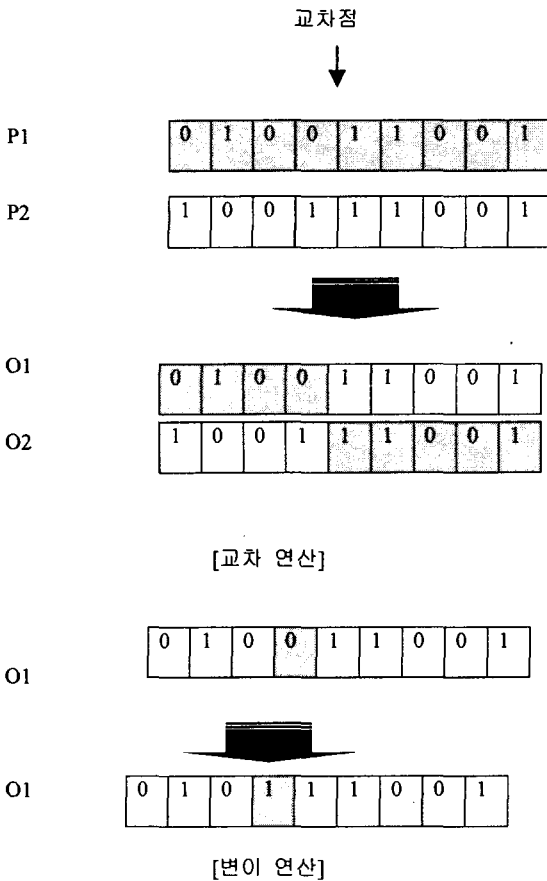
### 3.4 유전연산

교차 연산에서는 어떠한 부분을 교차할 것인가도 중요한데, 본 논문에서는 초기의 유전 알고리즘들은 대부분 사용하였고, 요즘에도 가장 많이 사용되는 교차 연산인 일점 교차를 사용했다. 길이가  $n$ 인 일차원 문자열로 된 염색체 상에서 일점 교차로 자르는 방법의 총 수는  $n-1$  가지이다.

비록 선택 연산과 교차 연산이 각각 효과적인 탐색과 재조합을 수행하였다고 하더라도 개체의 어떤 중요한 유전적 특징을 잃어 버렸을 수도 있다. 변이 연산은 이러한 회복할 수 없는 요소들을 극복할 수 있게 한다. 전형적인 변이 연산은 개체의 단일 위치값을 변경하는 경우로 설명되며 다음에 그 과정이 묘사되어 있다

참고문헌

- [1] G.Brassard and P. Bratley, "Algorithmics theory and Practice", PRENTICE HALL, 1988
- [2] H.Esbensen, Computing Near-Optimal Solutions to the Steiner Problem in a Graph Using a Genetic Algorithm, NETWORKS, Vol. 26 (1885), 173-185, 1995
- [3] P.Guitart and J.M. Basart, A Genetic Algorithm Approach for the Steiner Problem in Graphs, EURIT '98, 508-512, 1998
- [4] R. Sedgewick, "Algorithms in C++", Addison-Wesley, 1992
- [5] Y.Leung, G.Li, and Z.B. Xu, A Genetic Algorithm for the Multiple Destination Routing Problems, IEEE TRANS. ON EVOLUTIONARY COMPUTATION, VOL. 2, NO. 4 pp. 150-161, 1998
- [6] 강명주, 한치근, Rural Postman Problem 해법을 위한 Metaheuristic 알고리즘 비교 연구, 경희대학교 대학원 박사학위 논문, 1998
- [7] 정천용, 주가지수 선물거래 도입이 우리나라 주식시장에 미치는 영향에 관한 연구, 경희대학교 대학원 석사학위 논문, 1996
- [8] 서용만, 실시간 멀티캐스트 라우팅을 위한 유전자 알고리즘, 경희대학교 대학원 석사학위 논문, 2000



4. 결론 및 향후 연구방향

본 논문에서는 효과적인 Moving Average 예측 시스템을 위해 유전자 알고리즘 기법을 이용하였다. 기존 데이터를 토대로 적합한 해를 찾기 위한 개체의 염색체 구성과 유전 연산자를 사용하였다. 앞으로 다양한 시뮬레이션을 통해 기존의 알고리즘과의 비교를 통해 문제를 분석하고 적용하는 방안이 연구되어야 한다.