

기계학습을 이용한 파레토 프런티어의 생성[†]

윤예분¹ · 정나영² · 윤민³

¹간사이대학교 환경 도시공학부 · ²³부경대학교 통계학과

접수 2013년 3월 22일, 수정 2013년 4월 17일, 게재확정 2013년 5월 2일

요약

진화 알고리즘 계산 지능을 이용한 예측 방법이 다목적 최적화 문제에서 많이 이용되고 있고, 이러한 방법들은 많은 근사 파레토 최적해들을 좀 더 정확하게 생성하기 위해서 개선되고 있다. 본 논문은 다목적 최적화 문제에서 서포트 벡터기계를 이용하여 근사 파레토 프런티어를 찾는 방법을 제안한다. 또한 제안된 방법과 진화 알고리즘을 결합한 것이 파레토 프런티어를 더 잘 근사시킨다는 것과 두 개 혹은 세 개의 목적함수를 가진 의사결정은 제안된 방법으로 파레토 프런티어를 시각화한 것에 근거하여 더 쉽게 수행된다는 것을 보인다. 마지막으로 몇 개의 수치예제를 통해 제안된 방법의 효율성에 대해 보일 것이다.

주요용어: 다목적 최적화, 서포트 벡터기계, 진화연산, 파레토 프런티어.

1. 서론

많은 의사결정 문제는 다목적 최적화 문제 (multi-objective optimization problem; MOP)로 공식화할 수 있다. 모든 목적함수를 동시에 최적화 시키는 해가 반드시 존재하는 것은 아니다. 왜냐하면 다양한 목적이 서로 모순될 수 있기 때문이다. 그래서 파레토 최적해의 개념이 소개되었다. 일반적으로 많은 파레토 최적해가 존재하고 이것들은 최종적인 의사결정의 문제에서 해들의 후보로 고려된다. 의사결정자가 파레토 프런티어라 불리는 파레토 최적해들의 집합에서 어떻게 최종적으로 해를 결정하는지가 문제이다. 이러한 이유로 대화형 최적화 (interactive optimization) 방법들이 개발되었다 (Nakayama 등, 2009; Palli 등, 1998). 이 방법들은 다음의 두 단계를 반복적으로 수행함으로써 의사결정해들을 찾는다. 의사결정자의 목표수준 (aspiration level)에 가장 근접한 파레토 최적해를 얻기 위한 보조 최적화 문제를 해결하고, 트레이드-오프 (trade-off)분석으로 의사결정자의 목표수준을 수정한다.

하지만 많은 실제 공학 문제에 있어서, 블랙박스 목적함수의 형태들은 설계변수에 의해 정확하게 알 수 없으나, 이러한 함수값들은 표본의 실제적으로나 계산실험을 통하여 얻어진 값들로써 얻어지게 된다. 구조분석, 유체역학 그리고 열역학 등과 같은 여러 분야에서 목적함수들을 평가하기 위해 이용되었다. 일반적으로 이러한 분석은 비용이 많이 들고 아주 긴 실행시간이 소요된다. 따라서 의사결정자와 컴퓨터 사이의 상호작용에 의한 대화형으로 해를 구하는 것은 실제적으로 거의 불가능하다. 반면 최근에 진화연산과 같은 계산 지능을 이용한 근사 방법이 MOP를 해결하는데 이용되었다 (Coello Coello 등, 2002; Deb, 2001). 이러한 방법들은 많은 파레토 최적해들을 좀 더 정확하게 생성하기 위해 개선되고 있다.

[†] 이 논문은 2012학년도 부경대학교의 지원을 받아 수행된 연구임 (PK-2012-45).

¹ (564-8680) 일본 오사카 스이타시 3-3-35, 간사이대학교 환경 도시공학부, 교수.

² (608-737) 부산시 남구 용소로 45, 부경대학교 통계학과, 석사.

³ 교신저자 : (608-737) 부산시 남구 용소로 45, 부경대학교 통계학과, 부교수.

E-mail: myoon@pknu.ac.kr

본 논문은 MOP들에서 서포트 벡터기계 (Hwang과 Shim, 2012; Schölkopf 등, 2001; Steinwart와 Christmann, 2008)를 이용하여 근사 파레토 프런티어를 찾아내는 방법을 제안한다. 나아가 제안된 방법과 진화 알고리즘을 결합한 것이 파레토 프런티어를 더 잘 근사시킨다는 사실과 두 개 또는 세 개의 목적함수를 가지는 의사결정의 문제에서 제안된 방법을 이용하는 것이 더 쉽게 수행된다는 것을 파레토 프런티어를 시각화한 것에 근거하여 보인다. 마지막으로 수치예제를 통해 제안된 방법의 효율성에 대해 보인다. 본 논문은 다음과 같이 구성되었다. 서포트 벡터기계와 MOP에 대해 간단히 설명하고 서포트 벡터기계를 이용한 다목적 최적화 방법을 제안한 후 제안된 방법을 인공적, 실제 예제에 적용시키고 마지막으로 결론을 내린다.

2. 서포트 벡터기계

반응변수가 두 개의 범주를 갖는 패턴인식 문제에서 자료가 선형분리가 되지 않는 경우, 서포트 벡터 기계는 원 자료를 어떤 비선형 사상에 의해 고차원의 특징공간 (feature space)으로 변환한다. 특징공간으로 사상된 자료는 선형 분리 가능하고 서포트 벡터기계는 입력 공간에서는 비선형 결정 경계를 가지는 자료를 최대 마진을 가지고 선형으로 분리하는 초평면을 찾는다. 최근 서포트 벡터기계는 한 범주를 (혹은 분류되지 않는) 가지는 자료의 경우로 확장될 수 있고, 이 절에서 이와 같은 서포트 벡터기계를 간단하게 소개한다 (Schölkopf 등, 2001; Seok, 2010). 여기서 소개하는 알고리즘에서 사용된 서포트 벡터기계를 간단히 SVM이라고 표현하기로 한다. 표본자료 $\mathbf{x}_1, \dots, \mathbf{x}_\ell \in \mathcal{X}$ 를 고려하자. 여기서 $\ell (\in \mathbb{N})$ 은 관측치들의 수이고 \mathcal{X} 는 입력공간 상의 어떤 집합이다.

Φ 를 실제 입력 공간으로부터 고차원 특징공간으로 비선형 변환 사상이라 하자. 이 때 Φ 의 이미지에 서 내적은 커널을 이용하여 계산할 수 있다 (Park 등, 2012; Schölkopf와 Smola, 2002; Vapnik, 1995).

$$K(\mathbf{x}_i, \mathbf{x}_j) \equiv \langle \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) \rangle.$$

특히, 가우지안 (Gaussian) 커널함수는

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{2\sigma^2}\right)$$

와 같이 나타낼 수 있다. 커널함수를 사용함으로써 특징공간에 엄밀하게 사상시키지 않고도 분리 초평면을 계산하는 것이 가능하다.

SVM의 개념은 커널을 이용하여 고차원의 특징공간에 자료를 변환시켜 자료집합이 최대 마진을 가지도록 선형 분리시키는 것이다. 즉, 대부분의 자료 점들이 차지하고 있는 지역에서 양의 값을 가지고 나머지는 음수 값을 가지는 함수를 찾는 것이다. 새로운 점 \mathbf{x} 에 대해서, 특징공간에서 초평면의 어떤 위치에 있는지에 따라 함수 $h(\mathbf{x})$ 의 값이 평가된다.

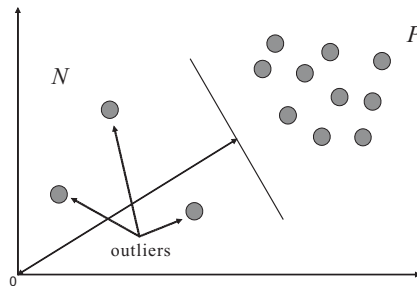


Figure 2.1 Geometric interpretation of support vector machines with one-category

SVM문제를 해결하는 의사결정 함수는 아래와 같이 표현될 수 있다.

$$h(\mathbf{x}) = \mathbf{w}^T \Phi(\mathbf{x}) - \rho. \quad (2.1)$$

특징공간 내에서 원점으로부터 자료를 분리하기 위하여, \mathbf{w} , ξ_i 그리고 ρ 에 대하여 아래의 이차계획문제 (SVM)를 해결해야 한다:

$$\begin{aligned} \underset{\mathbf{w}, \xi_i, \rho}{\text{minimize}} \quad & \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{1}{\nu \ell} \sum_{i=1}^{\ell} \xi_i - \rho & (\text{SVM}) \\ \text{subject to} \quad & \mathbf{w}^T \Phi(\mathbf{x}_i) \geq \rho - \xi_i, \\ & \xi_i \geq 0, \quad i = 1, \dots, \ell. \end{aligned}$$

여기서, $\nu \in (0, 1]$ 는 $\mathbf{w}^T \mathbf{w}$ 와 $\sum_{i=1}^{\ell} \xi_i$ 사이의 트레이드-오프를 나타내는 모수이다.

의사결정함수 h 의 값은 주어진 자료집합을 포함한 대부분의 표본자료 점 ξ 에서는 양의 값이 될 것이다. 또한 SVM문제의 목적함수에서 보여지는 바와 같이 ν 가 0에 접근하면 오차항 $\sum_{i=1}^{\ell} \xi_i$ 의 벌칙 (penalization)은 무한대가 된다. 이것은 모든 자료점들이 P 영역에 있다는 결과를 얻는다. 그러므로, 특징공간에서 모든 자료점들을 지지 (supporting)하는 초평면을 찾기 위해서 충분히 작은 양의 값 ν 를 취할 것이다. 예를 들면, 이후의 모의실험에서 소개되는 10^{-7} 과 같은 값이다.

SVM문제에 대하여 라그랑지안 함수를 \mathbf{w}, ρ, ξ_i 에 대해서 미분하면 아래와 같은 식들을 얻는다.

$$\mathbf{w} = \sum_i^{\ell} \alpha_i \Phi(\mathbf{x}_i), \quad (2.2)$$

$$\alpha_i = \frac{1}{\nu \ell} - \beta_i, \quad \sum_{i=1}^{\ell} \alpha_i = 1. \quad (2.3)$$

식 (2.2)와 (2.3)을 이용하면 라그랑지안 쌍대문제 (SVM_D)를 얻을 수 있다 :

$$\begin{aligned} \underset{\alpha_i}{\text{maximize}} \quad & -\frac{1}{2} \sum_{i,j=1}^{\ell} \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) & (\text{SVM}_D) \\ \text{subject to} \quad & \sum_{i=1}^{\ell} \alpha_i = 1, \\ & 0 \leq \alpha_i \leq \frac{1}{\nu \ell}, \quad i = 1, \dots, \ell. \end{aligned}$$

α^* 는 SVM_D에 대한 최적해라 하자. 그러면 의사결정함수(2.1)는 커널함수를 이용하여

$$h(\mathbf{x}) = \sum_{i=1}^{\ell} \alpha_i^* K(\mathbf{x}_i, \mathbf{x}) - \rho^*$$

으로 변환되고, 여기서 $0 < \alpha_j^* < \frac{1}{\nu \ell}$ 를 가지는 임의의 \mathbf{x}_j 에 대하여

$$\rho^* = \sum_{i=1}^{\ell} \alpha_i^* K(\mathbf{x}_i, \mathbf{x}_j)$$

이다. 여기서 자료점 $\mathbf{x} \in \{\alpha_i^* > 0 \text{ 를 가지는 } \mathbf{x}_i\}$ 를 서포트 벡터 (support vector)라고 부른다.

3. 서포트 벡터기계를 이용한 다목적 최적화 알고리즘

다음과 같은 MOP를 고려해보자.

$$\begin{aligned} & \underset{\mathbf{x}}{\text{minimize}} && \mathbf{f}(\mathbf{x}) := (g_1(\mathbf{x}), \dots, g_m(\mathbf{x}))^T && \text{(MOP)} \\ & \text{subject to} && \mathbf{x} \in S = \{ \mathbf{x} \mid g_j(\mathbf{x}) \leq 0, j = 1, \dots, n \}, \end{aligned}$$

여기서 $\mathbf{x} = (x_1, \dots, x_n)^T$ 는 설계변수의 벡터이고 S 는 모든 실행가능한 해들의 집합이라 하자.

일반적으로 단목적함수를 가진 전통적인 최적화 문제들과는 달리 MOP에서는 모든 목적함수 $g_i(\mathbf{x})$, $i = 1, \dots, m$,를 동시에 최소화시키는 해가 존재하는 경우는 드물다. 따라서 파레토 지배적인 관계 (Pareto domination)에 기초한 최적해의 개념이 다음과 같이 정의되었다.

정의 3.1 (파레토 최적해) 만일 $f_i(\mathbf{x}) \leq f_i(\hat{\mathbf{x}})$, $i = 1, \dots, m$ 그러나 $\mathbf{f}(\mathbf{x}) \neq \mathbf{f}(\hat{\mathbf{x}})$ 를 만족하는 $\mathbf{x} \in S$ 가 없다면, 점 $\hat{\mathbf{x}} \in S$ 은 MOP에서 파레토 최적해 (Pareto optimal solution)라고 부른다. 그리고 파레토 최적해의 집합을 파레토 프런티어 (Pareto frontier)라고 부른다.

만약 목적함수의 값은 쉽게 얻을 수 있다면 MOP에서 최종해는 예를 들어, 목표수준 방법 (Sawaragi 등, 1985)과 같은 기존의 방법으로 파레토 최적해들의 집합을 찾을 수 있다. 이 목적함수의 값을 평가하기 위하여 많은 계산 시간이 소요되는 경우에, 대화형 방법은 시간의 제약으로 인해 부적합하다. 최근, 유전자 알고리즘을 포함한 진화 알고리즘과 신경망 등의 계산지능 방법은 파레토 최적의 해를 생성하기 위해 연구되고 있다 (Coello Coello 등, 2002; Deb, 2001; Nakayama 등, 2009; Poles, 2003). 특히 두 개 또는 세 개의 목적함수를 가지는 문제들에 대하여 파레토 프런티어를 시각화할 수 있기 때문에 이러한 근사 최적화 방법이 적합하다.

3.1. 제안된 알고리즘

이제, MOP에서 SVM을 적용하여 파레토 프런티어 생성하고, 잘 근사된 파레토 프런티어를 생성하기 위해 SVM과 진화 알고리즘 (Yun 등, 2001)을 결합하고자 한다. 특히, 제안된 알고리즘은 다음과 같이 요약된다 :

1 단계 2에서 4를 여러 번의 반복을 통하여 서포트 벡터를 축적시킨다 :

1. (초기화) 설계변수공간에서 임의로 표본점 $\mathbf{y}_1, \dots, \mathbf{y}_\ell$ 을 생성한다.
2. (평가) 자료점들 $\mathbf{F}(\mathbf{x}_1), \dots, \mathbf{F}(\mathbf{x}_\ell)$ 을 갖는 SVM_D를 풀어서 결정함수 h 를 평가한다. 여기서, MOP에서 $\mathbf{f}(\mathbf{x})$ 의 확장된 목적함수 $\mathbf{F}(\mathbf{x}) := (F_1(\mathbf{x}), \dots, F_m(\mathbf{x}))^T$ 는

$$F_i(\mathbf{x}) = f_i(\mathbf{x}) + \sum_{j=1}^n p_j \times \max \{g_j(\mathbf{x}), 0\}^a,$$

로 나타내어진다. 여기서, p_j 는 벌칙 계수이고 a 는 벌칙의 지수이다. 확장된 목적함수를 이용하여, 실행불가능한 (infeasible) 자료 점들을 이용하고 동시에 다음 단계에서 자료집합으로부터 효과적으로 제거할 수 있다.

3. (선택) $h(\mathbf{F}(\mathbf{x}_i))$, $i = 1, \dots, \ell$ 의 값을 계산하고 그 값을 내림차순해서 $\ell/2$ 개의 점을 선택한다. $h(\mathbf{F}(\mathbf{x}))$ 의 작은 값은 모든 자료점이 포함된 지역과 가깝다는 것을 의미한다. 예를 들어, h 값이 0은 Figure 3.1(b)에서 점선상에 있는 것과 같다.

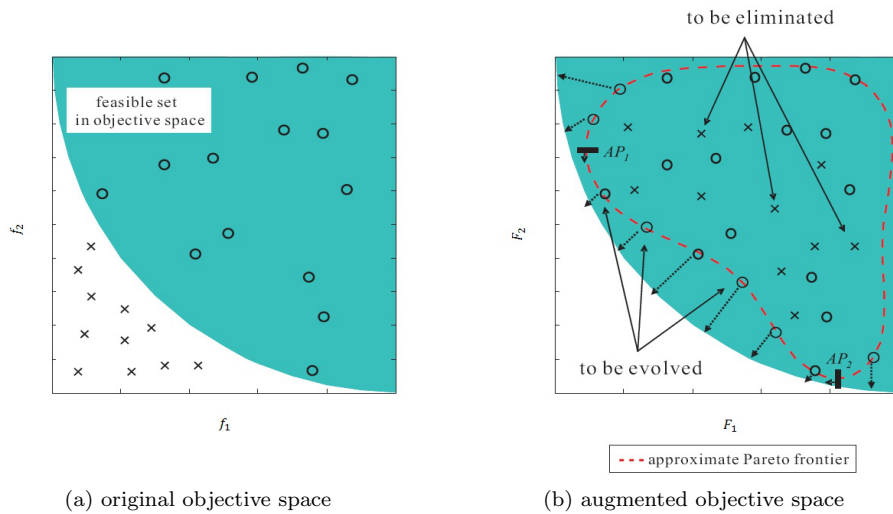


Figure 3.1 Geometric interpretation of the proposed method

4. (교차-돌연변이) 3.에서 선택된 점들 중에서 $\ell/2$ 개의 쌍을 임의로 만들고 새로운 집합을 생성한다. 주어진 돌연변이 확률에 따라 돌연변이를 시키고 2로 간다.

2 단계 1 단계에서 축적된 서포트 벡터 중에서 실행가능한 자료점을 가지고 SVM_D 를 해결하는 의사결정함수 $h(\mathbf{x})$ 를 계산한다. 즉, Figure 3.1(b)에서 점선으로 표시되어 있는 부분 ($h(\mathbf{x}) = 0$)에서 AP_1 에서 AP_2 까지의 선이 근사 파레토 프런티어가 된다.

Figure 3.1(b)에서 보는바와 같이 제안된 방법의 특징은 기계학습 알고리즘과 진화 알고리즘의 결합에 있다. 진화 알고리즘의 대표적인 방법인 유전자 알고리즘은 Holland (1975)에 의하여 제안되었고, Goldberg (1989)에 의하여 유전자 알고리즘은 최적화의 유용한 도구가 될 수 있다고 밝혔다. (유전자 알고리즘의 추가적인 설명은 원래의 참고문헌과 특히, Nakayama 등 (2009)을 참고하여라). 진화 알고리즘을 사용하여 가능한 많은 실행가능한 자료점들을 만들고 실행 시간에 따라 자료점들은 실제 파레토 프런티어로 다가간다. 제안된 알고리즘을 적용한 SVM은 원래 입력 공간에 자료점들을 모두 포락(envelopment)하는 영역을 제공한다. 각각 알고리즘의 최적을 활용하는 것은 효율적으로 수렴하고 빠르게 실제 파레토 프런티어에 다가간다.

3.2. 수치예제

다음은, 몇 개의 검정용 예제들을 통하여 제안된 방법의 효율성을 보인다. 일반적으로 계산지능을 이용한 다목적 최적화 방법들의 대부분의 연구는 둘 혹은 세 개의 목적함수들을 가지는 문제에서 시행되고 있고, 본 연구에서는 두 개의 목적함수들을 가지는 검정용 문제들을 다룬다.

그 이유는 우선, 파레토의 프런티어 근처에서 수렴하는 것과 광범위한 해들을 찾는 두 작업들을 적절하게 검정할 수 있으며, 둘째, 네 개 이상의 목적함수들을 가지는 경우에는 파레토 프런티어를 시각화하기가 어렵기 때문이다.

검정용 문제 1 (Yun 등, 2001)

$$\begin{array}{ll}
\text{minimize} & f_1(x_1, x_2) := x_1 \\
\text{minimize} & f_2(x_1, x_2) := x_2 \\
\text{subject to} & (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1, \\
& 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1.
\end{array}$$

검정용 문제 2 (Yun 등, 2001)

$$\begin{array}{ll}
\text{minimize} & f_1(x_1, x_2) := x_1 \\
\text{minimize} & f_2(x_1, x_2) := x_2 \\
\text{subject to} & x_1^3 - 3x_1 - x_2 \leq 0, \\
& -1 \leq x_1 \leq 1, -2 \leq x_2 \leq 2.
\end{array}$$

검정용 문제 3 (Deb, 2001)

$$\begin{array}{ll}
\text{minimize} & f_1(x_1, x_2) := x_1 \\
\text{minimize} & f_2(x_1, x_2) := 1 + x_2^2 - x_1 - 0.2 \sin(3\pi x_1) \\
\text{subject to} & 0 \leq x_1 \leq 1, -2 \leq x_2 \leq 2.
\end{array}$$

검정용 문제 4 (Binh 등, 1997)

$$\begin{array}{ll}
\text{minimize} & f_1(x_1, x_2) := 4x_1^2 + 4x_2^2 \\
\text{minimize} & f_2(x_1, x_2) := (x_1 - 5)^2 + (x_2 - 5)^2 \\
\text{subject to} & (x_1 - 5)^2 + x_2^2 \leq 25, \\
& (x_1 - 8)^2 + (x_2 + 3)^2 \geq 7.7, \\
& -15 \leq x_1, x_2 \leq 30.
\end{array}$$

위의 문제들에 대한 결과는 Figure 3.2에서 보여진다. 그림에서 어두운 부분은 목적공간에서 실행 가능한 부분이고 검정용 문제 4를 제외하고 점선의 부분은 각 문제에서 근사된 파레토 프런티어를 나타내고, 진한 실선이 실제 파레토 프런티어를 나타내고 어두운 부분은 자료점을 모두 포함한 것이다. 검정용 문제들은 Figure 3.2에서 보는 바와같이 다음과 같은 성질들을 가진다.

1. 검정용 문제 1, 2 그리고 4는 설계 변수들의 측면 제약식뿐만 아니라 비선형 제약식이 있다.
2. 검정용 문제 3은 복잡한 제약식들을 가지지는 않지만 파레토 프런티어는 불록하지 않고 연결되어 있지 않다.
3. 검정용 문제 1과 2에 대하여, 파레토 최적해들은 실행 가능한 영역의 경계에 있다.
4. 검정용 문제 2와 4에서 실행 가능한 영역은 설계공간에서 불록하지 않다.
5. 검정용 문제 1의 파레토 프런티어는 불록하지도 오목하지도 않다.
6. 검정용 문제 4의 어떤 파레토 최적해는 실행 가능한 영역의 경계에 있고 두 번째 목적함수의 약간의 감소가 첫 번째 목적함수의 큰 증가로 연결된다.

MOP들을 해결하는데 있어서 가능한 한 실제 파레토 프런티어상에 있거나 그 근방에 있고 잘 분포된 해들을 찾는 것이 중요하다. 계산지능 특히 진화 알고리즘을 이용한 기존의 알고리즘들은 위에서 서술한 것들 같이 다양한 특징들이 있으므로, 실제 파레토 프런티어에 수렴하고 파레토 최적해들의 다양성을 유지하는 어려움들에 직면해 왔고 또한 이러한 어려움들을 극복하기 위해 다양한 방법들이 시도되어 왔다. 이러한 관점에서 제안된 방법은 이러한 어려움을 극복 할 수 있으며, 또한 100(자료) × 5(반복)의 작은

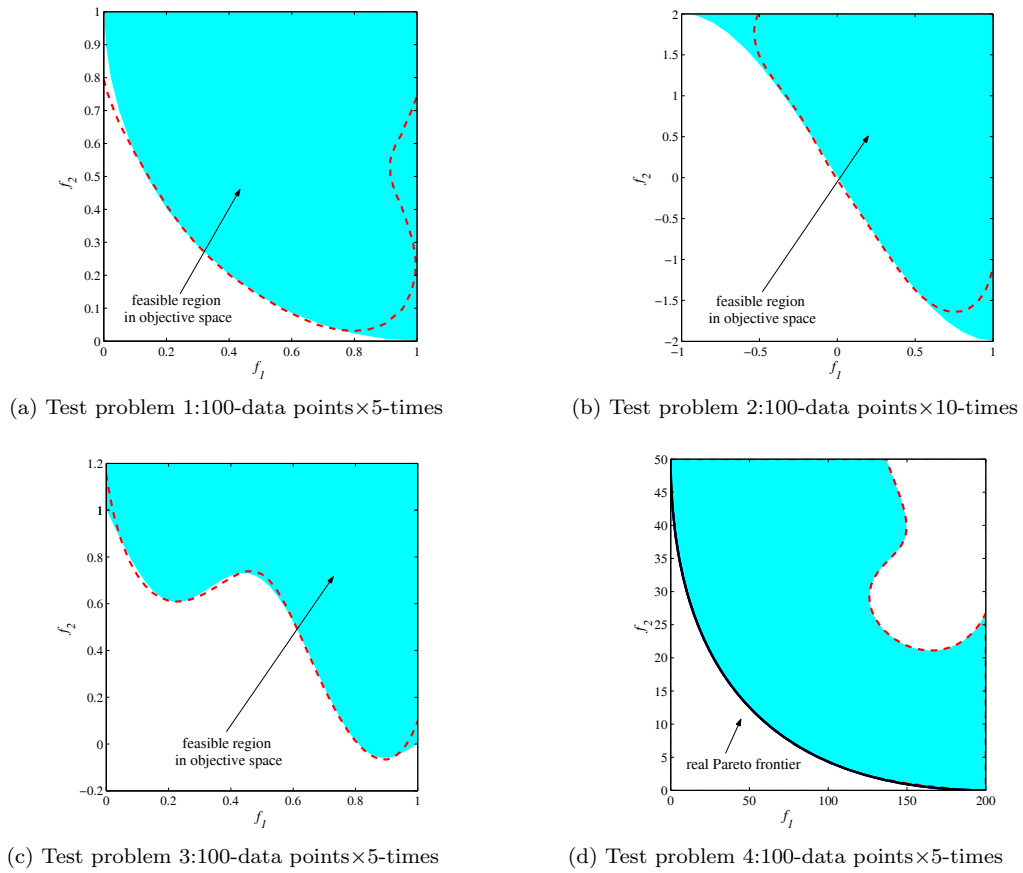


Figure 3.2 Pareto frontiers for test problems

표본실험 횟수 내에서 실제 파레토 프런티어와 거의 유사한 파레토 프런티어를 얻을 수 있으나 진화 알고리즘의 경우는 80(자료) × 30(반복)을 통해서 유사한 결과를 얻을 수 있었다.

4. 공학설계문제

이번 절에서 제안된 방법의 효율성을 보이기 위해 실제 공학 설계문제에서 실험할 것이다. 많은 실제적인 공학 문제에서 종종 설계변수들에 의해 정확하게 알려져 있지 않은 블랙박스 목적함수를 가지고 있다. 실제의 표본 또는 컴퓨터 실험 (= 표본 자료 점들의 수 × 실행 시간)에 의해 값이 주어지고, 구조 해석, 유체역학 해석 및 열역학적 해석과 같은 여러 종류의 분석의 경우에 목적함수의 값들을 평가하는데 적용된다. 이러한 분석은 비용의 측면에서 상당히 비싸다. 게다가 우리는 사전에 계산을 언제 중지하는지를 알 수 없고, 주어진 계산시간과 비용의 한계 때문에 상대적으로 일찍 계산을 종료한다. 이러한 상황에서 설계변수들에 의해 목적함수의 형태가 정확하게 주어지지 않은 목적함수들을 가진 최적화 문제에서 가능한 적은 실험으로 잘 근사된 해를 생성하는 것이 중요한 문제이다. 이와 같은 관점에서 제안된 방법은 다음과 같은 문제를 통해 바람직한 성능을 가지고 있다는 것을 보인다.

예제 4.1 : 캔틸레버 기둥 문제

우선 두 개의 설계변수를 가진 캔틸레버 설계문제를 고려하자. Figure 4.1(a)에서 보듯이 Deb (2000)에서 이용된 직경 (d)와 길이 (l)의 기둥 (beam)은 끝부분의 하중 P 를 지탱할 수 있다. 캔틸레버 설계문제는 설계에서 두 가지 목적함수가 서로 상충한다. 무게 f_1 의 최소화과 끝부분의 굴절 f_2 의 최소화, 그리고 두 개의 제약식들 즉, 개발된 최대 응력 σ_{\max} 은 허용강도 S_y 보다 작아야하고 끝부분의 굴절 δ 은 지정된 한도 δ_{\max} 보다 작아야 한다. 설명된 모든 것들을 고려한 최적화 문제는 아래와 같이 공식화할 수 있다.

$$\begin{aligned} & \text{minimize} && f_1(d, l) := \rho \frac{\pi d^2}{4} l \\ & \text{minimize} && f_2(d, l) := \delta = \frac{64Pl^3}{3E\pi d^4} \\ & \text{subject to} && \sigma_{\max} = \frac{32Pl}{\pi d^3} \leq S_y, \\ & && \delta \leq \delta_{\max}, \\ & && 10 \leq d \leq 50, \quad 200 \leq l \leq 1000, \end{aligned}$$

그리고 사용된 모수값들은

$$\begin{aligned} \rho &= 7800 \text{ kg/m}^3, \quad P = 1 \text{ kN}, \quad E = 207 \text{ GPa}, \\ S_y &= 300 \text{ MPa}, \quad \delta_{\max} = 5 \text{ mm} \end{aligned}$$

이다.

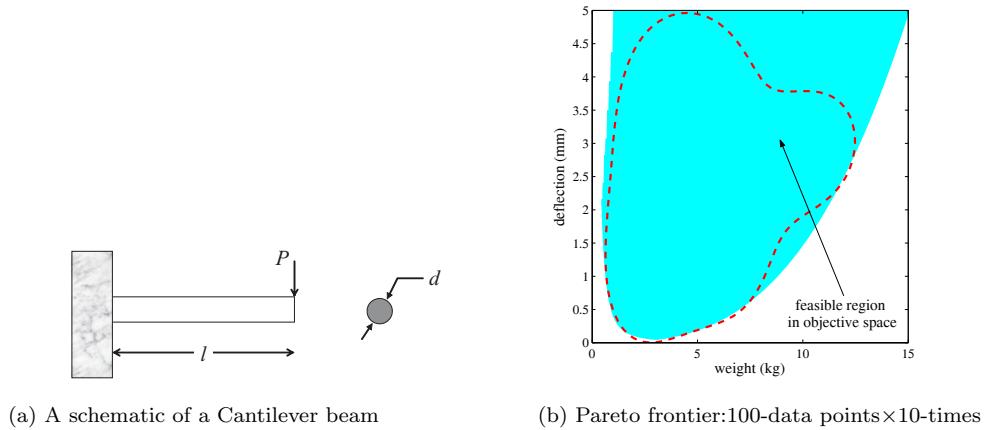


Figure 4.1 Cantilever beam design problem

5. 결론

본 논문은 MOP에서 근사 파레토 프런티어를 얻기 위하여 서포트 벡터기계를 이용한 방법을 제안하였고 더 정확하게 근사된 파레토 프런티어를 생성하기 위해 제안된 방법에 진화 알고리즘을 적용했다. 또한, 몇 가지 검정용 문제들 및 실제 문제에 제안된 방법을 적용하고, 시뮬레이션 결과를 통하여 제안

된 방법의 효율성을 보였다. 따라서, 제안된 방법이 기존의 알고리즘들에서 필요한 실험횟수의 1/10에서 1/2로 줄일 것으로 예상된다. 실제 공학 설계문제들에서 가능하면 실험의 수는 1000 번(가능하다면 100 번) 이하가 바람직하다. 다시 말하면 이것은 실험의 횟수를 가능한 줄이는 것이 얼마나 중요한지를 의미한다. 이런 관점에서 제안된 알고리즘은 기존의 방법보다 더 우수할 수 있다. 마지막으로 서포트 벡터기계와 진화 알고리즘을 결합한 제안된 방법이 좀 더 복잡한 형태의 목적함수들과 제약식들을 갖는 문제에서 잘 근사된 파레토 프런티어를 생성할 수 있다.

References

- Binh, T. T. and Korn, U. (1997). MOBES: A multiobjective evolution strategy for constrained optimization problems. *Proceedings of the 3rd International Conference on Genetic Algorithms*, 176-182.
- Coello Coello, C. A., Van Veldhuizen, D. A. and Lamont, G. B. (2002). *Evolutionary algorithms for solving multi-objective problems*, Kluwer Academic Publishers, New York.
- Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms*, John & Wiley Sons, New York.
- Deb, K., Pratap, A. and Moitra, S. (2000). Mechanical component design for multiple objectives using elitist non-dominated sorting GA. *Proceedings of the Parallel Problem Solving from Nature VI (PPSN-VI)*, 859-868.
- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley, Boston.
- Holland, J. H. (1975). *Adaptation in natural and artificial systems*, University of Michigan Press, Ann Arbor.
- Hwang, C. and Shim, J. (2012). Mixed effects least squares support vector machine for survival data analysis. *Journal of the Korean Data & Information Science Society*, **23**, 739-748.
- Nakayama, H., Yun, Y. B. and Yoon, M. (2009). *Sequential approximate multiobjective optimization using computational intelligence*, Springer Verlag, Berlin Heidelberg.
- Palli, N., Azram, S., McCluskey, P. and Sundararajan, R. (1998). An interactive multistage ϵ -inequality constraint method for multiple objectives decision making. *ASME Journal of Mechanical Design*, **120**, 678-686.
- Park, D. J., Yun, Y. B. and Yoon, M. (2012). Prediction of bankruptcy data using machine learning techniques. *Journal of the Korean Data & Information Science Society*, **23**, 569-577.
- Poles, S. (2003). *MOGA-II An improved multi-objective genetic algorithm*, Esteco Achieving Perfection Technical Report 2003-006, 1-14.
- Sawaragi, Y., Nakayama, H. and Tanino, T. (1985). *Theory of multiobjective optimization*, Academic Press Inc., Boston.
- Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J. and Williamson, R. (2001). Estimating the support of a high-dimensional distribution. *Neural Computation*, **13**, 1443-1471.
- Schölkopf, B. and Smola, A. J. (2002). *Learning with kernels*, MIT Press, New York.
- Seok, K. H. (2010). Semi-supervised classification with LS-SVM formulation. *Journal of the Korean Data & Information Science Society*, **21**, 461-470.
- Steinwart, I. and Christmann, A. (2008). *Support vector machines*, Springer, New York.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*, Springer Verlag, New York.
- Yun, Y. B., Nakayama, H., Tanino, T. and Arakawa, M. (2001). Generation of efficient frontiers in multi-objective optimization problems by generalized data envelopment analysis. *European Journal of Operational Research*, **129**, 586-595.

Generating of Pareto frontiers using machine learning[†]

Yeboon Yun¹ · Nayoung Jung² · Min Yoon³

¹Faculty of Environmental and Urban Engineering, Kansai University

^{2,3}Department of Statistics, Pukyong National University

Received 22 March 2013, revised 17 April 2013, accepted 2 May 2013

Abstract

Evolutionary algorithms have been applied to multi-objective optimization problems by approximation methods using computational intelligence. Those methods have been improved gradually in order to generate more exactly many approximate Pareto optimal solutions. The paper introduces a new method using support vector machine to find an approximate Pareto frontier in multi-objective optimization problems. Moreover, this paper applies an evolutionary algorithm to the proposed method in order to generate more exactly approximate Pareto frontiers. Then a decision making with two or three objective functions can be easily performed on the basis of visualized Pareto frontiers by the proposed method. Finally, a few examples will be demonstrated for the effectiveness of the proposed method.

Keywords: Evolutionary algorithm, multi-objective optimization, Pareto frontier, support vector machines.

[†] This work was supported by the Pukyong National University Research Fund in 2012 (PK-2012-45).

¹ Professor, Faculty of Environmental and Urban Engineering, Osaka 564-8680, Japan.

² Master of Science, Department of Statistics, Pukyong National University, Busan 608-737, Korea.

³ Corresponding author: Associate professor, Department of Statistics, Busan 608-737, Korea.

E-mail: myoon@pknu.ac.kr