제어 장벽함수에 기반한 안전성 보장 비선형 제어 기법 소개

자율주행 차량의 장애물 회피 등의 문제에서처럼, 제어 시스템의 안전성은 제어 공학자에게 더 이상 낯선 개념이 아니다. 이를 다루기 위한 노력의 일환으로, 최근 수 년 간 제어 학계에서는 제어 장벽함수라는 개념을 도입하고 이를 이용하여 제어 시스템의 안전성을 확보하는 방법론이 적극적으로 제안되고 있다. 필자는 제어 장벽함수를 정의하는 두 갈래의 서로 다른 관점들에 대해 소개하고, 이러한 관점의 차이로부터 관련 연구들이 어떤 방향성으로 뻗어나가고 있는지를 개념적으로 간략히 소개해보고자 한다.

박경훈(서울시립대학교 전자전기컴퓨터공학부)

Ⅰ. 서 론: 제어 시스템의 안전성이란?

기술의 발전과 함께 많은 제어 시스템들이 사람과 같은 환경을 공유하며 더욱 다양한 작업을 수행하기를 요구받고 있다. 이에 따라 제어 시스템의 안전성(safety)라는 개념이 지난 수 년간 큰 주목을 받아왔으며, 이러한 시대적 흐름에 발맞춰서 그간 많은 연구가 이루어져왔다. 본 고에서 필자는 제어 이론 관점에서 정의된 비선형 제어 시스템의 안전성의 개념을 간단히 소개하고, 이를 추구하는 제어 방법론들 중에 특히 제어 장벽함수(control barrier function)에 기반한 연구들에 대해 간략히 짚어보고자 한다.

소개에 앞서 한 가지 강조하고 싶은 부분은, 제어 공학에서 안전성이라는 개념(혹은 이를 위한 제어 이론) 자체는 그다지 낯설지 않다는 점이다. 예를 들어 모바일 로봇 혹은 자율주행 차량의 경우, 이 시스템들의 안전성이란 장애물과의 충돌을 회피하는 것으로 정의해볼 수 있다. 이러한 주행 안전성을 보장하기 위한 가장 잘 알려진 방법은 모델 예측 제어(model predictive control 혹은 MPC)인데, 최적화 기법에 기반하여 제약조건을 직접 다룰 수 있다는 큰 장점이 있다. 이를 활용하여 MPC를 제어기로 직접 사용하거나, 혹은 MPC를 경로 계획 알고리즘으로 이용하여 장애물을 회피하는 적절한 경로를 계획한 후 계획된 경로를 올바르게 따라가는 또 다른 (간단한 형태의) 제어기를 구성하는 식의 접근이 가능한데, 불행히도 MPC는 시스템 차수가 커지거나 시스템이 복잡한 경우에는 연산 복잡도가 매우 커질 수 있다는 큰 단점이 존재한다. 이로 인해 수학적 모델이 다소 복잡한 시스템에 대해서는 제어기로 직접 활용하기 어려울 수 있고, 또 급작스럽게 장애물이 끼어드는 등 동적 상황에서 빠른 대처가 어려울 수도 있다.

한편 본 고에서 다루게 될 제어 장벽함수에 기반한 방법론의 경우, 일단 제어 장벽함수가 특정 조건을 만족하도록 주어져 있다면 이를 이용한 제어기는 비교적 간단한 형태로 설계할 수 있다는 점에서 모델 예측 제어 기반의 접근법들에 비해 장점이 있다. 또한 우리에게 친숙한 비선형 제어 이론을 기반으로 설계하기 때문에 시스템의 비선형성을 MPC 등의 기법에 비해 손쉽게 다룰 수 있으며, 만약 볼록함(convexity)을 가정하지 않고 복잡한 형태로 존재하는 장애물에 대해서도 (장벽함수만 잘 정의가 된다면) 비교적 수월하게 다룰 수 있다.

제어 장벽변수에 기반한 안전성 기반 제어 기법을 살펴보기에 앞서, 본 장에서는 제어 입력이 인가되지 않은 다음의 비선형 시스템

$$\dot{x} = f(x), \qquad x \in \mathcal{X} \subset \mathbb{R}^n \tag{1}$$

에 대하여 시스템의 안전성을(최대한 간단한 형태로) 정의해보기로 한다. 여기서 $\mathcal{X} \subset \mathbb{R}^n$는 상태 공간이며, 함수 $f: \mathcal{X} \to \mathbb{R}^n$는 (편의상 전역적) 립쉬츠(Lipschiptz) 조건을 만족한다고 가정한다.

제어 공학의 관점에서는 식 (1)의 상태 변수 $x(t)$가 어떠한 원치 않는 영역에 도달하지 않는다면, 식 (1)의 시스템이 안전하다고 할 수 있을 것이다. 이를 좀 더 엄밀히 쓰기 위하여, 상태 변수가 $x(t)$가 계속 머무르기를 기대하는 $\mathcal{X} \subset \mathbb{R}^n$의 부분 집합 $C \subset \mathbb{R}^n$을 안전 영역(safe region)이라 하며, 이와 유사하게 $x(t)$가 진입하지 않기를 기대하는 $\mathcal{X} \subset \mathbb{R}^n$의 부분 집합 $D \subset \mathbb{R}^n$을 불안전 영역(unsafe region)이라 하자. 만약 우리에게 안전 영역 혹은 불안전 영역이 잘 정의되어 주어져 있다면, 이들을 통해 식 (1)의 비선형 시스템에 대한 안전성을 아래와 같이 정의할 수 있다.

정의 1 : 식 (1)의 시스템에 대하여 만약 안전 영역 $\mathcal{C} \subset \mathbb{R}^2$과 불안전 영역 $\mathcal{C} \subset \mathbb{R}^2$이 주어져있다고 할 때,

(a) $x(0) \in int(\mathcal{C}) \Rightarrow x(t) \in int(\mathcal{C}),\ \forall t \geq 0$를 만족하거나[1],

(b) $x(0) \notin \overline{\mathcal{D}} \Rightarrow x(t) \notin \overline{\mathcal{D}},\ \forall t \geq 0$를 만족한다면[3,4]

이 시스템은 안전(safe)하다고 한다.

한편 제어 입력이 존재하는 다음의 비선형 시스템

$$\dot{x} = f(x) + g(x)u \tag{2}$$

에 대하여, 안전성 보장 제어 문제를 다음과 같이 정의한다.

안전성 보장 제어 문제 [4]: 식 (2)의 비선형 시스템에 대하여, 전체 시스템이 정의 1의 관점에서 안전하면서, $x=0$이 점근적으로 안정(asymptotically stable)하도록 하는 상태 궤환 제어기 $u=k(x)$를 찾는다.

개념적으로 불안전 영역 D가 안전 영역 C의 여집합으로 선정될 것을 생각한다면, 정의 1의 안전성 조건인 (a)와 (b)는 매우 유사해 보인다. (물론 안전성을 따질 필요가 없는 제 3의 영역까지 고려한다면, 단순히 D가 C의 여집합의 형태로 표현되지 않을 수도 있다.) 다만 그럼에도 불구하고 시스템의 안전성 조건을 C와 D를 기준으로 따로 기술한 이유는, 관련 문헌들에서 안전성을 논할 때 둘 중 어떤 영역을 기준으로 장벽함수를 정의하느냐에 따라 전혀 다른 두 갈래의 접근법이 제안되었기 때문이다. 본 고에서 필자는 이들을 각각 "안전 영역 중심 접근법[1,2]"과 "불안전 영역 중심 접근법[3,4]"으로 표기한 후, 각 접근법의 주요 개념들에 대해 비교해가며 기술하고자 한다. 설명에 앞서, 본 고의 대부분의 내용은 [1]-[4]를 참고 및 정리하여 기술하였음을 미리 언급하며, 정의 1 역시 그러한 관점하에서 각 문헌에서 정의한 안전성을 재가공하여 기술하였음을 명시한다. 더불어 본 고에서는 "알기쉬운 제어이론"의 취지에 맞게 수학의 엄밀성보다는 개념 위주의 설명에 치중할 예정이며, 누락된 내용이나 구체적인 증명은 참고 문헌들을 확인해주기 바란다.

Ⅱ. 장벽함수를 정의하는 두 가지 방향성

1. 안전 영역 중심 접근법에서의 장벽 함수

안전 영역 중심 접근법에서는 어떠한 미분 가능한 함수 $h(x)$를 도입하고, 이를 이용하여 안전 영역을 다음과 같이 정의하도록 한다.

$$\mathcal{C} \coloneqq \{x \in \mathbb{R}^n : h(x) \geq 0\}$$

안전 영역 C가 $h(x)$로부터 정의된 덕분에, 집합의 전방 불변성(forward invariance)을 직접 따지지 않고 $h(x)$의 부호 변화 유무 정도만 확인함으로써 시스템의 안전성을 논할 수 있다는 장점이 있다. 즉, 만약 $h(x) \geq 0$을 항상 만족한다면 당연하게도 식 (1)의 시스템의 상태 변수는 C에 머물러있게 되어 시스템이 안전하게 된다.

그렇다면 $h(x)$가 항상 양의 값을 가지는 조건은 어떻게 기술할 수 있을까? 지금의 독자들은 여러 가지 창의적인 생각이 들 수 있지만, 그 중에서 우리는 먼저 $h(x)$를 이용하여 정의된 아래의 고전적인 형태의 장벽함수(barrier function)들을 살펴보고자 한다.

$$b_1(x) = -\log\left(\frac{h(x)}{1+h(x)}\right),\ b_2(x) = \frac{1}{h(x)} \tag{3}$$

(최적화 이론에 익숙한 독자들이라면 $b_1(x)$과 같은 로그-장벽 형태의 함수들이 더욱 친숙하게 느껴질 것이다.) 이러한 함수들은 상태 변수가 안전 영역 c내부에 있을 때는 항상 양수이며, 안전 영역의 경계 $\partial\mathcal{C}$로 다가갈수록 함수값의 크기가 커진다는 공통적인 특징을 가진다. 따라서 만약 위와 같은 장벽함수가 전 시간 영역에서 유한한 값으로 남아있을 수 있다면, 상태 변수가 항상 c의 내부에 존재한다고 결론내릴 수 있을 것이다.

위의 장벽함수들이 유한한 값으로 남아있으려면 $h(x)$에 대한 조건이 추가로 필요한데, 이를 간단히 설명하기 위해 $b_2(x)$를 예시로 전개해보자. $b_2(x)=1/h(x)$의 시간에 대한 미분은 $\dot{b}_2(x) = \frac{\partial b_2}{\partial h} L_f h(x) = -b_2^2(x) L_f h(x)$로 계산되는데, 만약 $h(x)$가 안전 영역을 포함하는 어떤 집합 $\mathcal{C}^+ \supset \mathcal{C}$에 대하여

$$L_f h(x) > -h(x)^3, \quad \forall x \in C^+ \tag{4}$$

을 만족한다면(식 (4)에 대한 자세한 설명은 추후에 하도록 하겠다), $x \in c^+$에 대하여 $\dot{b}_2(x)$는 다음과 같이 계산된다.

$$\dot{b}_2(x) = L_f b(x) = \frac{\partial b_2}{\partial h} L_f h(x) = -b_2^2(x) L_f h(x) \leq \frac{1}{b_2(x)} \tag{5}$$

이제 비선형 제어 이론에서 주로 등장하는 comparison lemma 등을 이용하면, 위의 장벽함수가 어떠한 class-KL 함수인 σ에 대하여,

$$b_2(x(t)) \leq \frac{1}{\sigma\left(\frac{1}{b(x(0))}, t\right)}, \quad \forall t \geq 0$$

를 만족하게 됨을 쉽게 알 수 있다. 이 때 위 부등식의 우변은 class-KL 함수의 특징으로 인해 시간이 지나도 유한한 값으로 남아 있기 때문에 전체 시스템은 정의 1의 관점에서 안전해진다. (지면 관계 상 자세한 도출은 생략하도록 하겠다. 이에 대한 자세한 내용이 궁금한 독자들은 [Theorem 1, 2]의 증명을 살펴보기 바란다.)

이제 위의 논의를 일반화한 다음의 정의와 정리를 차례로 소개한다.

정의 2 [Definition 1, 2] *:* 식 (1)의 비선형 시스템과 미분 가능한 함수 $b: \mathbb{R}^n \to \mathbb{R}$이 어떠한 class-$K$ 함수들 α_i, i = 1,2,3에 대하여 부등식

$$\frac{1}{\alpha_1(h(x))} \leq b(x) \leq \frac{1}{\alpha_2(h(x))}, \quad L_f b(x) \leq \alpha_3(h(x))$$

이 모든 $x \in \operatorname{int}(\mathcal{C})$에 대하여 만족한다면, 이 함수를 ($c$에 대한) 상보 장벽함수(reciprocal barrier function, RBF)이라고 한다.

앞선 예제에서 식 (4)-(5)를 정리하면 정의 2의 두 번째 부등식이 도출됨을 참고로 말해둔다.

정리 1 [Theorem 1, 2] : 식 (1)의 시스템에 대하여 안전 영역 c에 대한 상보 장벽함수인 $b(x)$가 존재한다면, 이 시스템은 정의 1(a)의 관점에서 안전하다.

정의 2(혹은 식 (3))에서 소개된 상보 장벽함수들은 전통적인 개념에서 기술한 장벽함수로써, 안전 영역의 경계 주변에서 장벽과 같이 큰 값을 주기 때문에 그 모양이 직관적이라는 장점이 있다. 하지만 이러한 직관성을 위해 기본적으로 시스템의 상태 변수가 경계로 다가갈수록 함수값이 커지며, 심지어 $h(x)$가 음수인 경우 정의가 잘 되지 때문에, 이를 이용하여 제어기를 설계할 때 구현 상 문제가 발생할 여지가 다분하다.

이를 해결할 방법으로 안전 영역 기반 접근법에서 제시한 방향성은 다소 (혹은 너무 싱겁게도) 간단한데, 앞서 안전 영역 c을 정의하기 위해 사용한 함수 $h(x)$를 새로운 장벽함수로 정의하는 것이다. 물론 이렇게 정의된 $h(x)$는 "장벽"으로써의 모양을 기대하기 어렵다. 즉, $b(x)$와는 달리 $h(x)$는 경계 부근에서도 유한한 값(더 자세히 말하자면 $h(x) = 0$)을 가진다. 그럼에도 이러한 특징 덕분에 $b(x)$를 활용한 제어기에 비해 $h(x)$를 활용한 제어기의 구현이 더 수월할 것이라는 기대가 있다.

그렇다면 (비록 장벽과 같은 생김새는 아니지만) $h(x)$가 장벽함수로써 역할하기 위한 조건은 무엇일까? 만약 $h(x)$가 적어도 경계에서는

$$L_f h(x) \geq 0, \forall x \in \partial C \tag{6}$$

를 만족한다면, $h(x)$의 값이 항상 0보다 크도록 할 수 있을 것이다. 문제는 위 식이 다소 보수적이라는 것인데, 앞서 상보 장벽함수에 대한 예제에서도 식 (4)에서 부등식의 우변이 음수가 됨을 어느 정도 허용하고 있었기 때문이다(따라서 어떤 의미에서 식 (4)는 식 (6)의 완화된 버전이라고도 생각해볼 수 있다). 그러므로 지금의 우리도, 식 (6)를 조금 더 완화한 조건으로 $h(x)$를 정의하고자 한다. $h(x)$가 음수인 경우도 고려하기 위해 먼저 양수인 입력값에 대해서만 정의되었던 class-K 함수의 정의를 조금 손 볼 필요가 있다.

정의 3 [Definition 2, 2] : 어떠한 양수 $a, b \in \mathbb{R}$에 대하여 정의된 연속 함수 $\alpha: (-b, a) \to \mathbb{R}$가 단조 증가(strictly increasing)하며 $\alpha(0) = 0$을 만족한다면, 이 함수를 확장된 class-K (혹은 class-K_+)라고 한다.

정의 4 [Definition 3, 2] : 식 (1)의 비선형 시스템과 미분 가능한 함수 $h: \mathbb{R}^n \to \mathbb{R}$ 이, 만약 $c^+ \supset c$와 class-K_+ 함수 α_4에 대하여

$$L_f h(x) \geq -\alpha_4(h(x)), \quad \forall x \in \mathcal{C}^+$$

를 만족한다면, 이러한 함수 $h(x)$를 (c에 대한) 영점화 장벽함수(Zeroing barrier function, ZBF)라고 한다.

정리 2 [Proposition 2, 1] : 식 (1)의 시스템에 대하여 영점화 장벽함수인 $h(x)$가 존재한다면, 이 시스템은 정의 1의 관점에서 안전하며, 집합 C는 점근적으로 안정하다.

만약 $h(x)$가 영점화 장벽함수가 된다면, 이를 통해 (조금 더

조건을 추가하여) $1/h(x)$를 상보 장벽함수로 만들 수 있다. 자세한 사항은 [Theorem 2, 2]를 참고하기 바란다.

2. 불안정 영역 중심 접근법에서의 장벽함수

안전 영역 중심으로 시스템의 안전성을 해석할 때는 어떤 어려운 점이 있을까? 가장 큰 어려운 점은, 안전성 뿐 아니라 시스템의 안정성(stability)를 위해 리아프노프 함수와 장벽함수를 동시에 고려해야 할 때 찾아온다. 앞서 정의한 영점화 장벽함수는 안전성을 위해 $h(x)$는 되도록 큰 값을 가져야 하나 고전적인 형태의 리아프노프 함수는 평형점으로의 수렴을 위해 작은 값을 가져야 하기 때문에, 이 둘을 함께 묶어서 생각해보기가 쉽지 않다. 이러한 어려움은 상보 장벽함수도 유사하게 겪고 있으며, 이 부분은 본 절에서 소개할 새로운 형태의 장벽함수와 대비된다.

이전 절에서 소개한 안전 영역 중심 접근법과는 달리, 참고문헌 [3,4]에서 소개하는 불안전 영역 중심 접근법에서는 불안전 영역 D가 먼저 주어지고 이에 대한 장벽함수를 정의한다는 점에서 근본적인 차이점이 있다(안전 영역 접근법에서는 $h(x)$가 먼저 정해진 이후에야 비로소 안전 영역을 정의함을 다시 한 번 상기하자.) 이제 참고 문헌 [3,4]에서 정의하는 형태의 장벽함수를 아래와 같이 소개한다. (더 나아가기 전 미리 안내하자면, 참고 문헌 [3]에서는 정의 5의 함수를 barrier certificate라는 이름으로 불렀으나, 본 고에서는 이전 장의 함수들과의 용어 상 통일성을 유지위해 "장벽함수"라 칭하고자 한다. 다만 다른 장벽함수들과의 혼동을 막기 위해, 본 고에서는 필요하다면 아래 정의의 장벽함수를 D-장벽함수로 지칭하여 이들을 구분하고자 한다.)

정의 5 [Definition 1, 3] : 식 (1)의 비선형 시스템에 대하여 불안전 영역 D이 주어졌을 때, 미분 가능한 함수 $\beta: \mathbb{R}^n \rightarrow \mathbb{R}$가

$$\beta(x) > 0, \qquad \forall x \in \mathcal{D},$$
$$L_f\beta(x) \leq 0, \qquad \forall x \in \mathcal{X}$$

를 만족한다면 이러한 함수를 (D-)장벽함수라고 한다.

쉽게 생각하면 $\beta(x)$는 Ⅱ-1절에서 소개한 영점화 장벽함수 $h(x)$에 대하여 $-h(x)$와 매우 유사한 성질을 가진다고 볼 수 있다. 특히 $h(x)$와 유사하게 $\beta(x)$도 영역의 경계 주변에서 무한한 값을 가질 필요는 없기 때문에, 상보 장벽함수와 비교하였을 때 구현 상 이점을 가진다.

한편 우리는 이제 불안전 영역 D에 집중할 예정이므로, ($h(x)$와는 달리) 안전 영역에서 $\beta(x)$의 함수값에 대한 특별한 조건이 요구되지 않는다는 점은 분명한 차별점이다. 일반적으로는 (정의 상에서 직접 등장하지는 않았지만) 안전 영역 내에서 $\beta(x)$는 음수이기를 기대하고 $\beta(x)$를 설계할 것이다. 하지만 안전 영역 내에서 그 함수값의 크기 $|\beta(x)|$는 작아도 크게 상관이 없는데, 이는 마치 고전적인 비선형 제어 이론에서 상태변수가 평형점(equilibrium) 주변에 있을 때 리아프노프 함수의 함수값이 거의 0에 가깝다는 점과 매우 유사하다. 이러한 유사성을 고려해볼 때 안전성 조건 역시 리아프노프 해석과 매우 유사한 형태로 표현할 수 있게 된다.

정리 3 [Theorem 2, 3] : 식 (1)의 시스템에 대하여 D- 장벽함수인 $\beta(x)$가 존재한다면, 이 시스템은 정의 1의 관점에서 안전하다.

이러한 성질들은 4장에서 제어기를 설계할 때 보다 자세히 설명할 예정이다.

Ⅲ. 안전 영역/불안전 영역 기반 접근법에서의 제어 장벽함수

이제부터는 본격적으로 제어 입력이 존재하는 식 (2)의 비선형 시스템에 대하여 안전성 보장 제어 문제를 다뤄보고자 한다. 흥미롭게도 본 고에서 살펴볼 두 방향의 접근법들에서는 모두, 제어 리아프노프 함수로부터 영감을 받아서 제어 장벽함수를 제안하고 이를 이용한 제어기를 설계한다. 다만 목표하는 제어 응용 대상과 제안한 장벽함수의 구조적인 특성 등으로 인해 제어 장벽함수 역시 서로 차이가 있다. 아래에서는 이를 순서대로 살펴보고자 한다.

아래 정의들에서는 정의 2와 정의 4에서 각각 소개한 상보 장벽함수 및 영점화 장벽함수에 대한 (제어 리아프노프 함수 관점에서의) 확장에 대해 짤막하게 기술한다.

정의 6 [Definition 1, 2] : 식 (1)의 비선형 시스템과 미분 가능한 함수 $b: \mathbb{R}^n \rightarrow \mathbb{R}$이 어떠한 class-$K$ 함수들 α_i, i =1,2,3에 대하여

$$\frac{1}{\alpha_1(h(x))} \leq b(x) \leq \frac{1}{\alpha_1(h(x))},$$
$$\inf_{u\in U}\left[L_f b(x) + L_g b(x)u\right] \leq \alpha_3(h(x))$$

이 모든 $x \in \mathrm{int}(\mathcal{C})$에 대하여 만족한다면, 이 함수를 상보 제어 장벽함수(reciprocal control barrier function, RCBF)이라고

한다.

정의 7 [Definition 5, 2] : 식 (2)의 비선형 시스템과 미분 가능한 함수 $h: \mathbb{R}^n \rightarrow \mathbb{R}$ 이, 만약 $c^+ \supset c$와 어떠한 class-K_+ 함수 α_4에 대해

$$\sup_{u \in \mathcal{U}} \left[L_f h(x) + L_g h(x) u \right] \geq -\alpha_4\left(h(x)\right)$$

를 만족한다면, $h(x)$를 영점화 제어 장벽함수(zeroing control barrier function, ZCBF)이라고 한다.

한편 불안전 영역 기반 접근법[3,4]에서도 위와 매우 유사한 관점에서 제어 장벽함수를 정의한다.

정의 8 [4] : 식 (2)의 비선형 시스템에 대하여, 불안전 영역 D 이 주어졌을 때, 미분 가능한 함수 $\beta: \mathbb{R}^n \rightarrow \mathbb{R}$ 가 다음을 만족할 때 이 함수를 제어 (D-)장벽함수라고 한다.

(a) $\beta(x) > 0, \forall x \in \mathcal{D}$,

(b) $L_f \beta(x) \leq 0, \ \forall x \in \left\{ z \in \mathbb{R}^n \setminus \mathcal{D} : L_g \beta(z) = 0 \right\}$

(c) $\breve{\mathcal{C}} := \{ x \in \mathbb{R}^n : \beta(x) \leq 0 \} \neq \emptyset$.

정의 8의 조건 (b)는 제어 리아프노프 함수가 가져야할 조건과 매우 유사함을 상기하자. 한편 조건 (c)에서 정의되는 $\breve{C}$는, 어떤 함수로부터 정의되었는지에 따른 약간의 차이만 있을 뿐 개념적으로는 Ⅱ-1절에서 다룬 안전 영역 C와 개념적으로 매우 비슷하다.

Ⅳ. 제어 장벽함수를 이용한 안전성 보장 제어 기법

본 장에서는 Ⅲ장에서 소개한 제어 장벽함수를 이용하여, 안전성을 보장하는 제어 기법들을 짧게 소개해보고자 한다. 특히 본 장에서는 제어 장벽함수의 정의의 방향성이 특히, 각 방법론에서는 공통적으로 안전성 보장 문제를 풀기 위하여 $x \in \mathbb{R}^n \setminus \mathcal{D}$에 대하여 다음의 조건들을 만족하는 제어 리아프노프 함수 $V(x)$를 추가로 고려한다.

$$\overline{\alpha}_1(||x||) \leq V(x) \leq \overline{\alpha}_2(||x||),$$
$$\inf_{u \in \mathcal{U}} \left[L_f V(x) + L_g V(x) u \right] < -\overline{\alpha}_3(V(x))$$

가장 마지막 줄의 부등식은, infimum의 정의에 의해 “$L_f V(x) + L_g V(x) u \leftarrow \overline{\alpha_3}(V(x))$를 만족하는 u가 존재한다” 는 문장으로 대체 가능함에 유의하자.

1. 2차 계획법 기반 안전성 보장 제어

앞서 언급한 바에 따르면 안전 영역 기반 접근법에서 제안한 제어 상보 장벽함수와 제어 영점화 장벽함수 모두 그 구조로 인해 제어 리아프노프 함수와 직접 연결짓기 어렵다. 대신 안전 영역 기반 접근법[1,2]에서는 다른 부분에 주목하였는데, 우리가 고려하는 식 (2)의 비선형 시스템이 control-affine한 덕분에 정의 6과 7의 마지막 조건들에 대한 부등식

$$L_f b(x) + L_g b(x) u \leq \alpha_3\left(h(x)\right)$$
$$L_f h(x) + L_g h(x) u \geq -\alpha_4\left(h(x)\right)$$

이 (만약 x가 고정되어 있다면) 제어 입력 u에 대한 1차식으로 쓰여진다는 점이다. 안전 영역 기반 접근법에서는 이 사실을 활용하여 2차 계획법(quadratic programming 혹은 QP)로 표현된 다음의 최적화 문제의 해(solution)를 구하여, 이를 제어 입력으로 취한다.

$$\begin{aligned} u_1^\star(x) := \operatorname{argmin}_{u,\delta} & \left(\frac{1}{2} u^\top H(x) u + p\delta^2 \right) \\ \text{subject to } & L_f V(x) + L_g V(x) u \leq -\overline{\alpha}_3\left(V(x)\right) + \delta \\ & L_f b(x) + L_g b(x) u - \alpha_3\left(h(x)\right) \leq 0 \end{aligned} \tag{7}$$

여기서 $H(x)$와 p는 가중 행렬(weighting matrix)로써 설계 변수가 된다. 한편 $b(x)$가 “제어” 장벽함수이기 때문에, 위 수식의 제일 아랫줄에 해당하는 제약조건을 만족하는 u는 정의에 의해 항상 존재한다. 하지만 그러한 u가 제어 리아프노프 함수에 대한 제약조건인 $L_f V(x) + L_g V(x) u \leq -\overline{\alpha}_3(V(x))$를 만족하리란 보장은 전혀 없기 때문에, 전체 QP 문제의 해의 존재성을 항상 보장하고자 추가적인 변수인 δ 를 리아프노프 함수에 대한 제약 조건에 도입하게 된다(따라서 식 (7)의 최적화 문제는항상 해가 항상 존재한다.) 최적화 이론에 따르면 2차 계획법은 전역적인 해가 존재하며 상대적으로 매우 빠른 시간 내에 풀 수 있다는 장점이 있다.

사실 최적화 문제의 해를 연속 시간에서 정의된 식 (2)의 비선형 시스템의 제어 입력으로 바로 적용할 경우, $u_1^*(x)$가 x에 대해 연속한지는 매우 중요하다. (비선형 시스템/제어 이론을 잘 알고 있는 독자라면, 함수의 연속성 혹은 립시츠 조건이 비선형 시스템의 해의 존재성과 유일성에 어떤 영향을 미치는지 잘 알고 있으리라 생각한다.) 다행스럽게도 우리는 식 (7)의 결과물인 $u_1^*(x)$가 (국소적으로) 립쉬츠 연속임을 보장할 수 있는데, 이에 대한 자세한 설명은 [2]에서 확인 가능하다.

한편 2차 계획법으로 제어기를 설계하는 관점에서 볼 때, 위와 같은 해석의 복잡성을 희생해서라도 좀 더 다양한 제약조

건들을 비교적 쉽게 다뤄볼 수 있다는 점에서 설계 상 큰 이점이 있다. 예를 들어 u의 크기에 대한 추가적인 조건을 고려하고 싶다면, 이를 QP의 제약조건으로써 식 (7)에 추가할 수 있을 것이다(물론 이 경우 올바른 해석을 위해서는 2차 계획법으로 기술된 최적화 문제가 항상 feasible한지부터 확인해봐야 할 것이다). 이러한 설계 상 장점으로 인해, 2차 계획법 기반 안전성 보장 제어는 휴머노이드 로봇과 같이 제어 목표가 많고 상태 변수의 차원도 높은 시스템에 대한 다목적 제어(multi-objective control) 기법으로 주로 활용된다.

한편 기존 시스템을 안정하게 하는(하지만 안전성을 고려하지 않은) 제어 입력 $k_{pre}(x)$를 가지고 있다면, 식 (7)과 유사하지만 조금 다른 방법으로 기술된 다음의 최적화 문제를 풀어서, 기존 제어 입력을 안전성을 고려하여 수정해볼 수도 있을 것이다.

$$u^{\star}(x) \coloneqq \mathrm{argmin}_u \quad \frac{1}{2}||u - k_{pre}(x)||^2 \quad \text{subject to } L_f h(x) + L_g h(x) u \geq -\alpha_3(h(x))$$

2. Sontag’s universial formula 기반 안전성 보장 제어

불안전 영역 기반 접근법에서는 앞서 소개한 장벽함수의 여러 특징들을 이용하여, 잘 알려진 Sontag’s universial control law[5]를 활용한 제어기 설계를 제안한다. 기본적으로 Sontag’s universial control law는 주어진 제어 리아프노프 함수 $V(x)$에 대하여, 다음과 같이 정의된다.

$$u = \kappa(\gamma, L_f V(x), (L_g V(x))^T) \tag{8}$$

여기서 함수 κ는 다음의 형태를 가진다.

$$\kappa(\gamma, a, b) \coloneqq \begin{cases} -\dfrac{a + \sqrt{a^2 + \gamma ||b||^4}}{b^\top b} b, & if \ \ b \neq 0 \\ 0, & otherwise \end{cases}$$

식 (8)의 제어기는 식 (2)의 비선형 시스템에 적용되었을 때 전체 폐루프 시스템의 원점(origin) $x = 0$을 점근적으로 안정하게 만든다는 사실이 알려져 있다.

앞서 소개한 대로 불안전 영역 중심 접근법의 가장 큰 장점은, 제어 장벽함수의 형태와 조건이 제어 리아프노프 함수와 매우 유사하다는 점이다. 이를 활용하여 따라서 만약 식 (8)에서 제어 리아프노프 함수 $V(x)$를 제어 장벽함수 $\beta(x)$로 대체하여 제어하게 되면 아래의 결과를 얻는다.

정리4 [Theorem 2, 4] : 만약 식 (2)의 비선형 시스템에 대하여 제어 장벽함수 $\beta(x)$가 주어졌을 때, 식 (2)와

$$u = \kappa(\gamma, L_f\beta(x), (L_g\beta(x))^T)$$

로 구성되는 폐루프 시스템은 안전하다. 그리고

$$\overline{\mathbb{R}^n \setminus (\mathcal{D} \cup \breve{\mathcal{C}})} \cap \overline{\mathcal{D}} = \emptyset$$

를 추가로 만족한다면, $x(0) \in \mathbb{R}^n \setminus \mathcal{D}$인 전체 시스템은 전역적으로 안전하다.

정리 4의 마지막 조건이 매우 이상하게 보일 수 있는데, 사실 이 조건은 개념적으로 단순하다. 앞선 논의에서 (정의 8에서 정의된) $\breve{C}$가 안전 영역 기반 접근법에서의 C에 대응되는 집합이라고 하였음을 다시 기억해보자. 이상적으로 $\breve{C}$가 D의 여집합임을 가정한다면, 정리 4의 마지막 조건은 자연스럽게 만족된다.

한편 $\beta(x)$는 안전성에 대한 함수이기 때문에, 정리 4의 제어기만으로는 $x = 0$의 점근적 안정성을 얻을 수는 없다. 대신 제어 리아프노프 함수와 제어 장벽함수는 그 성질이 매우 유사하기 때문에, 이 두 함수의 성질들을 합쳐놓아서 자연스러운 확장된 형태인 다음의 제어 리아프노프-장벽함수를 정의할 수 있다.

정의9 [2] : 만약 식 (2)의 비선형 시스템에 대하여 아래로 유계하며 미분 가능한 함수 $W(x)$가 다음 조건들을 모두 만족한다면, 이러한 함수를 제어 리아프노프-(D-)장벽함수라고 한다.

(a) $W(x) > 0, \forall x \in \mathcal{D}$,
(b) $L_f W(x) \leq 0, \forall x \in \{z \in \mathbb{R}^n \setminus \mathcal{D} : L_g W(z) = 0\}$
(c) $\breve{\mathcal{C}} \coloneqq \{x \in \mathbb{R}^n : W(x) \leq 0\} \neq \emptyset$.
(d) $\overline{\mathbb{R}^n \setminus (\mathcal{D} \cup \breve{\mathcal{C}})} \cap \overline{\mathcal{D}} = \emptyset$

만약 처음부터 제어 리아프노프-장벽 함수가 잘 정의되어 있다면, Sontag’s universial formula를 그대로 적용하여 다음의 결과를 얻는다.

정리4 [Theorem 3, 4] : 만약 식 (2)의 비선형 시스템에 대하여 제어 리아프노프-장벽함수 $W(x)$가 주어졌을 때, 식 (2)와

$$u = u_3^{\star}(x) \coloneqq \kappa\left(\gamma, L_f W(x), \left(L_g W(x)\right)^{\top}\right) \tag{9}$$

로 구성되는 폐루프 시스템에 대하여 다음이 성립한다.
(a) $u_{SUF}(x)$는 $x = 0$에서 연속적이다.
(b) $x = 0$는 전역적으로 점근적 안정하다.
(c) 전체 폐루프 시스템은 안전하다.

여기까지 읽은 독자라면 아마 “제어 리아프노프 함수와 제

어 장벽함수가 따로 주어져 있으면, 이를 통해서 제어 리아프노프-장벽 함수를 만들어낼 수 있을까?"에 대한 궁금함이 있을 것이다. 가장 간단한 접근법은 두 함수를 합하는 것일텐데, [4]에서는 아래와 같은 구조로 제안한다.

$$W(x) = V(x) + \lambda\beta(x) + \upsilon \tag{10}$$

여기서 $\lambda > 0$와 $\upsilon > 0$은 설계 변수이다. 하지만 단순하게 제어 리아프노프 함수와 제어 장벽함수를 식 (10)과 같이 그대로 더한다면 함수값 $W(x)$가 최소가 되는 지점이 $x=0$이 아닐 수 있다는 큰 문제가 발생한다. 이 경우 Sontag's universal formula를 이용하여 제어를 하게 될 경우 시스템의 상태 변수가 $x=0$이 아닌 엉뚱한 점(정확히는 $W(x)$의 함수값이 최소인 곳)으로 수렴해갈 가능성이 존재한다.

이러한 문제를 해결하려면 $\beta(x)$에 대한 조건을 강화하는 것이 필요하다. 즉, $\beta(x)$가 적어도 안전 영역에서는 무시될만큼 작은 크기의 음수값으로 계산된다면, 우리는 안전 영역에서 다음의 근사식을 얻을 것이다.

$$V(x) + \lambda\beta(x) \approx V(x)$$

따라서 식 (9)는 적어도 불안전 영역 밖에서는 기존 제어 리아프노프 함수와 유사하게 역할할 것이라 기대할 수 있다. 이러한 아이디어를 조금 더 구체화하면 다음의 정리를 얻는다.

정리 5 [Proposition 3, 4] : 식 (2)의 비선형 시스템에 대하여 다음 조건을 모두 만족하는 제어 리아프노프 함수 $V(x)$, 제어 장벽함수 $\beta(x)$, 어떠한 옹골 집합(compact set) B가 존재한다고 하자.

(a) $c_1\|x\|^2 \le V(x) \le c_2\|x\|^2,\ \forall x \in \mathbb{R}^n$
(b) $\beta(x) = -\epsilon < 0,\ \forall x \in \mathbb{R}^n \setminus \mathcal{B},$
(c) $\mathcal{D} \subset \mathcal{B},\ 0 \notin \mathcal{B}$
(d) $L_f W(x) \le 0, \forall x \in \{z \in \mathbb{R}^n \setminus \mathcal{D}: L_g W(z) = 0\}$

그리고 다음과 같이 정의된 완화된 불안전 영역을 생각하자.

$$\mathcal{D}_{relaxed} \coloneqq \{x \in \mathcal{X}: W(x) > 0\} \supset \mathcal{D}$$

이 때 식 (2)의 비선형 시스템과 식 (9)-(10)의 제어기로 구성된 전체 폐루프 시스템은 (새로운 불안전 영역 $D_{relaxed}$에 대하여) 안전하며 $x=0$를 점근적으로 안정하도록 하는 λ와 υ이 항상 존재한다.

정리하자면, 기존 제어 리아프노프 함수에 붙이기 위해서는 제어 장벽함수 $\beta(x)$을 좀 더 잘 설계해야 한다. 하지만 매번 이 조건을 만족하는 장벽함수를 찾기가 쉽지 않을 수 있고, 또 복수의 장벽함수를 더하고자 하면 문제가 더 복잡해질 것이라 예상한다. 이와 관련해서 참고 문헌[4]에서는 정리 5의 조건 (b)를 만족하지 않는 제어 장벽함수를 잘 가공하여 해당 조건을 만족하도록 수정하거나, 복수의 장벽함수를 더하는 법에 대해서도 소개하고 있다. 따라서 이러한 기법들에 관심 있는 독자라면 살펴봐도 좋을 것 같다.

V. 결론 및 고찰

본 고에서는 제어 시스템의 안전성 보장 문제를 소개하고, 제어 장벽함수를 활용한 두 갈래의 서로 다른 접근법들에 대해 간략히 비교해보았다. 이 중 안전 영역 기반 접근법에서는 두 종류의 장벽함수를 소개하였으며, 이들을 이용한 2차 계획법 기반의 안전성 보장 제어 기법에 대해 살펴보았다. 한편 불안전 영역 기반 접근법에서는 리아프노프 함수와 유사한 특징을 보이는 제어 장벽함수를 도입하였으며, 이 덕분에 Sontag's universal formula와 같이 잘 알려진 비선형 제어 기법을 이용한 또 다른 안전성 보장 제어 기법을 도출해내었다.

최근에는 작은 크기의 외란에 대한 강인 안정성을 논의하기 위한 입력-상태 안정성(input-to-state stability)[6,7]과 모델 불확실성에 대한 강인성[8]과 관련된 연구들이 발표되고 있다. 최근에는 A* 알고리즘 혹은 RRT* 알고리즘과 같은 경로 생성 알고리즘, 강화학습, 모델 예측 제어[9] 등에도 제어 장벽함수를 적극 도입하려는 연구 결과들이 많이 제안되어 오고 있는데, 상당히 흥미진진한 결과들이 많으니 관심이 있다면 살펴봐도 좋을 것 같다.

마지막으로 필자의 노력 부족으로 그림이나 예시를 통한 설명이 부족하였음에 죄송함을 전하며, 그럼에도 본 고가 평소 제어 장벽함수에 궁금하였거나 제어 이론을 사랑하는 많은 독자들에게 부디 도움이 되기를 바라본다.

REFERENCES

[1] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," *Proceedings of European Control Conference*, pp. 3420-3431, 2019.

[2] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," *IEEE Transactions on Automatic Control*, vol. 62,

no. 8, pp. 3861-3876, 2017.

[3] P. Weiland, F. Allgower, "Constructive safety using control barrier functions," *Proceedings of IFAC Symposium on Nonlinear Control Systems*, pp. 473-478, 2007.

[4] M. Z. Romdlony, B. Jayawardhana, "Stabilization with guaranteed safety using control Lyapunov-barrier function," *Automatica*, vol. 66, pp. 39-47, 2016.

[5] E. D. Sontag, "A universal construction of Artstein's theorem on nonlinear stabilization," *Systems & Control Letters*, vol. 13, no. 117-123.

[6] S. Kolathaya, A. D. Ames, "Input-to-state safety with control barrier functions," *IEEE Control Systems Letters*, vol. 3, no. 1, pp. 108-113, 2019.

[7] M. Z. Romdlony, B. Jayawardhana, "Robustness analysis of systems' safety through a new notion of input-to-state safety," *International Journal of Robust and Nonlinear Control*, vol. 29m no. 7, pp. 2125-2136, 2018.

[8] Y. Zhang, S. Walters, X. Xu, "Control barrier function meets interval analysis: Safety-critical control with measurement and actuation uncertainties," *ArXiv*, 2021. Available at https://arxiv.org/abs/2110.00915

[9] J. Zeng, B. Zhang, K. Sreenath, "Safety-critical model predictive control with discrete-time control barrier function," *Proceedings of American Control Conference*, 2021.

저자약력

박경훈

- 2018년 서울대학교 전기공학부에서 박사 학위를 수여
- 2018-2019년 서울대학교 자동화시스템공동연구소 박사후 연구원,
- 2019-2021년 한국과학기술연구원 지능로봇연구단에서 박사후 연구원으로 재직
- 2021년 3월부터 서울시립대학교 전자전기컴퓨터공학부에서 재직 중이며, 외란 관측기, 외란 관측기, 안전성 보장 제어, 다목적 제어 등 제어 시스템의 강인성, 안전성, 보안성과 관련된 제어 이론을 연구하고, 이를 로봇 매니퓰레이터, 모바일 로봇, 2족 보행로봇 등의 다양한 로봇 플랫폼에 적용하는 연구