HJB방정식과 최대원리

이 도 성*

경제학 문헌목록 주제분류 C6, C7

I. 서론

동적계획법은 Bellman[2]에 의해 개발된 최적제어문제에 접근하는 당시로는 전혀 새로운 방법이었다 종전까지의 고전적 변분법적 분석방법의 대표적인 결과는 2계 상미분방정식(ODE)의 형태인 Euler-Lagrange 조건이었으며 변분법적 모형을 다양한 제약을 포함하도록 확장시킨 Pontryagin등[16]에 의해 개발된 최적제어의 대표적 결과는 수반함수를 포함한 1계 ODE의 형태인 최대원리(maximum principle)이었다 Bellman의 일차적인 관심은 공학의 피드백제어의 개념을 수학적으로 추상화하고 일반화하는 것이었으므로 출발점을 고정시키는 최적제어모형을 피드백제어분석이 가능하도록 모든 가능한 출발점과 출발시점을 고려하도록 확장시켰으며 그 결과 가치함수의 시간과 상태에 대한 편도함수간의 관계를 최적조건으로 도출하였다 이것이 동적계획법의 대표적인 결과인 1계 비선형 편미분방정식(PDE)인 Bellman 방정식이다 그러나 실용적 결과

* 서강대학교 경제학과 교수

를 중시하였던 Bellman은 수학적 형식에는 그다지 치중하지 않았음을 다음과 같이 표현하고 있다[1)]

"그러나, 앞에서 하던 방식대로, 우리는 먼저 형식을 비논리의 경계까지 미루어두고 그리고 이렇게 용감한 방식으로 얻은 결과의 해석과 입증에 관해 고심한다"

이렇게 개발된 동적계획법의 실용적 측면에 대해서 Lee-Markus[14]는 동적계획법의 Bellman 방정식은 최적의 충분조건을 확인하는 역할 이외에도 일반적인 비선형 최적제어문제의 1, 2차 함수의 근사치를 사용하여 원문제의 근사치적인 해법을 얻는데에 유용하며, 또한 동적계획원리는 수치적 계산에서 닫힌 루프를 얻는데에 유용하다고 평가하고 있다.

반면에 동적계획법의 논리적 근거에 대한 의문도 제기되었다. 이러한 문제의 중심은 Bellman 방정식이 가치함수의 연속미분가능성을 전제로 한다는 점이다 이것이 사실일 경우에는 최대원리는 Bellman 방정식으로부터 즉시 도출할 수 있으며 이 경우 동적계획법의 이론체계가 최대원리보다 더 간명함을 알 수 있다 Pontryagin 등은 이러한 조건이 지켜질 경우에는 최적제어의 조건들과 동적계획의 HJB 방정식이 대등하다는 것을 인식하고 있었으나 이러한 조건이 간단한 문제에서도 지켜지기가 어렵다는 사실을 그들의 저서 초반부에 소개하는 3개의 예를 통해 강조하고 있다[2)] 이러한 이유에서 초기의 동적계획법은 Pontryagin 등에 의해 최적제어문제의 직관적인 설명 방법에 불과하다는 평을 받는다[3)]

"동적계획법은 미분방정식체계로 묘사할 수 있는 것보다 훨씬 더 일반적인 최적제어과정의 필요성에 의해 개발되었다 그러므로 동적계획법은 최대원리보다 더 광범위한 성격을 지닌다 그러나 후자에 비해 이 방법은 그것이 가치있는 설명도구로 성공적으로 사용될 수 있는 모든 경우에도 엄밀한 논리적 근거를 가지고 있지

---

1) Bellman [1967], Introduction to the Mathematical Theory of Control Processes, Academic Press.p.106

2) L S Pontryagin et al [1962], *The Mathematical Theory of Optimal Process*, Wiley pp 69-73

3) 전게서 p 69

않다"

이러한 상반된 평가속에서 동적계획법과 최대원리이론은 각각 분리된 상태에서 독립적으로 연구되었으나 Rockafellar[20]를 위시하여 Clarke[5]등에 의해 70년대에 시작된 비미분분석과 뒤이어 Crandall[8], Bardi[3]등에 의해 시작된 PDE의 일반화된 해의 개념이 소개되면서 두 방법의 대등한 관계와 나아가서 고전적 변분법과의 대등한 관계가 밝혀짐으로써 대등한 분석수단으로서 인식되기 시작하였다

이러한 흐름을 파악하기 위해서 고전적 변분법의 발전과정을 먼저 살펴보는 것이 필요하다 고전적인 변분법의 기본이 되는 Euler-Lagrange 조건은 상태공간이 n차원일 경우 n 개의 2계 ODE 의 형태이다[4] 이를 Hamilton은 Legendre 변환을 사용하여 2n 개의 좀더 풀기 쉬운 1계 ODE 로 변환시켰으며 이것을 역학에서는 해밀톤체계(Hamiltonian sys- tem)라 부른다 통상 ODE 보다 PDE 가 풀기 어렵지만 그 반대가 사실인 경우도 있다 Jacobi 는 Hamilton의 ODE체계를 풀 수 있는 하나의 PDE를 제시하며 이것이 오늘날의 Hamilton-Jacobi 방정식이다.[5] 이들은 변분법적 문제의 필요조건을 제시하고 있지만 Bellman은 최적제어의 필요조건을 PDE를 사용하여 표현하고 있으며 비미분을 포함하는 일반적인 상황에서는 변분법과 최적제어가 대등한 문제구성의 형식이라는 것이 알려지면서 Bellman방정식도 Hamilton-Jacobi 방정식과 동일한 효력을 가졌다는 점에서 오늘날 함께 Hamilton-Jacobi-Bellma(HJB)방정식으로 부른다 그러나 최적제어 문제의 가치함수가 일반적으로 연속미분가능하지 않기 때문에 Bellman 당시에는 최적제어문제를 HJB 방정식을 통해 분석하는 것에 대한 엄밀성에 대한 의문이 제기되었으며 Pontryagin 등은 이 문제를 다시 수반함수를 포함한 2n 개의 ODE를 사용하는 최대원리로 나타낸다 이것은 형식상으로는 변분법의 해밀톤체계로의 회귀를 의미하지만 분석대상인

4) Sussman and J C Willems [1997], 300 Years of Optimal Control From the Brachystochrone to the Maximum Principle, *IEEA Control Systems Magazine*, v 17 no 3, pp32-44 은 변분법과 최대원리의 발전과정을 자세히 소개하고 있다

5) V I Arnold [1989], *Mathematical Methods of Classical Mechanics second edition,* Springer pp 260-266

문제가 당시로는 변분법보다 다양한 제약을 미분가능성의 가정이 없이 다룰 수 있는 최적제어문제라는 점에서 이러한 형식상의 회귀는 내용면에서의 큰 진전으로 인식되었다 그러나 Sussman[18], Sussman-Willems[19]는 Hamilton 이나 그 이후의 Weierstrass가 Legendre변환[6]에 집착하지 않고 미극대 Hamiltonian을 사용하였다면 최대원리는 훨씬 일찍이 발견되었을 것이라는 주장을 함으로써 고전적인 변분법의 결과들이 최대원리에 지근의 거리에 있었음을 강조한다 이와 같이 동일한 고전적 변분법 문제를 기본으로 하는 역학, 최적제어, 동적계획 및 이와 관련된 분석수단인 ODE 와 PDE 는 변분법의 긴 역사만큼 많은 상호작용을 하고 있다[7]

특히 비미분분석과 PDE의 일반환된 해의 개념이 등장하면서 동적계획법과 HJB 방정식의 엄밀한 기초가 확립되었고 최적제어의 분석과 아울러 미분게임의 연구가 활발해지면서 동적계획법의 중요성은 더한층 커지게 되었다 최적제어의 기법을 응용하고 있는 미분게임모형에서는 최적제어와는 달리 시간에만 의존하는 열린 루프전략과 시간과 상태에 의존하는 닫힌 루프전략이 다른 정보구조를 나타내고 있으므로 각각 다른 균형으로서의 역할을 한다 반면에 단독 의사결정인 최적제어문제에서는 열린 루프와 닫힌 루프 제어는 계산 범위와 방법의 차이일 뿐이며 시작점이 주어지면 같은 최적경로를 실현한다 따라서 단독 최적제어문제에서는 동적계획법의 사용이나 이에 따른 피드백 즉 닫힌 루프제어의 계산은 필수적인 것은 아니지만 미분게임에서는 상태에 대한 정보가 있는 게임구조의 분석에 필수적이다 이러한 이유에서 본고는 동적계획법의 다양한 형태의 흐름을 개관하고 최대원리와의 비교결과들을 소개하며 간단한 예시를 병행한다.

본고는 먼저 동적계획모형을 소개하고 이 모형에서 사용되는 다양한 HJB 방정식의 관행을 소개한다 다음으로 주어진 HJB 방정식을 최대원리의 ODE 로 전환시키는 PDE 분야의 잘 알려진 특성방법을 소개한다 다음으로 이러한

6) Sussman and J C Willems [1997], 300 Years of Optimal Control From the Brachystochrone to the Maximum Principle, *IEEA Control Systems Magazine, v 17 no 3*, p39 참조

7) 이 관계에 관한 상세한 설명은 J Yong and X Y Zhou [1999], *Stochastic Controls Hamiltonian Systems and HJB Equations*, Springer pp 217-232 을 참조할 것

방법을 통해 두 접근법의 대등성을 감응도관계를 통해서 보며 비미분문제를 포함할 경우 약화되는 감응도관계에 관한 최근의 결과를 소개한다 끝으로 비미분 및 불연속 상황을 포함하는 일반화된 해의 개념을 소개한다

Ⅱ. 동적계획모형

동적계획법은 최적제어 문제에 대해 가치함수라는 도구를 사용하여 HJB 방정식에 접근한다는 데에 특징이 있다 최적제어문제 P 를

$$(P)\quad \begin{cases} \max \displaystyle\int_0^T L(t,x(t),u(t))dt + g(x(T)) \\ x(t) = f(t,x(t),u(t)), \qquad x(0)=x_0 \\ u(t) \in U(t) \end{cases}$$

라 하고 문제 P_{tx} 를

$$(P_{tx})\quad \begin{cases} \max \displaystyle\int_t^T L(s,x(s),u(s))ds + g(x(T)) \\ x(s) = f(s,x(s),u(s)), \qquad x(t)=x \\ u(s) \in U(s) \end{cases}$$

라 할 때 가치함수를 $V(t,x) = \max(P_{tx})$로 정의하면 임의의 $t \in [0,T), t+h \in (t,T), x \in X$ 에 대해

$$\begin{aligned} V(t,x) = \max & \left\{ \int_t^{t+h} L(s,x(s),u(s))ds + V(t+h, x(t+h)) \right\} \\ & \dot{x}(s) = f(s,x(s),u(s)), \quad x(t)=x \\ & u(s) \in U(s) \end{aligned}$$

이 성립한다는 것이 동적계획원리(DPP)이고 이 가치함수가 미분가능할 경우 성립하는 다음 PDE

$$V_t(t,x) + \max_u [V_x(t,x) f(t,x,u) + L(t,x,u)] = 0$$
$$V(T,x) = g(x)$$

를 HJB 방정식이라 한다.

Ⅲ. HJB 방정식의 제형태

1. 극대극소간의 부호의 문제

최적제어는 다양한 분야에서 연구 및 사용되고 있으며 그 분야의 관심에 따라 다양한 형태를 취한다 먼저 목적함수가 극대와 극소의 형태로 나뉜다. 고전적 미분 상황에서는 원칙적으로 반대부호를 취함으로써 두 형태의 문제간의 변환이 간단히 이루어지지만 목적함수로부터 파생되는 함수 및 변수들이 많고 비미분의 상황에서는 이러한 변환에 적지 않은 혼란이 따르게 되므로 여기서는 부호에 따른 최적조건들을 각각 별도로 정리하기로 한다 극대화 문제의 가치함수를 $V(t,x)$ 라 하고 해밀토니언을 H 라 하고 이 경우의 수반변수를 p 라 하자 그러면 최적상황을 나타내는 최적제어의 최대원리와 동적계획법의 HJB 방정식과 감응도관계는 다음과 같다

$$\max \int L dt + g \quad x = f, \ x(T) \in C$$

$$\exists (\lambda, p) \neq 0, \lambda \geq 0$$

$$H = \max[pf + \lambda L]$$
$$-\dot{p} = H_x$$
$$(p(0), -p(T)) \in \lambda \nabla(-g) + N_C$$

$$\phi_t + \max[\phi_x f + L] = 0, \ \phi(T) = g$$
$$(H, p) \in co\partial V$$

또한 문제가 극소화로 주어졌을 경우에는 원칙적으로 목적함수의 부호를 반대로 하면 극대화로 변환이 가능하지만 다른 관행들의 복합적인 영향으로 실

제로 사용되는 최적조건들은 다음 형태이다

$$\min \int L dt + g \quad x = f,\ x(T) \in C$$

$$\exists (\lambda, p) \neq 0, \lambda \geq 0$$

$$\begin{aligned} &H = \max[pf - \lambda L] \\ &-\dot{p} = H_x \\ &(p(0), -p(T)) \in \lambda \nabla g + N_C \\ &\phi_t + \min[\phi_x f + L] = 0,\ \ \phi(T) = g \\ &(H, p) \in co\partial(-V) \end{aligned}$$

또한 비미분분석에서는 부호를 바꿔서 하부미분을 취한 결과와 본래의 하부미분의 부호를 바꾼 집합이 다를 수가 있으며 PDE의 일반화된 해는 방정식의 부호를 바꿀 경우 달라지는 것이 가능하므로 부호의 문제는 처음부터 극대나 극소의 한 형태에 따라 일관성을 유지하는 것이 안전하다

2 HJB방정식의 표현에 따른 제관행

동적계획법의 HJB 방정식의 표현에 있어서도 저자에 따라 해밀토니언이나 가치함수를 다르게 정의하여 동일한 문제의 HJB 방정식이 달라지는 일이 있다 그 몇가지 예들을 여기에 정리한다 먼저 주어진 문제가

$$\begin{aligned} &\min_u \int_0^T L(t, x(t), u(t))dt + g(x(T)) \\ &s.t.\ \ x(t) = f(t, x(t), u(t)),\ \ x(0) = x_0 \end{aligned}$$

일 경우 이 문제의 DP 의 기본 관계식인 HJB 방정식

$$\phi_t + \min_u \{\phi_x \cdot f + L\} = 0$$

은 해밀토니언의 정의에 따라 부호의 관습이 달라진다 즉

$H=\max_u\{-p\cdot f-L\}$로 정의하면 $\phi_t-H(t,x,\phi_x)=0$ 가 되고 또 $H=\max_u\{p\cdot f-L\}$로 정의하면 $\phi_t-H(t,x,-\phi_x)=0$ 가 된다.

또한 가치함수의 변수를 선택하는 관행도 다양하다. Vinter 등[20]이 사용하는 통상적인 가치함수는 시작시점을 변수로 갖도록

$$V(t,x)=\min_u\left[\int_t^T L(x(s),u(s),s)ds+g(x(T))\right]$$
$$s.t.\ \ x(s)=f(s,x(s),u(s)),\quad x(t)=x$$

와 같이 정의한다 이 경우 $(0,x)$ 로 주어진 문제의 최적경로가 $\bar{x}$ 라 하면 임의의 $t\in(0,T)$ 에 대해 $(t,\bar{x}(t))$ 로 주어진 문제의 최적경로도 $\bar{x}$ 의 해당부분이라는 것을 DPP 로 알 수 있으므로 문제가 충분히 미분가능한 경우를 상정하면

$$\int_t^{t+h} L(s,\bar{x}(s),u)ds+V(t+h,\bar{x}(t+h))\geq V(t,\bar{x}(t))\quad \forall u$$

이고 등식은 $u=\bar{u}$ 에서 성립하므로 이로부터 모든 $v\in U(t)$에서

$$-V_t(t,\bar{x}(t))-V_x(t,\bar{x}(t))f(t,\bar{x}(t),\bar{u}(t))-L(t,\overline{x(t)},v)\leq 0$$

가 성립한다 등식은 $u=\bar{u}$ 에서 성립하므로

$$-V_t+sup\{-V_xf(t,\bar{x}(t),\bar{u}(t))-L(t,\overline{x(t)},v)\}=0$$

해밀토니언을 $H=\max\{p\cdot v-L\}$로 정의하면 윗 식은

$$-\phi_t+H(t,x,-\phi_x)=0,\ \ \phi(T,x)=g(x)$$

이 된다

다음으로 Sagan 등[17]은 다음과 같이 가치함수를 시작점을 고정시키고 최종점을 변수로 갖는 것으로 정의한다 즉

$$V(t,x)=\min_u \left[\int_0^t L(x(s),u(s),s)ds+g(x(t)) \right]$$
$$s.t.\ \ \dot{x}(s)=f(s,x(s),u(s)),\quad x(t)=x$$

이 경우 (T,x) 로 주어진 문제의 최적경로가 $\bar{x}$ 라 하면 임의의 $t\in(0,T)$ 에 대해 $(t,\bar{x}(t))$ 로 주어진 문제의 최적경로도 $\bar{x}$ 의 해당부분이라는 것을 DPP로 알 수 있으므로 문제가 충분히 미분가능한 경우를 상정하면 모든 u 에 대해

$$V(t,\bar{x}(t))+\int_t^{t+h} L(s,\bar{x}(s),u)ds\geq V(t+h,\bar{x}(t+h))$$

이고 등식은 $u=\bar{u}$ 에서 성립하므로 이로부터 모든 $v\in U(t)$에서

$$V_t+V_x f(t,\bar{x}(t),\bar{u}(t))-L(t,\bar{x}(t),v)\leq 0$$

가 성립한다 해밀토니언을 $H=\max\{p\cdot v-L\}$ 로 정의하면 윗식은

$$V_t+H(t,x,V_x)=0,\quad V(0,x)=g(x)$$

이 된다.

Bardı 등[3]은 다음과 같이 가치함수를 시작점의 상황과 끝점의 시간을 변수로 갖는 것으로 정의한다 즉

$$V(t,x) = \min_u \left[\int_0^t L(x(s), u(s), s)ds + g(x(t)) \right]$$
$$s.t.\ \ \dot{x}(s) = f(s, x(s), u(s)),\quad x(0) = x$$

이 경우 (T, x) 로 주어진 문제의 최적경로가 $\bar{x}$ 라 하면 임의의 $t \in (0, T)$ 에 대해 (t, x) 로 주어진 문제의 최적경로도 $\bar{x}$ 의 해당부분이라는 것을 DPP 로 알 수 있으므로 문제가 충분히 미분가능한 경우를 상정하면

$$V(t,x) + \int_t^{t+h} L(s, \bar{x}(s), u)ds \geq V(t+h, x) \quad \forall u$$

이고 등식은 $u = \bar{u}$ 에서 성립하므로 이로부터 모든 $v \in U(t)$에서

$$V_t + V_x f(t, \bar{x}(t), \bar{u}(t)) - L(t, \overline{x(t)}, v) \leq 0$$

가 성립한다. $H = \max\{p \cdot v - L\}$ 로 정의하면 윗식은

$$V_t + H(t, x, V_x) = 0,\quad V(0, x)$$

가 된다.

Ⅳ. 확인정리

확인정리(verification theorem)란 주어진 최적제어문제의 한 제어방법이 최적인가를 확인하는 조건이다. HJB 방정식의 일반화된 해가 등장하기 전까지는 HJB 방정식의 전통적인 역할은 이러한 임무를 수행하는 것이었으며 실제로 이러한 기능을 사용하여 추측한 가치함수의 형태를 확인하기도 한다. 이와 관련된 주요정리를 소개하면 다음과 같다.

정리(Vinter[21, p31]) $(\bar{x}, \bar{u})$ 가 문제 P의 제약을 만족시키며 $\phi \in C^1$ 이 존재하여 HJB 방정식을 만족시키며 또한

$$\begin{aligned}&\phi_x(t, \bar{x}(t))f(t, \bar{x}(t), \bar{u}(t)) + L(t, \bar{x}(t), \bar{u}(t))\\&= \min_u\{\phi_x(t, \bar{x}(t))f(t, \bar{x}(t), u)) + L(t, \bar{x}(t), u)\}\end{aligned}$$

을 만족시키면 $(\bar{x}, \bar{u})$ 는 문제 P 의 최소점이며 $\phi(0, x_0)$ 는 최소치이다

이 정리에서 확인함수 ϕ 의 당연한 대상은 가치함수 V 이며 $V \in C^1$ 일 경우 V 는 HJB방정식의 해이다 다음 예를 통해서 추측한 형태의 전략이 최적이 되는지를 확인정리를 사용하여 알아보기로 한다 이 예는 본고에서 계속 사용된다

예 시간수평과 초기상태가 각각 $T \in (0, 1)$, $x_0 \geq 0$이라 할 때 최적제어 문제가

$$\max\left[\int_0^T\left(ux - \frac{u^2}{2}\right)dt + x(T)\right] \quad s.t.\ \dot{x} = u,\ x(0) = x_0$$

이다 상수 u 형태의 제어가 최적제어가 되는지를 확인정리를 통해 알아보기로 하자 먼저 상수 중에서 최적 상수를 구하기 위해 상수 u 를 제어로 선택할 경우의 목적함수의 값을 계산하자 이 경우 $u(t) = c \quad \forall t \in [0, T]$ 라고 하면 $x(t) = x_0 + ct$ 가 된다 따라서

$$\begin{aligned}&\int_0^T\left\{c(x_0 + ct) - \frac{c^2}{2}\right\}dt + x_0 + cT\\&= \frac{c^2}{2}T(T-1) + cT(x_0+1) + x_0\end{aligned}$$

이 되어 이 값은 $T > 1$ 일 경우 상계가 존재하지 않으며 $T \in (0, 1)$ 일 경우에

만 최대값이 존재한다. 최대점 c 는

$$c=\frac{x_0+1}{1-T}$$

이다 이 문제는 제어의 선택에 제약이 없기 때문에 $T>1$ 일 경우 해가 존재하지 않으며 이를 위해서는 제어의 제약이 필요하다 다음으로 확인정리를 사용하기 위해서 임의의 (t,x) 에서 시삭할 경우 나머지 기간 $[t,T]$ 동안 상수제어를 택하는 경우의 목적함수의 값은

$$\int_t^T\left\{c(x+c(s-t))-\frac{c^2}{2}\right\}ds+x+c(T-t)$$
$$=\frac{c^2}{2}(T-t)(T+t-1)+c(T-t)(x+1)+x$$

이다 이제 이 값이 항상 극대값을 갖기 위해서는 c의 2차항의 계수가 음이어야 한다 즉

$$(T-t)(T-t-1)<0 \quad \forall t\in[0,T)$$

이 되는 조건은 $T\in(0,1)$ 이다 이를 극대화하는 상수제어 $u(t,x)$ 와 그 경우의 목적함수의 값 $V(t,x)$ 는 각각

$$u(t,x)=\frac{1+x}{1-T+t}, \quad V(t.x)=x+\frac{(1+x)^2}{2(1-T+x)}$$

이 된다 이 전략이 최적인가를 HJB 방정식으로 확인하기로 한다.

$$V_t=\frac{(1+x)^2}{2(1-T+t)^2}, \quad V_x=\frac{1-x(T-t)}{1-T+t}$$

이고 이 V_x 값을 사용하면 $V_x u + ux - \frac{u^2}{2}$ 는 $u = x + V_x$ 에서 극대화되며 이 경우 HJB방정식

$$V_t + \frac{1}{2}(x + V_x)^2 = 0$$

가 성립함을 알 수 있다. 따라서 상수전략은 최적전략이다

V. 열린 루프와 닫힌 루프

최적제어문제에서 제어가 시간에만 의존하는 열린 루프(open loop)와 시간과 상태에 의존하는 닫힌 루프(closed loop)가 있다는 사실은 잘 알려져 있다. 동적최적화문제를 개인의 기본문제로 상정하는 미분게임의 발전과 함께 열린 루프와 닫힌 루프 전략의 균형상의 차이가 분명해짐으로써 닫힌 루프 전략의 계산을 용이하게 하는 동적계획법의 유용성이 더 한층 부각되었다. 개인의 동적최적화문제에서는 초기상태가 주어지면 이 상태로부터 열린 루프 제어를 따라 진행되는 상태경로는 동일한 출발점에서 닫힌 루프 제어를 따라 진행되는 상태경로와 같으므로 두 제어방식의 본질적인 차이는 없고 표현형식의 차이만이 있을뿐이다

물론 최적제어모형의 최대원리로부터 닫힌 루프전략을 계산하는 것도 가능하나 이는 일반적인 경우 다음과 같이 복잡한 과정을 거쳐야 한다 문제 P의 경우 초기상태백터 x_0 가 주어지면 최대원리의 ODE로부터 최적 $x(t, x_0), p(t, x_0)$ 가 결정된다 그 다음 임의의 x, t 에 대해 $x = x(t, x_0)$ 가 되는 x_0 를 역함수정리의 조건이 지켜지면 구할 수 있다. 이를 $x_0(x, t)$ 라 하면

$$u(t, x) = \arg\max_{u \in U(t)} \{ p(t, x_0(x, t)) f(t, x, u) + L(t, x, u) \}$$

이 닫힌 루프제어이다

예 앞의 예를 다시 검토하기로 하자 여기서 $T\in(0,1)$, $x_0\geq 0$ 이고

$$\max\left[\int_0^T\left(ux-\frac{u^2}{2}\right)dt+x(T)\right] \quad s.t.\ \dot{x}=u,\ x(0)=x_0$$

이다 이 문제의 미극대 해밀토니언은 $p\in R, \lambda\geq 0$ 일 때

$$H=pu+\lambda(ux-\frac{u^2}{2})$$

이며 최대조건은 $\lambda u=\lambda x+p$, $\dot{p}=-\lambda u$ 이고 횡단조건은 $C=\{x_0\}\times R$, $N_C=R\times\{0\}$, $\partial g=\{(0,1)\}$ 이므로

$$(p(0),-p(T))\in-\lambda\partial g+N_C=R\times\{-\lambda\}$$

로부터 $p(T)=\lambda$임을 알 수 있다 위의 결과를 정리하면 정규방정식 체계는 $\dot{p}=\lambda x+p$, $\lambda\dot{x}=-\lambda x-p$, $p(T)=\lambda, x(0)=x_0$이다 이제 $\lambda=0$일 경우를 먼저 살펴보자 이 경우 $p(T)=\lambda=0, \dot{p}=-\lambda u=0$ 이므로 $p(0)=p(T)=0$가 되어 $p(t)\equiv 0\ \forall t\in[0,T]$ 가 되어 $(p,\lambda)\neq(0,0)$ 에 위배되므로 불가능하다 그러므로 $\lambda>0$ 이어야 하며 정규방정식 체계를 λ 로 나누어

$$\frac{\dot{p}}{\lambda}=x+\frac{p}{\lambda},\ \dot{x}=-x-\frac{p}{\lambda},\quad \frac{p(T)}{\lambda}=1, x(0)=x_0$$

과 같이 나타낼 수 있으므로 종전의 기호를 사용하며 $\lambda=1$ 로 놓는 것이 편리하다 이에 따라 정규방정식은

$$\begin{pmatrix}\dot{p}\\ \dot{x}\end{pmatrix}=\begin{pmatrix}-1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}p\\ x\end{pmatrix},\quad \begin{pmatrix}p(T)\\ x(0)\end{pmatrix}=\begin{pmatrix}1\\ x_0\end{pmatrix}$$

이다 일반해를 구하면

$$\begin{pmatrix} p(t) \\ x(t) \end{pmatrix} = \begin{pmatrix} c_1 - c_2 + c_2 t \\ -c_1 - c_2 t \end{pmatrix}$$

이며 경계조건을 이용하면

$$c_1 = -x_0, \quad c_2 = \frac{x_0 + 1}{T - 1}$$

이다 이로부터 최적치를 구하면

$$p(t) = \frac{(T - t)x_0 + (1 - t)}{1 - T}, \quad x(t) = \frac{(1 - T + t)x_0 + t}{1 - T}$$

$$u(t) = \frac{x_0 + 1}{1 - T}$$

이와 같이 열린 루프전략은 현재시간 t에 의존하지 않고 초기값 x_0 에 의존하므로 이 의존관계를 $u_O(x_0)$ 로 나타내기로 한다 다음으로 상태식 $x(t)$ 에서 현재상태가 (t, x) 가 되게 하는 초기값 x_0 를 $x(t, x_0) = x$ 가 되는 $x_0(x, t)$ 의 함수형태로 구하면

$$x_0(x, t) = \frac{(1 - T)x - t}{1 - T + t}$$

이며 이 값이 정의되는 영역 즉 $\{(t, x) \mid (1 - T)x - t \geq 0\}$ 에서 닫힌 루프 전략 $u_C(t, x)$ 은

$$u_C(t, x) = \frac{1 + x}{1 - T + t}$$

이 된다. 열린 루프전략에서는 초기값에 따라 일정한 u 값이 선택되었지만 닫힌 루프 전략에서는 현재의 상태와 시간에 따라 u 값이 달라진다. 그러나 일정한 초기값 x_0 에서 출발한 최적경로의 상태와 시간에서 닫힌 루프 전략에 따라 선택되는 전략은 일정하며 이는 $u_O(x_0)$ 와 일치한다. 즉

$$\begin{aligned} u_C(t, x(t, x_0)) &= u_C(t, x_0 + tu_O(x_0)) \\ &= \frac{1 + x_0 + tu_O(x_0)}{1 - T + t} \\ &= \frac{x_0 + 1}{1 - T} = u_O(x_0) \end{aligned}$$

이 성립한다 이제 닫힌 루프전략으로부터 가치함수를 다음과 같이 계산할 수 있다 출발이 (t, x) 인 문제의 가치함수

$$V(t, x) = \int_t^T \left[u(t, x)x(s) - \frac{u(t, x)^2}{2} \right] ds + x(T)$$

에서 최적상태의 경로는

$$x(s) = x + \left(\frac{1 + x}{1 - T + t} \right)(s - t), \qquad s \in [t, T]$$

이므로 이를 대입하여 계산하면

$$V(t, x) = x + \frac{(1 + x)^2 (T - t)}{2(1 - T + t)}$$

를 얻는다. 가치함수의 상태변수에 대한 미분과 수반변수를 비교하면

$$V_x(t, x) = \frac{(T - t)x + 1}{1 - T + t}$$

이며 이 값은 $p(t, x_0(x, t))$ 의 값과 일치함을 확인할 수 있다

Ⅵ. 특성방법

PDE의 일부는 특성(characteristic)곡선을 따라 ODE로 풀 수 있다는 사실은 특성방법이라는 이름으로 PDE분야에서는 잘 알려진 기법이지만 미분게임을 포함한 최적제어의 분석에 그 유용성이 인식된 것은 PDE의 일반화된 해의 이론이 발전되기 시작한 비교적 최근의 일이다 최적제어의 최대원리와 HJB 방정식의 관계를 이해하는 데에는 특성방법이 필수적이므로 이 장에서는 그 주된 내용을 간략히 소개한다 자세한 내용은 Evans[9], Melikyan[15] 및 Courant-Hilbert[7]을 참고하기 바란다

$M \subset R^n$, $\Gamma \subset bdyM$ 이고 $\phi\ M \to R$, $D\phi \cdot M \to R^n$, $g\ \Gamma \to R$ 이고 $F\ M \times R \times R^n \to R$ 이며 여기서 $bdyM$ 은 M 의 경계이고 의 형태이고 $D\phi(x) = (D_{x_1}\phi(x), \ldots, D_{x_n}\phi(x))$ 는 편도함수의 벡터를 나타낸다 이제 다음의 1계 비선형 Cauchy PDE 문제를 보자

$$
\begin{aligned}
&F(x, v, p) = 0, && x \in M \subset R^n \\
&u(x) = g(x), && x \in \Gamma \subset bdyM
\end{aligned}
$$

앞으로의 분석을 위해 $bdyM$ 을 $Z \subset R^{n-1}$ 를 정의역으로 갖는 함수 $\psi\ Z \to R^n$ 를 사용하여 나타내기로 한다 그러면 $g(\psi(z)) = h(z)$와 같이 h 함수를 Z 에 정의할 수 있다

1. 특성방정식

주어진 PDE $F(x, v, p) = 0$ 의 특성방정식은 이 방정식의 독립변수들간에 성립하는 ODE로서 이 경우에는 다음 $2n+1$ 개의 식

$$
\begin{aligned}
x &= F_p \\
\dot{v} &= \langle p, F_p \rangle \\
p &= -F_x - pF_v
\end{aligned}
$$

이며 여기서 각형괄호는 내적을 나타낸다 이 특성방정식들은 다음과 같이 도출된다 독립변수에 포함되어 있는 시간변수와는 별도의 매개변수 $s \in R$ 를 사용하여 임의의 미분가능한 경로 $x(s)$를 설정하고 PDE를 매개변수 s 에 대해 미분하면 $(F_x + F_v p + D^2 v F_p) x(s) = 0$ 를 얻는다 이는 임의의 $\dot{x}(s)$ 에 내해 성립하여야 하므로 PDE 가 충분히 연속미분가능한 해 $\phi(x)$ 를 가질 경우 이 해를 따라

$$
F_x + F_v D\phi + D^2 \phi F_p = 0
$$

이어야 한다 이와 같이 해가 $\phi(x)$ 일 경우 경로 $x(s)$ 를 택하고 이와 함께 다른 변수들의 경로도 $v(s) = \phi(x(s))$, $p(s) = D\phi(x(s))$ 로 정의하면

$$
\begin{aligned}
\dot{v}(s) &= D\phi \dot{x}(s) \\
\dot{p}(s) &= D^2 \phi \dot{x}(s)
\end{aligned}
$$

이 된다 여기서 $D^2\phi$ 는 2계도함수의 행렬이므로 이 복잡한 항을 제거하도록 $x(s)$ 를 선택하는 것이 중요하다 여기서 $x(s) = F_p$ 로 선택하면 $p(s) = D^2\phi x = D^2 \phi F_p = -F_x - F_v D\phi$ 와 같이 $D^2\phi$ 항이 제거된다. 이에 따라 $\dot{v} = Dvx = \langle p, F_p \rangle$ 가 되어 위의 특성방정식 체계가 도출된다 앞의 특성방정식체계의 변수들을 명시하면

$$
\begin{aligned}
\dot{x}(s) &= F_p(x(s), v(s), p(s)) \\
v(s) &= \langle p(s), F_p(x(s), v(s), p(s)) \rangle \\
p(s) &= -F_x(x(s), v(s), p(s)) - p(s) F_v(x(s), v(s), p(s))
\end{aligned}
$$

와 같이 $2n+1$ 개의 함수 $(x(s), v(s), p(s))$ 의 ODE 체계를 얻는다

2. 초기조건과 국지적 해의 도출

이제 이 ODE 의 초기조건을 도출하기로 한다 경계조건을 간접적으로 나타내는 식 $v(\psi(z))=h(z)$ 를 $z\in Z$ 로 미분하면 $D\psi(z)p=Dh(z)$을 얻으며 이 식들을 주어진 PDE와 연립한

$$F(\psi(z), v(\psi(z)), p) = 0$$
$$\begin{pmatrix} \psi_{11}, \dots\dots, \psi_{1n} \\ \\ \psi_{(n-1)1}, \dots, \psi_{(n-1)n} \end{pmatrix} \begin{pmatrix} p_1 \\ \\ p_n \end{pmatrix} = \begin{pmatrix} D_1 h(z) \\ \\ D_{n-1} h(z) \end{pmatrix}$$

로부터 p의 n 성분을 구할 수 있다 이를 위해서는 이 체계의 Jacobian 인

$$\begin{vmatrix} F_{p_1}, \dots\dots\dots, F_{p_n} \\ \psi_{11}, \dots\dots\dots, \psi_{1n} \\ \\ \psi_{(n-1)1}, \dots, \psi_{(n-1)n} \end{vmatrix} \neq 0$$

이어야 한다 이는 $i=1,\dots,n-1$ 에 대해 행벡터 $D_i\psi(\bar{z})$는 Γ 의 $\psi(\bar{z})$에서의 접공간(tangent space)에 속하며 $n-1$ 개의 행벡터 $D_i\psi(\bar{z})$ 는 접공간의 기저를 이루므로 이들과 F_p 가 선형독립이라는 것은 F_p 와 Γ 가 횡단관계에 있음을 의미한다

이 경우 z 가 주어질 경우 음함수 정리에 의해 시초점 $p(z)$가 일의적으로 존재하며 따라서 z 에 대응하는 ODE 의 시초점들은

$$x=\psi(z), \quad v=h(z), \quad p=p(z), \quad z\in Z$$

과 같이 정해진다 따라서 임의의 $z \in Z$ 가 주어지면 이에 대응하는 시초점 $(x(0), v(0), p(0)) = (\phi(z), h(z), p(z))$ 에서 출발하는 특성방정식 ODE의 해 $(x(s), v(s), p(s))$ 가 결정되므로 이를 $x(s, z)$, $v(s, z)$, $p(s, z)$ 와 같이 $(s, z) \in R^n$ 에 의존하도록 나타내기로 하자. 그러면 근방의 임의의 $x \in M \subset R^n$ 가 주어지면 $x(s, z) = x$ 가 되는 $(s, z) \in R^n$ 는 x 의 함수이다 이는 $x(0) = \phi(z)$에서

$$\begin{pmatrix} x_{1t}, \dots\dots, x_{nt} \\ x_{1z_1}, \dots\dots, x_{nz_1} \\ \dots\dots\dots\dots \\ x_{1z_{n-1}}, \dots, x_{nz_{n-1}} \end{pmatrix} = \begin{pmatrix} F_{p_1}, \dots\dots, F_{p_n} \\ \phi_{11}, \dots\dots, \phi_{n1} \\ \dots\dots\dots\dots \\ \phi_{1n-1}, \dots\dots, \phi_{nn-1} \end{pmatrix}$$

이 행렬식이 0 이 아니었으므로 0 이 아닌 근방이 존재하기 때문이다. 따라서 역함수정리에 의해 $x(s, z) = x$ 가 되는 (s, z) 를 국지적으로 $s(x), z(x)$ 의 함수로 니타낼 수 있으므로 이 근방의 x 에 대해 $\phi(x) = v(s(x), z(x))$ 로 정의하면 PDE의 해를 국지적으로 얻는다 지금까지의 논의는 1계 비선형 PDE에 대한 일반적인 특성방법이었다. 다음에서는 1계 비선형 PDE의 특별한 형태인 HJB방정식에 특성방법을 적용하여 최대원리을 도출하기로 한다

3. 특성방법과 최대원리

이 방법의 적용범위는 비선형 1계 PDE를 포함하므로 가치함수가 연속미분가능한 최적제어문제의 HJB 방정식에 적용하면 이 문제의 최대원리의 ODE를 특성방정식으로 얻는다 즉 문제 P 의 HJB 방정식은

$$\begin{aligned} &\phi_t(t, x) + \max{}_u[\phi_x(t, x) f(t, x, u) + L(t, x, u)] = 0 \\ &\phi(T, x) = g(x) \end{aligned}$$

이었다 특성방식의 논의에서 사용하던 독립변수 $x \in R^n$ 는 지금의 최적제어문제에서는 시간성분을 별도로 $(t, x) \in [0, T] \times R^n \subset R^{n+1}$ 와 같이 구분하고

함수값을 나타내는 v 를 ϕ 로 바꾸고 편미분값을 나타내는 p 를 (ϕ_t, ϕ_x) 로 역시 성분을 구분하여 표시하고 이 문제의 해밀토니언을

$$H(t, x, \phi_x) = \max_{u \in U(t)} [\phi_x f(t, x, u) + L(t, x, u)]$$

로 정의하여 HJB 방정식을 나타내면

$$F(t, x, \phi, \phi_t, \phi_x) = \phi_t + H(t, x, \phi_x) = 0$$

가 되며 이 PDE의 특성 ODE 는

$$\begin{aligned} \dot{t} &= 1 \\ \dot{x} &= H_{\phi_x} \\ \phi &= \phi_t + \phi_x H_{\phi_x} \\ \phi_t &= -H_t \\ \phi_x &= -H_x \end{aligned}$$

이다 이 중 2식과 5식이 최대원리의 ODE 이다 최대원리의 횡단조건을 확인하기 위해서 PDE 의 경계를 $\psi\ R^n \rightarrow R^{n+1}$ 인 함수 $\psi(x) = (\psi_1(x), \psi_2(x)) = (T, x)$ 로 나타내기로 하면 $h(x) = \phi(\psi(x)) = \phi(T, x) = g(x)$ 가 된다 이로부터 상황의 초기값을 $(t(0), x(0)) = (T, x)$ 라고 하면 $\phi(0) = \phi(t(0),\ x(0)) = \phi(T, x) = g(x)$ 가 된다 이제 남은 것은 $(\phi_t(0), \phi_x(0))$ 즉 $\phi_t(T, x)$ 와 $\phi_x(T, x)$ 의 값을 구하는 것이다

이는 다음 두 식

$$\begin{aligned} &\phi_t(T, x) + H(T, x, \phi_x(T, x)) = 0 \\ &\phi(\psi(x)) = g(x) \end{aligned}$$

에서 먼저 미극대 해밀토니언의 (T, x) 에서의 극대점을 $u(T, x)$ 로 나타내어 첫 식을 고쳐쓰고 두 번째 식을 미분한 다음의 $n+1$ 개의 연립식

$$\phi_t(T, x) + \phi_x(T, x) f(T, x, u(T, x)) + L(T, x, u(T, x)) = 0$$
$$\phi_t(T, x) D\psi_1(x) + \phi_x(T, x) D\psi_2(x) = \nabla g(x)$$

에 $D\psi_1 = 0$, $D\psi_2 = I$ 를 대입하면

$$\begin{pmatrix} 1 & f \\ 0 & I \end{pmatrix} \begin{pmatrix} \phi_t \\ \phi_x \end{pmatrix} = \begin{pmatrix} -L \\ \nabla g \end{pmatrix}$$

이 되어 $\phi_x = \nabla g$ 와 $\phi_t = -f\nabla g - L$ 을 얻는다 그러면 이 경우의 횡단조건 $\phi_x = \nabla g$ 임을 확인할 수 있다 물론 이와 같은 도출은 충분한 미분가능성을 가정한 것이므로 이 방법을 그대로 일반적인 상황에 적용할 수 있는 것은 아니다

4 특성방법의 예시

앞의 예의 HJB 방정식 $\phi_t + \max\{\phi_x f + L\} = 0$ 로부터 이를 u 로 극대화하면 $\phi_t + \frac{1}{2}(x + \phi_x)^2 = 0$, $\phi(T, x) = x$ 임을 알 수 있다 따라서 이 경우 $n = 1$ 이므로 HJB방정식은

$$F(t, x, \phi, \phi_t, \phi_x) = \phi_t + \frac{1}{2}(x + \phi_x)^2 = 0$$

이며 다음 5개의 특성 ODE

$$
\begin{aligned}
\dot{t} &= F_{\phi_t} = 1 \\
\dot{x} &= F_{\phi_x} = x + \phi_x \\
\dot{\phi} &= \phi_t F_{\phi_t} + \phi_x F_{\phi_x} = \phi_t + (x+\phi_x)\phi_x \\
\dot{\phi}_t &= -F_t - F_\phi \phi_t = 0 \\
\dot{\phi}_x &= -F_x - F_\phi \phi_x = -x - \phi_x
\end{aligned}
$$

를 갖는다. 경계조건은 $\phi(T,x)=x$ 이므로 종료시점 $T\in(0,1)$ 값이 주어지면 경계는 $\Gamma=\{(T,x) \mid x\in R\}$ 와 같이 시간 상태의 평면에서 시간축 좌표를 T 로 갖는 상태직선이므로 1차원의 다양체이다 따라서 Γ 를 $\psi\ R\rightarrow R^2$ 함수 $\psi(x)=(\psi_1(x),\psi_2(x))=(T,x)$ 로 나타내기로 하면 Γ 에 속하는 임의의 (T,x) 에서는 HJB 방정식과 경계조건이 동시에 성립하므로 다음의 두 등식

$$
\begin{aligned}
&\phi_t + \frac{1}{2}(x+\phi_x)^2 = 0 \\
&\phi(\psi_1(x),\psi_2(x)) = \phi(T,x) = x
\end{aligned}
$$

이 성립한다 두 번째식을 x 로 미분하면 $\phi_t\psi_1{}' + \phi_x\psi_2{}' = 1$ 이 되므로 이 식과 첫 식을 함께 풀면 $\phi_t = -\frac{1}{2}(x+1)^2$, $\phi_x=1$ 을 얻는다 따라서 Γ 에 속하는 임의의 점 (T,x) 에서 시작하는 특성 ODE 의 경로를 $(t(s),x(s),v(s),p_1(s),p_2(s))$ 라 하면 시작점의 값은

$$
(t(0),x(0),v(0),p_1(0),p_2(0)) = (T,x,x,-\frac{1}{2}(x+1)^2,1)
$$

이다 위의 두 ODE를 풀면 일반해는

$$
\begin{pmatrix} p_2(s) \\ x(s) \end{pmatrix} = \begin{pmatrix} c_1 - c_2 + c_2 s \\ -c_1 - c_2 s \end{pmatrix}
$$

또한 $x(0)=x$ 이므로 이 경계조건을 만족시키는 값들은

$c_1 = -x$, $c_2 = -x-1$ 이 되고 x 에서 출발하는 특성 ODE 의 해는

$$\begin{pmatrix} p_2(s) \\ x(s) \end{pmatrix} = \begin{pmatrix} 1-(x+1)s \\ x+(x+1)s \end{pmatrix}$$

이 된다 따라서 가치함수인

$$\begin{aligned} &\phi(t(0), x(0)) - \phi(t(t-T), x(t-T)) \\ &\quad = \int_{-(T-t)}^{0} [p_1(s) + (x(s) + p_2(s))p_2(s)]ds \end{aligned}$$

로부터

$$\begin{aligned} x - \phi(t, x) &= \int_{-(T-t)}^{0} [p_1(s) + (x(s) + p_2(s))p_2(s)]ds \\ &= -\frac{1}{2}(1+x)(t-T) + (1+x)(t-T) + \frac{1}{2}(1+x)^2(t-T)^2 \end{aligned}$$

이 된다 이 식을 정리하면

$$\phi(t, x) = x + \frac{(1+x)^2(T-t)}{2(1-T+t)}$$

가 되어 앞의 결과와 일치한다. 또한 이 가치함수를 사용하여 닫힌 루프 최적 전략을 구하면 앞의 결과와 역시 일치함을 쉽게 확인할 수 있다.

Ⅶ. 최대원리와의 관계

동적계획법의 가치함수와 최대원리와의 사이에는 이른바 감응도관계 (sensitivity relationship)인 $p(t) = V_x(t, x(t))$ 가 성립한다는 것을 V 가 충분히 미분가능한 경우에는 앞의 특성방법등의 방법을 사용하여 보일 수 있다 또한 HJB 방정식에서 $H(t, x(t), u(t)) = V_t(t, x(t))$ 가 되므로 이 관계는

$(H(t,x(t),u(t)),-p(t))=\nabla V(t,x(t))$ 로 나타낼 수 있다. 그러나 동적계획법의 이론적 근거를 확보하기 위해서 HJB 방정식의 해의 개념이 일반화되어 가치함수의 범위가 불연속함수를 부분적으로 포함하는 LSC 함수로 확장됨에 따라 최대원리와의 관계는 그만큼 불분명해진다. 가치함수가 충분히 미분가능할 경우에 성립하였던 대등한 두 이론 체계간의 감응도관계는 좀더 일반적인 상황에서는 어떠한 형태가 되는가에 대한 최근의 결과가 다음에 소개하는 Vinter의 감응도관계에 관한 정리이다.8) Vinter 는 다음의 Mayer 형의 미분포함식 극소문제

$$\min g(x(T)) \quad s.t. \quad x(t)\in F(t,x(t)) \quad a.e. \quad x(0)=x_0$$

를 분석하며 여기서 미분포함식에 대해서는 다음을 가정한다

(HF1) F 는 연속다가함수 비공, 폐볼록치이다

(HF2) $\exists c>0 \ s.t.$ $F(t,x)\subset c(1+|x|)B \quad \forall(t,x)\in[S,T]\times R^n$

(HF3) $\exists k>0 \ s.t. \ F(t,x)\subset F(t,x')+k|x-x'|B$ $\quad \forall(t,x),(t,x')\in[S,T]\times R^n$

(HF1)에서 특기할 점은 F 의 연속성만을 요구하고 미분가능성을 요구하지 않는다는 점이다

정리(Vinter[21, p 478]) 윗 문제에서 F 가 (HF1), (HF2) 및 (HF3)을 만족시키고 $g\in LLC$ 이고 V 가 가치함수이고 x 가 최적경로이면 $\exists\ p\in W^{1,1}$ s t. $a.e.$ $t\in[0,T]$ $\forall v\in F(t,x(t))$

① $p(t)\in co\{q \mid (q,p(t))\in N_{GrF(t,\ )}(x(t),\dot{x}(t))\}$

② $p(t)\cdot x(t)\geq p(t)\cdot v$

8) R. B Vinter [2000], *Optimal Control*, Birkhauser,pp 478-483

③ $\left(\max_{v\in F(t,x(t))} p(t)\cdot v,\ -p(t)\right)\in co\partial V(t,x(t))$

④ $(p(0), -p(T))\in\partial_x(-V)(0,x(0))\times\partial g(x(T))$.

이 정리는 F, g 의 연속성만을 요구하므로 가치함수 V 의 연속미분가능성은 보장되지 않으며 이러한 경우 수반변수가 따르는 미분포함식 ①은 등식이 아닌 포함식형태의 수반관계로 성립하고 해밀토니언과 수반함수의 값과 가치함수의 편도함수와의 관계인 ③도 등식이 아닌 관계식인 감응도관계로 성립한다. 이와 같이 가치함수의 연속미분가능성이 보장되지 않는 비미분상황에서는 최대원리의 수반함수가 다수 존재하며 동적계획법의 가치함수의 하부미분도 다수 존재한다

이러한 상황에서 이 정리는 수반관계를 만족시키는 수반아크 중의 적어도 하나는 가치함수의 하부미분의 볼록포에 속한다는 것을 밝히고 있다 이와 같이 일반적인 비미분상황에서는 동적계획법의 결과와 최대원리의 결과간에는 가치함수가 연속미분가능한 경우처럼 일의적인 대응관계가 성립하는 것을 기대할 수 없다. Vinter[21, p475]의 예를 통해서 수반관계에서 얻은 절대연속함수중 수반아크가 될 수 없는 것이 있으며 수반아크중에도 감응도관계를 지키지 않는 것이 있다는 것을 보임으로써 이 점을 분명히 하고 있다

Ⅷ. HJB 방정식의 일반화된 해

주어진 최적제어문제의 가치함수와 HJB방정식을 놓고 고전적인 견해는 가치함수가 HJB방정식의 일의적인 해가 되는 것이었다 이는 가치함수가 연속적으로 미분가능한 경우 성립한다 그러나 제어문제의 주어진 함수들이 이상적으로 미분가능하더라도 가치함수는 그렇지 않은 일이 종종 발생하기 때문에 가치함수를 HJB방정식의 일의적 해로 만드는 데에는 추가적인 노력이 필요하게 되었다 미분가능성이 보장되지 않는 연속함수를 PDE의 해로 허용하기 위해 해의 개념을 확장시킨 것이 Crandall 등[8]의 비스코시티 해이다 이들은 상부미분과 하부미분의 개념을 사용하여 PDE가 하부미분에 대해 상부해이면 비스코시티 상부해, 상부미분에 대해 하부해이면 비스코시티 하부해로 정의하고 두

조건을 모두 만족시키면 비스코시티 해로 정의하여 이 분야에 새로운 분석의 길을 열었다 9)

그러나 제어문제에 상태제약이 포함되면 이러한 문제의 가치함수는 연속성 마저 보장되지 않는다 또한 최종상태에 대한 다양한 제약을 목적함수가 제약집합 밖에서는 무한대의 값을 갖고 제약집합 내에서만 유한한 값을 갖는 형태로 분석하는 것이 횡단조건 등의 도출에 편리하기 때문에 이러한 분석방법을 사용하려면 불연속인 가치함수를 포함할 필요가 생긴다 불연속인 가치함수가 HJB 방정식의 일의적 해가 되는 해의 개념과 조건을 제시한 것이 Frankowska-Plaskacz[12]와 Frankowska-Vinter[13] 등의 연구이다 이들은 외향조건이 지켜질 경우 가치함수는 HJB 방정식의 일반화된 해라는 것을 증명하였다 여기서 이들이 사용하는 해들 중에서 epi-미분해를 소개하기로 한다

다음은 Frankowska-Vinter[13]의 정의이다 먼저 이들이 분석하는 문제는

$$(P^1_{t,x})\begin{cases} \min g(x(T)) \\ s.t.\ \dot{x}(s)\in F(s,x(s)), a.e. s\in[t,T] \\ \quad x(s)\in A \quad \forall s\in[t,T] \\ \quad x(t)=x \end{cases}$$

이며 $\inf(P^1_{t,x})=V(t,x)$ 라 하면 이 문제의 HJB 방정식은

$$\phi_t(t,x)+\min_{v\in F(t,x)}\phi_x(t,x)\cdot v=0 \qquad (t,x)\in[0,T]\times \mathrm{int}A$$
$$\phi(T,x)=g(x) \qquad x\in R^n$$

이다 여기서 A 는 r 개의 LLC 구배를 갖는 C^1 함수인 $h_j\ R^n\to R$ 에 의해 $A=\bigcap_{j=1}^{r}\{x \mid h_j(x)\le 0\}$ 과 같이 정의되며 A 의 경계점을 $bdyA$ 라 하면 $x\in bdyA$ 일 때 $I(x)$ 는 $h_j(x)=0$ 이 되는 j 의 집합을 나타낸다

정의 (epi-미분 또는 Dini 하부미분) $\phi\ R^k\to R\bigcup\{\infty\}$ 가 LSC이고 $\phi(x)<\infty$

9) 비스코시티해에 관한 설명은 [3,4,9]를 참조할 것

일 때

$$D\uparrow \phi(x;u)=\lim_{h\downarrow 0}\inf_{w_i\to u}\frac{\phi(x+hw_i)-\phi(x)}{h}$$

를 x에서 u방향으로의 부수 epi-미분(contingent epiderivative) 또는 디니 하부미분(lower Dini deriavative) 이라 한다 또한 ϕ 가 부분집합에만 정의되었을 경우에는 나머지 집합에서는 ∞의 값을 갖는 것으로 정한다 다음은 Vinter-Frankowska[13]의 정의 및 관련된 정리이다

정의· (epi-미분해 또는 하부 Dini해) HJB 방정식

$$\phi_t-H(t,x,-\phi_x)=0,\quad \phi(T,x)=g(x)$$

이 주어졌을 때 LSC함수 $\phi\ [0,T]\times R^n\to R\cup\{\infty\}$ 가 다음을 만족시키면 이 방정식의 epi-해라 한다

① $\phi(t,x)<\infty$ 이고 $t\in[0,T)$ 인 모든 (t,x) 에서
$$\inf_{v\in F(t,x)} D\uparrow \phi((t,x)\cdot(1,v))\le 0$$

② $\phi(t,x)<\infty$ 이고 $t\in(0,T]$ 인 모든 (t,x) 에서
$$\sup_{v\in F(t,x)} D\uparrow \phi((t,x)\cdot(-1,-v))\le 0$$

③ $\phi(T,x)=g(x)\quad \forall x\in R^n$

이 정의를 살펴보기 위해서 먼저 $\phi\in C^1$ 인 경우 임의의 $h=(1,v)$ 방향에 대한 기울기는

$$D\phi((t,x),(1,v)) = \phi_t(t,x) + \phi_x(t,x) \cdot v$$

이므로 HJB 방정식은 정의역의 임의의 (t,x) 에서 선택 가능한 방향기울기의 최소값이 0 이라는 것을 요구하고 있음을 알 수 있다. 이러한 조건을 만족시키는 함수는 각 점에서 등고선의 방향이 선택가능한 가장 작은 기울기의 방향이며 다른 방향을 선택하면 함수가 증가하여야 한다 이로써 위의 정의의 조건 (1)은 등고선 이하의 영역으로의 선택이 가능하여야 한다는 것이고 조건 (2)는 모든 가능한 선택의 반대방향은 등고선 이하의 영역이어야 한다는 것임을 알 수 있다

정리. F 가 앞의 (HF1)(HF2)(HF3) 를 만족시키고 다음 제약자격(CQ)를 만족시키면 가치함수는 HJB 방정식의 epi-해이고 역도 사실이다

(CQ) $\forall x \in A, \forall t \in [0, T], \exists v \in F(t,x)$ $s.t.$
$\nabla h_j(x) \cdot v > 0, \ \forall j \in I(x)$

이 외향조건은 유효상태제약의 경계에서 밖으로 움직일 수 있어야 한다는 것을 요구하고 있다

예 (예의 계속) 이제는 앞의 예에서 $T \in (0,1)$ 로 고정시키고 제어변수를 $U=[0,2]$ 의 범위로 제한하자 그러면 최적제어는 종전의 닫힌 루프 전략이 2 를 초과하는 영역에서는 2 가 되고 그 이하의 영역에서는 그대로가 됨은 자명하다 즉

$$u(t,x) = \min\left[2, \frac{1+x}{1-T+t}\right]$$

이고 이 경우 HJB 방정식은

$$\max\left(V_x u + ux - \frac{u^2}{2}\right) \ s.t. \ u \in [0,2]$$

의 해를 $\hat{u}$ 라 하면 $\hat{u} = \min[\max[V_x + x, 0], 2]$ 이므로 이 기호를 사용하여 HJB 방정식을 나타내면

$$V_t + \left(V_x \hat{u} + \hat{u}x - \frac{\widehat{u^2}}{2} \right) = 0, \quad V(T, x) = x$$

이 된다 최적제어의 형태가 알려졌으므로 이에 따라 가치함수를 직접 계산하면

$$V(t, x) = x + (T - t) \min\left[2(x + T - t), \frac{(1+x)^2}{2(1 - T + t)} \right]$$

이다 이 함수는 $\frac{x+1}{1-T+t} = 2$ 가 되는 (t, x) 를 따라 제어제약하의 가치함수와 제어제약이 없을 때의 가치함수는 같은 값을 가지며 또한 같은 미분을 가지며 편도함수들은 연속이므로 C^1 함수로서 HJB방정식의 고전해이다 따라서 비스코시티해와 epi-해이다 그러나 $x \geq 0$ 이라는 상태제약이 주어질 때 현재의 예는 $U = [0, 2]$ 로서 외향성분이 없으므로 현재의 해는 제약하의 epi-해는 해당되지 않는다 그러나 $U = [-2, 2]$ 로 주어질 경우 외향성분이 있으므로 위의 해는 epi-해이다

〈참고문헌〉

[1] V I Arnold [1989], *Mathematical Methods of Classical Mechanics second edition*, Springer

[2] R Bellman [1967], *Introduction to the Mathematical Theory of Control Processes, vol 1* Academic Press

[3] M Bardi and I Capuzzo-Dolcetta [1997], *Optimal Control and Viscosity solutions of Hamilton-Jabcobi-Bellman Equations*, Birkhauser

[4] M Bardi, M G Crandall, L C Evans, H M Soner, and P E Souganidis [1997], *Viscosity Solutions and Applications,* Lecture Notes in Mathematics 1660, Springer-Verlag

[5] F H Clarke [1983], *Optimization and Nonsmooth Analysis*, Wiley

[6] ________, Y Ledayaev, R Stern, P Wolenski [1998], *Nonsmooth Analysis and Control Theory*, Springer-Verlag

[7] R Courant and D Hilbert [1962], *Methods of Mathematical Physics, vol II Partial Differential Equations* Wiley-Interscience

[8] M G. Crandall, H Ishii, and P-L. Lions [1992], User's Guide to Viscosity Solutions of Second Order Partial Differential Equations, *Bull Amer Math Soc* 27, pp 1-67

[9] L C Evans [1998], *Partial Differential Equations*, AMS

[10] W H Fleming and R W Rishel [1975], *Deterministic and Stochastic Optimal Control*, Springer

[11] _______ and H M Soner [1991], *Controlled Markov Processes and Viscosity Solutions*, Springer

[12] H Frankowska and S Plaskacz [2000], Semicontinuous Solutions of Hamilton-Jacobi-Bellman Equations with Degenerate State Constraints, *J Math Anal Appl*,251,pp 818-838

[13] _______ and R B Vinter[2000], Existence of Neighboring Feasible Trajectories Applications to Dynamic Programming for State-Constrained Optimal Control Problems, *J Opt Appl*, *104*, 1, pp 21-40

[14] E B. Lee and L Markus [1967], *Foundatıons of Optımal Control Theory, reprınt edıtıon*[1986], Krıeger.

[15] A A. Melikyan [1998], *Generalızed Characterıstıcs of Fırst Order PDEs Applıcatıons ın Optımal Control & Differential Games*, Birkhauser.

[16] L S. Pontryagın, V C. Boltyanskı III, R V Gamkrelidge, E.F. Mishchenko [1962], *The Mathematical Theory of Optımal Process*, Wıley & Sons.

[17] H. Sagan, [1969], *Introductıon to the Calculus of Varıatıons*, Dover

[18] H J. Sussmann [1999], Geometry and Optımal Control ın *Mathematical Control Theory* edited by J. Baıllieul and J. C Willems, Sprınger

[19] _______ and J C Willems [1997], 300 Years of Optımal Control. From the Brachystochrone to the Maximum Prıncıple, *IEEA Control Systems Magazıne, v 17 no 3*, pp.32-44.

[20] R. T Rockafellar [1970], *Convex Analysıs*, Prınceton Unıv Press

[21] R. B Vınter [2000], *Optımal Control*, Birkhauser

[22J J Yong and X Y. Zhou [1999], *Stochastıc Controls Systems and HJB Equatıons*, Springer

Abstract

HJB equation and the Maximum Principle

Dosung Lee

Evolving relationships between the main results of dynamic programming and optimal control are surveyed Their common predecessor is the classical calculus of variations and their major results, i e , HJB equation and the maximum principle are essentially two different ways of expressing the Euler-Lagrange condition of the calculus of variations This fact is accounted by the method of charateristics that converts the PDE of HJB equation into the system of ODEs of the maximum principe. The relationship between these two methods of dynamic optimization weakens as nonsmoothness in value function is permitted and the equalities that prevailed with smoothness becomes those of inclusions

JEL Classification: C6, C7