독립심화학습 7주차

2017103580 김정운

이전의 헤밀턴-야코비-벨만 방정식(HJB)은 optimal control에 대한 필요충분조건이지만, value function이 infimum으로 정의되어 있기 때문에 명확한 형태의 control을 구하기 어렵다고 했다. 또한 HJB를 근사시키는 방법으로 강화학습과 policy iteration이 있는데, 이들이 같은 해로 수렴한다는 것을 증명하기 쉽지 않으며 연산량이 매우 많다. 이러한 한계를 극복할 수 있는 한 가지 방법은 optimal control에 대한 필요조건을 제시한 다음에, 그것이 유일하다는 것을 증명하여 최적제어항을 얻는 것이다. 이러한 방법을 pontraygain’s minimum principle이라고 한다.

Min subject to 라는 최적제어문제가 있다고 하자(U는 bounded set). 이때 함수 와 는 연속함수이고, x에 대한 편미분 , 는 모든 변수 (t,x,u)에 대해 연속이어야 한다. 이러한 조건을 classical regularity hypotheses라고 하며, 이는 optimal control의 존재성 등을 증명하기 위해 필요하다. 헤밀토니안 을 정의하는데, 이때 은 0 또는 1의 값을 취한다. 0을 취한다는 것은 제약조건이 매우 엄격하여 오직 하나의 control 만을 구할 수 있음을 의미한다. 이때 cost function에 대한 정보를 사용하지 않기 때문에 이 된다. 이러한 현상이 발생한 원인 중 하나는 제약조건이 매우 많은 상황이기 때문에, 이를 줄이거나 다른 방식으로 문제를 설정하기도 한다. 이는 cost function을 활용하는 것을 목적으로 둔다. 이라는 것은 사용할 수 있는 control 중에서 cost function을 최소화하는 control을 찾고자 하는 것이다.

만약 u\*가 위의 문제에 대한 optimal control이며 는 이에 대응되는 optimal state trajectory라고 한다면, 와 가 성립하고 는 , u\*서 maximum을 취한다. 첫 번째 방정식을 adjoint equation이라고 하며, 두 번째 조건과 세 번째 조건을 각각 transversality condition과 maximum condition이라고 한다. 이러한 정리를 pontraygain’s maximum(or minimum) principle(이하 PMP)고 한다.

이때 adjoint equation의 p가 value function이라고 한다면, 해당 방정식은 HJB 방정식이 된다. 이는 PMP의 특별한 경우가 HJB 방정식이라는 것을 의미한다. p가 취할 수 있는 후보와 이에 대응되는 control 및 state들을 모두 찾은 다음에, 그 중에서 cost function을 최소화할 수 있는 control을 구할 수 있다. 만약 cost function이 u에 대해 convex한다면 cost function을 최소화할 수 있는 control은 유일하기 때문에, PMP로 구한 해가 최적해가 될 수 있다. 이는 HJB방정식과 달리 value function을 사용하지 않기 때문에 미분가능성을 보장할 수 있으며, infimum을 사용하지 않아 보다 명확하게 방정식을 기술할 수 있다. 이는 policy iteration이나 강화학습을 사용하지 않음으로써 연산량을 줄일 수 있다.

물리학이나 기계공학 등에서 cost function이 control u에 대한 이차함수라고 함으로써 u에 대한 convex를 보장한다. 반면 경제학이나 사회학에서는 이러한 이차함수나 convex 함수를 사용하기 어려울 수 있는데, 이들을 최소화하는 control은 평균과 관련될 수 있기 때문이다. 분산이 큰 상황에서의 평균은 robust가 떨어지는데, 이를 방지하는 한 가지 방법은 절대오차를 cost function으로 사용하는 것이다. 그러면 중앙값과 관련된 control이 산출되며, 이는 평균보다 상대적으로 분산에 덜 민감하다. 하지만 cost function의 convex는 control의 미분가능성을 보장하기 때문에, smooth control이 필요한 분야에서는 convex function으로 cost function을 설정하는 것이 좋다.