

Title	非線形最適レギュレータ問題への強化学習の適用
Author(s)	内藤, 浩行
Citation	
Issue Date	2000-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/670
Rights	
Description	Supervisor: 吉田 武稔, 知識科学研究科, 修士

An Application of Reinforcement Learning to Non-linear Optimal Regulatory Problems

Hiroyuki Naitou

School of Knowledge Science,
Japan Advanced Institute of Science and Technology

March 2000

Keywords: Reinforcement Learning , Optimal Control , Non-linear System , Order of Truncation.

In this research , we propose to derive by using the reinforcement learning the order of truncation concerning the control law of the non-linear system which the control law is given the formal power series . Therefore , we assume the system which the control law is designed the formal power series and attempt to run by the numerical value simulation . In that case , we should truncate the order to assume the formal power series to be the limited series to mount the control law . Moreover , the degree of the approximation concerning the order differ from the non-linear system which is the object . So , we design as the reinforcement learning problem which derives the order of truncation for the control law relies on a experience and a intuition and plan the design of regulator . At the same time , the examination for the order of truncation by the numerical value simulation is presented .

Next , For the purpose of the above , we shall introduce the reinforcement learning which is technique used by this research . The reinforcement learning , from a point of view technological , is the system which imitated the adjustment phenomenon which is that the learner study a suitable action which escapes from the punishment by obtaining the reward from the unknown environment . When we catch this system from the side of machine learning , it is the class of learning problem which presened by R.S.Sutton and C.J.C.H.Watkins . This learning method become a center of attraction as the method which an adaptive solution is obtained for the system which exists in the evaluation of the state the delay , and includes the discrete discontinuity and the uncertainty in the state transition . That is , The reinforcement learning is considered the solution method which applied to the problem to solve the problem . We shall describe the component of such reinforcement learning , the environment , the agent , the policy , the reward and the value function , in addition the connection of the component is explained . Moreover we will discuss

about the mounting method , that is Q -learning proposed by C.J.C.H.Watkin and used by this article .

Next , we shall report about the optimal regulator problem , which is the problem is formulated by reinforcement learning in this research and run . First of all , so far a lot of researches are done , we shall simply describe about the linear optimal regulator problem which is the optimal control problem that the evaluation function consists of the quadratic form evaluation function and the state equation is linear ordinary differential equations . So , we shall explain to become clear about the optimal regulator problem of the non-linear system that the plant , which becomes the object of the research by this article , is expressed in the formal power series . Moreover , the non-linear optimal regulator problem is defined as the suboptimal regulator problem . The numerical value simulation for this problem is presented . The property of this system is reported . The main results of the non-linear optimal regulator problem which is the controled plant of this research is presented by Yoshida . So , the presented figure and numerical value simulation is reported by Dr . tanaka . As a result , when is given a big initial value , even if the order of truncation is mode higher-order , it is clear that the response of the system causes the overshoot from the figure .

From a point of view described above , the formulation to solve the order of truncation concerning the control law of the non-linear system which is object of reinforcement learning is run . Defines for the component of reinforcement learning the following : the enviroment is the non-linear system , the state is the state space , the action is the state feedback rule that the order of truncation differ , the reward is the quadratic form for the control input and the state variable , the agent is the controller , the value function is expressed in the degree of the order of truncation and the state variable . Using these definition , we shall define the reinforcement learning problem as the deriving problem of the order of truncation of non-linear optimal regulator problem . As a result , the order of truncation using the solution of the better policy is presented . Moreover , the numerical value simulation about the defined problem is presented In that case , the coefficient matrix of the mounted system uses the numerical value introduced in the preceding chapter . And , the time response of the system when learning is done and the trajectory are shown in figure , the evaluation value is presented , compare with the result when the order of truncation is fixed , the effectiveness of this technique is presented .

Therefore , with the better order of truncation , simply , unless it only has to be a higher-order number , it is clarified that a necessity for switching the order of truncation according to the situation is need . Moreover , the situation for switching the order of truncation is considered . Finally , the relation between the over shot and the discontinuance degree is considered by simulating for using a coefficient different from coefficient matrix used when the above-mentioned is simulated .