

Support Vectors for Reinforcement Learning

Thomas G. Dietterich and Xin Wang

Oregon State University, Corvallis, Oregon, USA

tgd@cs.orst.edu

<http://www.cs.orst.edu/~tgd>

Abstract. Support vector machines introduced three important innovations to machine learning research: (a) the application of mathematical programming algorithms to solve optimization problems in machine learning, (b) the control of overfitting by maximizing the margin, and (c) the use of Mercer kernels to convert linear separators into non-linear decision boundaries in implicit spaces. Despite their attractiveness in classification and regression, support vector methods have not been applied to the problem of value function approximation in reinforcement learning. This paper presents three ways of combining linear programming with kernel methods to find value function approximations for reinforcement learning. One formulation is based on the standard approach to SVM regression; the second is based on the Bellman equation; and the third seeks only to ensure that good actions have an advantage over bad actions. All formulations attempt to minimize the norm of the weight vector while fitting the data, which corresponds to maximizing the margin in standard SVM classification. Experiments in a difficult, synthetic maze problem show that all three formulations give excellent performance. However, the third formulation is much more efficient to train and also converges more reliably. Unlike policy gradient and temporal difference methods, the kernel methods described here can easily adjust the complexity of the function approximator to fit the complexity of the value function.