6.231 Dynamic Programming and Stochastic Control
Fall 2008

6.231 Dynamic Programming

Approximate Dynamic Programming Lectures Outline

▼ Lecture 1: Overview of main approaches
- ▼ Problem approximation
- Simplification - Aggregation
- Use in limited lookahead
- ▼ Approximation in value space
- Rollout
- Approximate policy iteration
- Q-Learning
- Bellman error approach
- Approximate linear programming
- ▼ Approximation in policy space
- Direct parametrization
- Indirect parametrization through the cost

▼ Lecture 2: Cost approximation - Discounted cost (Section 6.2, Vol. II)
- Policy evaluation - Policy improvement
- Actor-critic paradigm
- Direct policy evaluation using cost samples
- Batch and incremental methods
- TD(1) and TD(lambda)
- Regular and optimistic versions
- Q-Factor evaluation
- The problem of exploration

▼ Lecture 3: Projected equation methods (Section 6.3, Vol. II + Refs)
- The projected Bellman equation
- Contraction properties and error bounds
- Direct solution and projected value iteration
- Simulation-based versions - LSTD(0) and LSPE(0)
- LSTD(lambda), LSPE(lambda), and TD(lambda)

▼ Lecture 4: More on projected equations - Q-learning (Sections 6.3, 6.4, Vol. II + Refs)
- Basis function selection and tuning
- Practical aspects
- Convergence analysis
- Bellman error alternatives
- Q-Learning

- Optimal stopping

▼ Lecture 5: Extensions to SSP and average cost (Section 6.5, 6.6, Vol. II + Refs)
- Contraction properties for SSP
- Approximation policy iteration for SSP
- Q-Learning for SSP
- Average cost approximations
- Approximation policy iteration for average cost
- Q-Learning for average cost

▼ Lecture 6: Gradient methods for approximation in policy space (Section 6.7, Vol. II + Refs)
- Problem formulations
- Gradient formulas
- Gradient evaluation using simulation
- Feature selection
- Comparison of approximation in value and policy space