Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

On Approximate Dynamic Programming in Switching Systems

Anders Rantzer

***Abstract*— In order to simplify computational methods based on dynamic programming, an approximative procedure based on upper and lower bounds of the optimal cost was recently introduced. The convergence properties of this procedure are analyzed in this paper. In particular, it is shown that the computational effort in finding an approximately optimal control law by relaxed value iteration is related to the polynomial degree that is needed to approximate the optimal cost. This gives a rigorous foundation for the claim that the search for optimal control laws requires complex computations only if the optimal cost function is complex. A computational example is given for switching control on a graph with 60 nodes, 120 edges and 30 continuous states.**

I. Introduction

Optimal switching between linear systems is in many respects as challenging as optimal control of general nonlinear or hybrid systems. It is rarely possible to find exact expressions for optimal control laws or the optimal cost. Instead approximative solutions need to be sought. Already in Bellman's pioneering work on dynamic programming [3], the need for approximate solutions was recognized and discussed. Since then, a variety of methods have been developed, with application to discrete optimization as well as Markov processes, differential equations and hybrid systems. Of particular significance for this paper is the inequality version of the Hamilton-Jacobi-Bellman equation, used by Leake and Liu [12] to derive bounds on the optimal cost function. It turns out that the inequality for lower bounds on the optimal (minimal) cost is convex. This gives a natural connection to convex duality theory in optimal control, an idea introduced by Kantorovich [11] for mass transportation problems, which has been recently been further explored [26], [20], [21], [22]. An application to image databases is described in [24]. Computational methods based on convex optimization were pursued in [23], [10] and the idea of relaxed dynamic programming was introduced in [14], [13].

Numerical solutions to the Hamilton-Jacobi-Bellman equation in a continuous state space are often based on discretization [8], [9]. This gives a connection to the rich literature on optimal control in discrete state spaces [4]. In particular, error bounds for approximate dynamic programming were given in [25], [7]. An alternative method which avoids discretization is to use Galerkin's spectral method to approximate the optimal cost function without prior discretization [2]. Altogether, existing methods have proved effective for many small scale problems, but the complexity grows exponentially with increasing state dimension.

In contrast to general nonlinear methods with exponential growth, it is well known that linear-quadratic optimal control problems grow only polynomially with state dimension and can be solved with hundreds of state variables. It is therefore challenging to search for general nonlinear synthesis procedures that reduce to Riccati equations in the special case of linear-quadratic control and to linear programming in the case of network optimization on a finite graph. One step in this this direction was taken in [15]. This paper proceeds towards the goal in a more general setting.

Recent research on model predictive control and optimal control of hybrid systems is also connected to this work [16], [17], [6], [5]. In fact, our approach resulted from an effort to treat hybrid systems by merging methods and experiences from the two fields of network optimization and control theory. In particular, convex inequality relaxations commonly used in network optimization are combined with computational tools from the control field, such as linear matrix inequalities and sum-of-squares optimization.

The next section of the paper reviews some of the basic results on dynamic programming before stating the main results on global convergence in approximate value iteration. Then the focus is moved to the special case of switching systems in section IV. For such systems, the general results are concretized and a computational example is completed in section V.

II. Approximate value iteration

Let X, the set of states, and U, the set of inputs, be arbitrary. Given $f : X \times U \to X$ consider the dynamical system

$$x(k+1) = f(x(k), u(k)) \qquad x(0) = x_0 \qquad (1)$$

with $k = 0, 1, 2, \ldots$ Combining this with the control law $\mu : X \to U$ gives the closed loop dynamics

$$x(k+1) = f(x(k), \mu(x(k))) \qquad (2)$$

To measure the performance of the system, we introduce a non-negative step cost $l : X \times U \to \mathbf{R}$ and define

A. Rantzer is with Department of Automatic Control, LTH, Lund University, Box 118, SE-221 00 Lund, Sweden, `rantzer at control.lth.se`.

The paper was written during a sabbatical at California Institute of Technology.

0-7803-9568-9/05/$20.00 ©2005 IEEE

the value function

$$V_\mu(x_0) = \sum_{k=0}^{\infty} l(x(k), \mu(x(k)))$$

where x is given by (2). The optimal cost function V^* is defined as

$$V^*(x_0) = \inf_\mu V_\mu(x_0)$$

and can be characterized as follows:

Proposition 1 (Dynamic programming [3]): Suppose that $V : X \to \mathbf{R}$ satisfies

$$0 \le V(x) = \min_u [V(f(x,u)) + l(x,u)] \qquad \forall x \tag{3}$$

and $\lim_{j\to\infty} V(x_j) = 0$ for every $\{(x_j,u_j)\}_{j=1}^\infty$ with $\sum_{j=1}^\infty l(x_j,u_j) < \infty$. Then $V = V^*$ and the formula

$$\mu^*(x) = \arg\min_u [V^*(f(x,u)) + l(x,u)] \tag{4}$$

defines an optimal control law.

An iterative approach to solution of the Hamilton-Jacobi-Bellman equation (3) is known as value iteration. Next, we give a bound on the convergence rate of this scheme.

Proposition 2 (Value iteration convergence): Suppose the condition $0 \le V^*(f(x,u)) \le \gamma l(x,u)$ holds uniformly for some $\gamma < \infty$ and that $0 \le \eta V^* \le V_0^* \le \delta V^*$. Then the sequence defined iteratively by

$$V_{j+1}^* = \min_u \left[V_j^*(f(x,u)) + l(x,u)\right] \qquad j \ge 0 \tag{5}$$

approaches V^* according to the inequalities

$$\left[1 + \frac{\eta - 1}{(1+\gamma^{-1})^j}\right] V^* \le V_j^* \le \left[1 + \frac{\delta - 1}{(1+\gamma^{-1})^j}\right] V^* \tag{6}$$

In particular, if $0 \le V_0^* \le V^*$, then

$$\left[1 - \frac{1}{(1+\gamma^{-1})^j}\right] V^* \le V_j^* \le V^*$$

The proof is given in Section VI.

The main limiting factor in applications of value iteration is the complexity in computation and representation of the functions $V_j^*(x)$. Many schemes for approximation have therefore been developed. In this paper, we will use the following statement to quantify the effects of approximation errors in the Hamilton-Jacobi-Bellman equation.

Proposition 3 (Approximate dynamic prog. [13]): Suppose $0 \le \alpha \le 1 \le \beta$. Let $V : X \to \mathbf{R}$ satisfy

$$\begin{aligned}&\min_u \Big\{V(f(x,u)) + \alpha l(x,u)\Big\}\\ &\le V(x) \le \min_u \Big\{V(f(x,u)) + \beta l(x,u)\Big\}\end{aligned} \tag{7}$$

and $\lim_{j\to\infty} V(x_j) = 0$ for every $\{(x_j,u_j)\}_{j=1}^\infty$ with $\sum_{j=1}^\infty l(x_j,u_j) < \infty$. Then

$$\alpha V^*(x) \le V(x) \le \beta V^*(x) \qquad \forall x$$

Moreover, $\mu(x) = \arg\min_u [V(f(x,u)) + \alpha l(x,u)]$ has a value function V_μ satisfying $\alpha V_\mu \le V$.

Solutions to the inequalities (7) can be found by approximate value iteration:

Proposition 4 (Approximate value iteration): Suppose $\{V_j\}_{j=0}^\infty$ and $\{V_j^*\}_{j=0}^\infty$ start from $V_0 \equiv V_0^*$ and

$$\begin{aligned}&\min_u \Big\{V_j(f(x,u)) + \alpha l(x,u)\Big\}\\ &\le V_{j+1}(x) \le \min_u \Big\{V_j(f(x,u)) + l(x,u)\Big\}\end{aligned} \tag{8}$$

while V_j^* satisfies (5). Then $\alpha V_j^* \le V_j \le V_j^*$ for all j.

Proof. The statement follows by induction over j. □

Combining Proposition 4 with the convergence bound of Proposition 2, we get that the following bound on the distance from optimality.

Theorem 1: Given $0 \le \alpha \le 1$, assume that $0 \le V^*(f(x,u)) \le \gamma l(x,u)$ uniformly, $\gamma < \infty$ and that the sequence $V_0, V_1, V_2, \ldots$ starting with $0 \le V_0 \le V^*$ satisfies (8). Then

$$\alpha_j V^* \le V_j \le V^* \qquad \alpha_j = \left[1 - (1+\gamma^{-1})^{-j}\right]\alpha \tag{9}$$

Moreover, $\mu_j(x) = \arg\min_u \Big\{V_j(f(x,u)) + \alpha l(x,u)\Big\}$ gives a value function $V_{\mu_j}(x)$ satisfying

$$[\alpha + \gamma(\alpha_j - 1)]\, V_{\mu_j}(x) \le V^*(x) \tag{10}$$

Remark 1. The inequality (10) gives an upper bound on the cost function for the policy μ_j provided that the bracket in front of V_{μ_j} is positive. This will happen for large values of j whenever $\alpha > \gamma/(1+\gamma)$.

Proof. The inequalities (9) follows directly from Proposition 4 and Proposition 2. Hence

$$V_j(f(x,\mu_j(x))) + \alpha l(x,\mu_j(x)) \le V_{j+1}(x) \le V^*(x)$$

Using $\alpha_j V^* \le V_j$ and $V^*(f(x,u)) \le \gamma l(x,u)$, we get

$$\begin{aligned}\alpha_j V^*(f(x,\mu_j(x))) + \alpha l(x,\mu_j(x)) &\le V^*(x)\\ V^*(f(x,\mu_j(x))) + [\alpha + \gamma(\alpha_j - 1)] l(x,\mu_j(x)) &\le V^*(x)\end{aligned}$$

For trajectories of (1) with $u(k) = \mu_j(x(k))$, we get

$$[\alpha + \gamma(\alpha_j - 1)]\, l(x,\mu_j(x)) \le [V^*(x(k)) - V^*(x(k+1))]$$

Summing over k gives (10). □

III. Iterations in a finite-dimensional space

When X has an infinite number of elements, the search for the optimal cost V^* is a search in an infinite-dimensional space. It is often natural to limit this search to a finite-dimensional subspace $\mathcal{L}$, for example polynomials of a fixed degree. A natural question to ask is whether existence of a solution to (7) in $\mathcal{L}$ has any implications on feasibility of the iterative inequalities (8). A striking result of this kind is given next, but for a slightly modified algorithm:

Theorem 2: The conclusions of Theorem 1 remain valid if the conditions (8) are replaced by

$$\min_u \Big\{ V_j(f(x,u)) + \alpha l(x,u) \Big\}$$
$$\leq V_{j+1}(x) \leq \min_u \Big\{ V_{j+1}(f(x,u)) + l(x,u) \Big\} \qquad (11)$$

Proof. Every solution V_{j+1} to the right inequality in (11) must be bounded from above by V^* as shown in Proposition 3. Moreover, the lower bound from Proposition 4 remains valid with the same proof. The rest of of the proof is identical to the proof of Theorem 1. □

Remark 2. Suppose that V^* has a simple approximation in the sense that $V^s \in \mathcal{L}$ satisfies

$$\min_u \Big\{ V^*(f(x,u)) + \alpha l(x,u) \Big\}$$
$$\leq V^s(x) \leq \min_u \Big\{ V^s(f(x,u)) + l(x,u) \Big\} \qquad (12)$$

Then, with $V_0 \equiv 0$, the iterative inequalities (11) define feasible convex conditions on $V_{j+1} \in \mathcal{L}$ at every step.

Remark 3. Time-varying linear quadratic optimal control problems, usually solved by Riccati equations, and shortest-path network problems solved by linear programming are two well-known special cases of our framework. One consequence of Theorem 2 is that also other problems with an optimal cost function close to one of these special cases will be solvable with small computational effort.

Remark 4. Notice that the right hand side of (12) is bounded from above by $\min_u \{V^*(f(x,u)) + l(x,u)\}$. Comparing this to the left hand side shows that the only difference is the coefficient in front of $l(x,u)$. Hence the assumption (12) implicitly puts a constraint on the relative sizes of the cost in the next step $l(x,u)$ and the remaining cost $V^*(f(x,u))$. For optimal control problems with slow decay rate of the terms in the sum $\sum_k l(x(k),u(k))$ at optimality, this means that V^s needs to approximate V^* very accurately in order for the theorem to apply.

This observation has a natural interpretation in economic language. Let $V^*(x)$ be the value of a product with quality and location specified by x. The changes due to the business transaction u are given by $f(x,u)$. The transaction generates profit quantified by $l(x,u)$. The problem to maximize $\sum_k l(x,u)$ is then aimed to find the most profitable sequence of business transactions. In this context, the comparison of $l(x,u)$ and $V^*(f(x,u))$ says that small profit margins in each transaction increases the need for exact representation of the cost function at each step.

Remark 5. The difference between (8) and (11) is that in the second case, V_{j+1} appears also in the right hand side, not just in the middle expression. This enables us to guarantee feasibility in every iteration. The condition (11) is slightly more complicated than (8) but is still a convex condition on V_{j+1}. A disadvantage in some applications is that the new condition leaves less room for distributed computations.

Combination of Theorem 2 with the previous bounds on value iteration convergence gives the following main result of the paper.

Theorem 3: Assume $0 \leq V^*(f(x,u)) \leq \gamma l(x,u)$ uniformly with $\gamma < \infty$. Let $\mathcal{L}$ be a linear space of functions $X \to \mathbf{R}$. Suppose that there exists a $U \in \mathcal{L}$ such that $(1-\epsilon)V^*(x) \leq U(x) \leq V^*(x)$ where $0 \leq \epsilon < (1+\gamma)^{-2}$. Then, with $V_0 \equiv 0$ and $\alpha = 1 - \epsilon(1+\gamma)^2$, the iterative convex inequalities (11) have a solution sequence $V_0, V_1, V_2 \ldots \in \mathcal{L}$ and the conclusions of Theorem 1 remain valid.
The proof is given in Section VI.

Remark 6. Combining this result with $\mathcal{L}$ as a set of polynomials and using the sum-of-squares technique [18], [19] for verification of the inequalities (11) gives a very general computational setting for optimal control. In this context, it is natural to apply the theorem with a modified interpretation of the inequalities, namely that the differences between left and right hand sides can be written as sums of squares.

In particular, the theorem proves an attractive feature of the algorithm defined by iteration of (11). This is that the computational effort in finding an approximately optimal control law (the polynomial degree needed in the approximate value iteration) is related to the polynomial degree that is required to approximate the optimal cost. It also quantifies the accuracy of the outcome in terms of two fundamental parameters related to the difficulty of the problem, γ and ϵ.

IV. A model of switched linear systems

To concretize the results for switched linear systems, consider a graph defined by a set of nodes $\mathcal{N}$ and a set of edges $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$. A matrix $A_{ij} \in \mathbf{R}^{n \times n}$ is assigned to each edge $(i,j) \in \mathcal{E}$. The state $x = (z,i)$ has two components, $z \in \mathbf{R}^n$ and $i \in \mathcal{N}$ and the system dynamics are

$$\begin{aligned} z(k+1) &= A_{i(k)u(k)} z(k) \qquad & z(0) &= z_0 \\ i(k+1) &= u(k) & i(0) &= i_0 \end{aligned} \qquad (13)$$

Note that z evolves according to a linear equation defined by A_{ii} as long as the discrete state i remains constant. The role of the input u is to induce changes in the discrete state.

The step cost is defined by a set of matrices $Q_{ij} \geq 0$ for $(i,j) \in \mathcal{E}$ such that

$$l((z,i),u) = z^T Q_{iu} z$$

Thus, the cost is given by Q_{ii} when the discrete state i remains unchanged and by Q_{iu} when the step switches to u.

Taken together, this gives the following problem statement for switched linear systems:

$$\text{Minimize} \sum_{k=0}^{\infty} z(k)^T Q_{i(k)u(k)} z(k) \text{ subject to (13).} \quad (14)$$

Example 1 (Linear time-varying systems with quadratic cost) In the special case of a graph with only one path, i.e. for every $i \in \mathcal{N}$ there is just one j with $(i,j) \in \mathcal{E}$, the cost function is a quadratic function $V^*(z,i) = z^T P^i z$ uniquely determined by the initial state. The Hamilton-Jacobi-Bellman equation then reduces to a time-varying Lyapunov equation

$$P^k = A_k^T P^{k+1} A_k + Q_k$$

with $P^k = P^{i(k)}$, $A_k = A_{i(k)i(k+1)}$ and $Q_k = Q_{i(k)i(k+1)}$. □

Computation of the optimal control law for (14) is generally NP-hard. In fact, the classical travelling salesman problem is a special case.

V. Computations for switched linear systems

Let us now specialize the results of section II to the case of switched linear systems. Define

$$V^*(z_0, i_0) = \min_{u(0),u(1),\ldots} \sum_{l=1}^{\infty} z(l)^T Q_{i(l)u(l)} z(l)$$

where the relationship between u, i and z is defined by the dynamics (13). Then the Hamilton-Jacobi-Bellman equation becomes

$$V^*(z,i) = \min_u \left\{ V^*(A_{iu} z, u) + z^T Q_{iu} z \right\} \quad (15)$$

For approximate solutions, a natural space $\mathcal{L}$ for a first approximation of the optimal cost is the space of quadratic forms $V(z,i) = z^T P^i z$. For example, if $P^1, \ldots, P^m$ are symmetric matrices satisfying the matrix inequalities

$$P^i \le A_{iu}^T P^u A_{iu} + Q_{iu} \qquad \forall (i,u) \in \mathcal{E}$$

then Proposition 3 shows that $z^T P^i z \le V^*(z,i)$ for every z, i.

With this parameterization, the inequalities (11) can equivalently be written

$$\begin{aligned} &\min_u \left\{ z^T A_{iu}^T P_j^u A_{iu} z + \alpha z^T Q_{iu} z \right\} \\ &\le z^T P_{j+1}^i z \le z^T A_{iu}^T P_{j+1}^v A_{iu} z + z^T Q_{iv} z \end{aligned} \quad (16)$$

for all $z \in \mathbf{R}^n$, $(i,v) \in \mathcal{E}$ and the minimization is over all u with $(i,u) \in \mathcal{E}$. At each step of the iteration, these inequalities should be solved for the matrices $P_{j+1}^1, \ldots, P_{j+1}^m$. The second inequality reduces to standard linear matrix inequalities on the independent variables. The first inequality is also a convex constraint on P_{j+1}^i, but more cumbersome, since the minimum expression on the left hand side does not have a simple representation.

A more conservative, but often useful, alternative to (16), is to instead require existence of scalar parameters $\theta_{j+1}^1, \ldots, \theta_{j+1}^m \ge 0$ with $\sum_{j=1}^m \theta_{j+1}^j = 1$ and such that

$$\sum_u \theta_{j+1}^u \left(A_{iu}^T P_j^u A_{iu} + \alpha Q_{iu} \right) \le P_{j+1}^i \le A_{iv}^T P_{j+1}^v A_{iv} + Q_{iv} \quad (17)$$

for all $(i,v) \in \mathcal{E}$. The parameters θ_{j+1}^i can be interpreted as the probabilities of a stochastic control law, which ignores the value of the continuous state z, hence the conservatism. The inequalities can be solved for θ_{j+1}^u and P_{j+1}^i by semi-definite programming in order to generate a sequence $P_0^i, P_1^i, P_2^i, \ldots$ that converges to a solution of the inequalities

$$\sum_u \theta^u \left(A_{iu}^T P^u A_{iu} + \alpha Q_{iu} \right) \le P^i \le A_{iv}^T P^v A_{iv} + Q_{iv} \quad (18)$$

for all $(i,v) \in \mathcal{E}$. A precise statement is given in the following corollary, stated similarly to Theorem 3.

Corollary 1: Assume $V^*(A_{iu} z, u) \le \gamma z^T Q_{iu} z$ for all z, i, u. Suppose there exist matrices $P^1, \ldots, P^m$ such that

$$(1-\epsilon) V^*(z,i) \le z^T P^i z \le V^*(z,i) \quad 0 \le \epsilon \le (1+\gamma)^{-2}$$

Let $\alpha = 1 - \epsilon(1+\gamma)^2$. Then, with $P_0^i = 0$ for $i \in \mathcal{N}$, the iterative convex inequalities (17) have solutions P_{j+1}^i and θ_{j+1}^u for every $j \ge 0$. All such solutions generate approximations to the optimal cost according to the inequalities

$$\begin{aligned} \alpha_j V^*(z,i) &\le z^T P_j^i z \le V^*(z,i) \\ \alpha_j &= \left[1 - (1+\gamma^{-1})^{-j} \right] \alpha \end{aligned}$$

Moreover, the control law $\mu_j(z,i) = \arg\min_u z^T (A_{iu}^T P_j^u A_{iu} + \alpha_j Q_{iu}) z$ has a value function V_{μ_j} satisfying $[\alpha + \gamma(1-\alpha_j)] V_{\mu_j} \le V^*$.

Remark 7. In general (17) is significantly more conservative than (16), but equivalence holds for example if the sum on the left has only two terms, i.e. if there are only two options for u at every switch instance.

Let us conclude the section with a major computational example to demonstrate the power of the proposed algorithms.

Example 2 First we generate a graph by randomly distributing 60 nodes in a square and defining edges by assigning two possible jumps from each node. The resulting graph is shown in Figure 1.

We will use 30 continuous states in each node. The step costs are chosen as $Q_{ij} = d_{ij} I$, where d_{ij} is the distance between two nodes. The dynamics, defined by the matrices A_{ij} will be chosen randomly, but with significant restrictions. Recall that if A_{ij} are all equal to the identity, then we recover the shortest-path-problem (provided that there is "target node" where it is possible to stay with step cost zero). The value iteration then works without need for approximation.

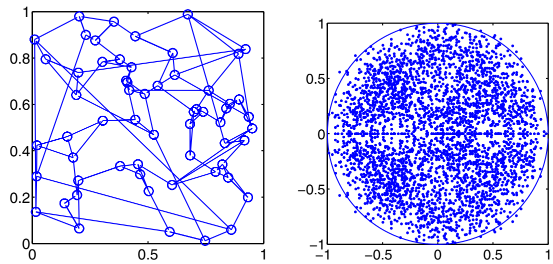

Fig. 1. A graph with 60 nodes has been randomly generated. From each node, there are two edges defining possible switches. For each of the 120 edges, a 30×30 matrix A_{ij} is used to define the dynamics of the continuous states along that edge. To the right, all eigenvalues of the 120 A_{ij} matrices are shown in one plot.

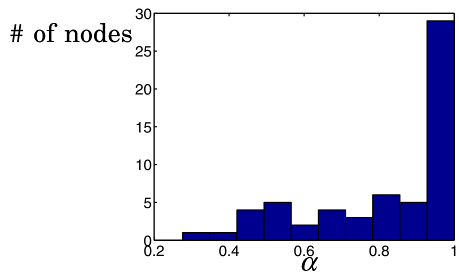

Fig. 2. For each node, the Hamilton-Jacobi-Bellman equation needs a certain amount of relaxation to be satisfied. This histogram reflects the fact that in most nodes of the graph, the equation can be satisfied with α around 0.9, much better than what is indicated by the worst case value $\alpha = 0.26$.

Similarly, if the A_{ij} are very small, then the cost function is essentially determined by the cost of the first step, and therefore close to quadratic. Relaxed value iteration will then work well with quadratic approximations.

We will consider a case somewhere in between these two extremes. Each A_{ij}-matrix is randomly generated, but with a spectrum varying within a disc of diameter 0.5 arbitrarily positioned with a center at most 0.9 from the origin. As a consequence, some of the matrices have eigenvalues outside the unit disc and are therefore expanding the continuous state in some directions. See the eigenvalue plot in Figure 1. Once the graph and matrices Q_{ij} and A_{ij} are defined, we are ready to run the value iteration algorithm. In each iteration let α_j be the maximal value of α for which (17) holds and let α^j be the maximal number of α for which the resulting P^i_j also satisfy (18). We then get the sequence

$$\begin{array}{ll} \alpha_1 = 0.58 & \alpha^1 = -7.12 \\ \alpha_2 = 0.34 & \alpha^2 = -4.13 \\ \alpha_3 = 0.28 & \alpha^3 = -0.42 \\ \alpha_4 = 0.29 & \alpha^4 = 0.26 \end{array}$$

Hence, after only four value iterations, we have found a quadratic approximation to the optimal cost satisfying

$$0.26 V^*(z,i) \le z^T P^i z \le V^*(z,i) \qquad \forall x, i \qquad (19)$$

and the corresponding control law yields a cost which is necessarily within a factor 4 from optimality:

$$V^*(z_0,i_0) \le \sum_k z(k)^T Q_{i(k)u(k)} z(k) \le \frac{1}{0.26} V^*(z_0,i_0)$$

It is interesting to look closer at some details of the solution. It turns out, as indicated in Figure 2, that in most of the nodes the inequalities (19) actually hold with a much higher value of α than 0.26. These are usually the nodes where one jump direction is clearly preferable to the other, regardless of the continuous state. Compare to Figure 3.

The source files of this example are available on the web site [1]. □

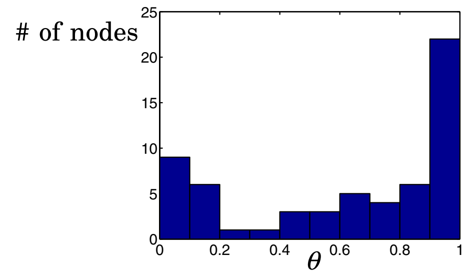

Fig. 3. For each node, there is a number θ^i, which appears in the left hand side of (18) and indicates the optimal switch. The histogram over the θ-values shows a preference for $\theta = 1$, which corresponds to switching to the nearest node in the graph. This is natural, since the nearest node has lowest step cost. Values between 0 and 1 can be interpreted as probabilities for jumps in different directions.

VI. Proofs

Proof of Proposition 2 From $V^*(f(x,u)) \le \gamma l(x,u)$ it follows that

$$\begin{aligned} V_1^*(x) &= \min_u \left[V_0^*(f(x,u)) + l(x,u) \right] \\ &\ge \min_u \left[\eta V^*(f(x,u)) + l(x,u) \right] \\ &\ge \min_u \left[\left(\eta + \frac{1-\eta}{\gamma+1} \right) V^*(f(x,u)) + \left(1 - \gamma \frac{1-\eta}{\gamma+1} \right) l(x,u) \right] \\ &= \frac{\eta\gamma+1}{\gamma+1} \min_u \left[V^*(f(x,u)) + l(x,u) \right] \\ &= \frac{\eta\gamma+1}{\gamma+1} V^*(x) \end{aligned}$$

The lower bound in (6) is obtained by repeating the argument j times. The upper bound in (6) is obtained analogously. □

Proof of Theorem 3 Define $V^s := (1-\epsilon\gamma)U$. Repeating the argument of Proposition 2, we have

$$\begin{aligned}
&\min_u \Big\{ V^s(f(x,u)) + l(x,u) \Big\} \\
&= \min_u \left[(1-\epsilon\gamma)U(f(x,u)) + l(x,u)\right] \\
&\geq \min_u \left[(1-\epsilon\gamma)(1-\epsilon)V^*(f(x,u)) + l(x,u)\right] \\
&\geq \min_u \left[((1-\epsilon\gamma)(1-\epsilon)+\epsilon)V^*(f(x,u)) + (1-\epsilon\gamma)l(x,u)\right] \\
&\geq (1-\epsilon\gamma)\min_u \left[V^*(f(x,u)) + l(x,u)\right] \\
&= (1-\epsilon\gamma)V^*(x) \geq (1-\epsilon\gamma)U(x) = V^s(x)
\end{aligned}$$

This proves the right inequality in (12). Similarly

$$\begin{aligned}
&\min_u \left[V^*(f(x,u)) + \alpha l(x,u)\right] \\
&\leq \min_u ([1-\epsilon(1+\gamma)]\, V^*(f(x,u)) + [\alpha + \epsilon(1+\gamma)\gamma]\, l(x,u)) \\
&= [1-\epsilon(1+\gamma)] \min_u \left[V^*(f(x,u)) + l(x,u)\right] \\
&= [1-\epsilon(1+\gamma)]\, V^*(x) \\
&\leq (1-\epsilon\gamma)(1-\epsilon)V^*(x) \\
&\leq (1-\epsilon\gamma)U(x) = V^s(x)
\end{aligned}$$

which proves the left inequality in (12). Hence, the convex constraints (11) on V_{j+1} are feasible at every step and Theorem 2 can be applied. □

VII. Conclusions

The main conclusion in this paper, as expressed in Theorem 3, is that finding approximately optimal control laws requires complex computations only if the cost function is complex.

Algorithms for control synthesis should therefore be designed to take advantage of this fact. They should give a simple answer quickly whenever there is one, and enter into more involved computations only when simpler alternatives have been exhausted.

Let us finally remark that although Example 5 was generated randomly within some restrictions, those restrictions were indeed essential. For a vast majority of problems in the class defined in section IV, quadratic approximations of the optimal cost will most likely not be sufficient for convergence of the value iteration.

VIII. Acknowledgments

The author is grateful to many colleagues for comments on this work, in particular the PhD students Peter Alriksson, Bo Lincoln, Ritesh Madan and Andreas Wernrud. The research was supported by the European Commission through grant IST-2001-33520 and the HYCON Network of Excellence. Much of the writing was done during a sabbatical supported by the Swedish Foundation of Strategic Research. An excellent sabbatical environment was provided by Control and Dynamical Systems at Caltech.

References

[1] http://www.control.lth.se/articles/article.pike?artkey=ran05.
[2] Randal Beard, George Saridis, and John Wen. Approximate solutions to the time-invariant Hamilton-Jacobi-Bellman equation. *Journal of Optimization Theory and Application*, 96(3), March 1998.
[3] Richard E. Bellman. *Dynamic Programming*. Princeton Univ. Press, 1957.
[4] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control, 2nd edition*. Athena Scientific, 2001.
[5] Vincent D. Blondel and Yurii Nesterov. Computationally efficient approximations of the joint spectral radius. *SIAM Journal of Matrix Analysis*. To appear.
[6] F. Borrelli, M. Baotic, A. Bemporad, and M. Morari. An efficient algorithm for computing the state feedback optimal control law for discrete time hybrid systems. In *American Control Conference*, pages 4717–4722, Denver, Colorado, USA, 2003.
[7] D.P. de Farias and B. Van Roy. The linear programming approach to approximate dynamic programming. *Operations Research*, 51(6):850–865, 2003.
[8] M. Falcone. A numerical approach to the infinite horizon problem of deterministic control theory. *Applied Mathematics and Optimization*, 15:1–13, 1987. Corrigenda, 23:213–214, 1991.
[9] R. Gonzales and E. Rofman. On deterministic control problems: an approximation procedure for the optimal cost. I. the stationary problem. *SIAM Journal on Control and Optimization*, 23:242–266, 1985.
[10] Sven Hedlund and Anders Rantzer. Convex dynamic programming for hybrid systems. *IEEE Transactions on Automatic Control*, 47(9):1536–1540, September 2002.
[11] L.V. Kantorovich. On a problem of Monge. *Uspekhi Mat. Nauk.*, 3:225–226, 1948.
[12] R.J. Leake and Ruey-Wen Liu. Construction of suboptimal control sequences. *SIAM Journal of Control*, 5(1):54–63, 1967.
[13] Bo Lincoln and Anders Rantzer. Relaxing dynamic programming. To appear in IEEE Transactions on Automatic Control.
[14] Bo Lincoln and Anders Rantzer. Suboptimal dynamic programming with error bounds. In *Proceedings of the 41st Conference on Decision and Control*, December 2002.
[15] Wei-Min Lu and J.C. Doyle. Robustness analysis and synthesis for nonlinear uncertain systems. *IEEE Transactions on Automatic Control*, 42(12):1654–1662, 1997.
[16] J. Lygeros, C. Tomlin, and S. Sastry. Controllers for reachability specifications for hybrid systems. *Automatica*, 35(3):349–370, 1999.
[17] D.Q. Mayne, J.B. Rawlings, C.V. Rao, and P.O.M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814, 2000.
[18] P.A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming Ser. B*, 96(2):293–320.
[19] S. Prajna, A. Papachristodoulou, and P.A. Parrilo. Introducing sostools: a general purpose sum of squares programming solver. In *Proceedings of the 41st IEEE Conference on Decision and Control*, Las Vegas, USA, 2002.
[20] S. Rachev and L. Rüschendorf. *Mass Transporation Problems, Volume I: Theory*. Probability and its Applications. Springer, 1998.
[21] Anders Rantzer. A dual to Lyapunov's stability theorem. *Systems & Control Letters*, 42(3):161–168, February 2001.
[22] Anders Rantzer and Sven Hedlund. Duality between cost and density in optimal control. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, 2003.
[23] Anders Rantzer and Mikael Johansson. Piecewise linear quadratic optimal control. *IEEE Trans. on Automatic Control*, 2000.
[24] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas. A metric for distributions with applications to image databases. In *Proceedings of the 1998 IEEE International Conference on Computer Vision*, Bombay, India, 1998.
[25] J. N. Tsitsiklis and B. Van Roy. Feature-based methods for large scale dynamic programming. *Machine Learning*, 22:59–94, 1996.
[26] R. Vinter. Convex duality and nonlinear optimal control. *SIAM J. Control and Optimization*, 31(2):518–538, March 1993.