Idempotent Methods for Control of Diffusions

Ben G. Fitzpatrick and Li Liu

***Abstract*—In this paper we discuss the application of max-plus arithmetic to stochastic control problems. The dynamic programming equation is not max-plus linear in the stochastic case, but a max-plus distributivity property permits efficient value function and control computation. We illustrate the technique by controlling the van der Pol equation.**

I. INTRODUCTION

The recent successes of max-plus and more general idempotent structures for attacking nonlinear control problems offer the potential for revolutionary improvements in our ability to design and implement nonlinear controls for real applications. The key ingredient in max-plus approaches to nonlinear control is the max-plus linearity of the Bellman equation of dynamic programming. One of the primary difficulties in the stochastic setting has been the lack of commutation of expectation and maximization. In this paper, we consider a max-plus distributivity property that, when coupled with max-plus finite element expansion, permits an efficient propagation of the Bellman equation. We illustrate the concepts with a simple nonlinear control example using the van der Pol equation.

II. PROBLEM FORMATION

We begin with a standard stochastic differential equation of the form

$$dX = f(X,u)dt + \sigma(X)dW, \quad X(t_0) = X_0$$

in which W is a standard Brownian motion. We define the cost functional

$$J(u,x_0,t_0) = E\left[\int_{t_0}^{t_f} g(X(t),u(t))dt\right]$$

which is to be maximized over admissible controls, which are progressively measureable and satisfy

$$u \in U(t_0,t_f) \subset \left\{u:[t_0,t_f] \to \mathbf{R}^m \mid E\int_{t_0}^{t_f}|u|^2 < \infty\right\}.$$

The Bellman equation of dynamic programming (DPE) takes the form

$$V(y,t) = \max_{u\in U(t,s)}\left\{E\left[\int_t^s g(X(\tau),u(\tau))d\tau + V(X(s;t,y,u),s)\right]\right\}$$

$$V(x_0,t_0) = \max_{u\in U(t_0,t_f)}\left(J(u,x_0,t_0)\right).$$

The value function is usually characterized through the Hamilton-Jacobi-Bellman partial differential equation (HJB)

$$\frac{\partial V(x,t)}{\partial t} + \max_u\left\{\sum_{ij}\partial_{x_i}\partial_{x_j}\left(\sigma(x,u)\sigma^T(x,u)V(x,t)\right) + f(x,u)\bullet\nabla V(x,t) + L(x,u,t)\right\} = 0,$$

in which the maximization is over $\mathbf{R}^m$. The optimal control is determined, as a function of the time and state variables, as the argument that attains the max in the HJB, or alternatively, piecewise as the argument that attains the max in the DPE. Numerical solution schemes tend to fall into two basic approaches: numerical PDE approaches (e.g., [1,2]) and stochastic process approaches [3]. We are seeking a third approach here, using idempotent algebraic techniques that have been demonstrated very effective in the deterministic setting (see, e.g., [4,5]). However, it is at this very point that the max-plus linearity fails for the stochastic control problem: the semigroup for backward propagation

$$S_{s,t}(\phi) = \max\left\{E\left[\int_t^s g(X(\tau),u(\tau)d\tau + \phi(X(s;t,y,u))\right]\right\}$$

is not (necessarily) linear, because the maximization and expectation cannot be interchanged in order. The crucial observation to move forward involves the max-plus distributive property, which we discuss after introducing max-plus algebra.

III. MAX-PLUS ALGEBRA AND DYNAMIC PROGRAMMING

The max-plus algebra, as discussed in [6,7,8,9,10,11], involves a redefinition of arithmetic operations, for computational and analytical benefit. We consider the real numbers, augmented by $-\infty$: $\mathbf{R}^- = \mathbf{R}\cup\{-\infty\}$. On this set, we define two operations, $\oplus$ and $\otimes$, by

$$a \oplus b = \max\{a,b\},$$

$$a \otimes b = a + b.$$

It is well known that $\mathbf{R}^-$ forms a commutative semi-ring under these operations. The additive identity is $-\infty$ while the multiplicative identity is 0. Except for the additive identity, every element has a multiplicative inverse, suggesting that one might be able to extend the structure to a field structure. However, addition in this semi-ring is idempotent, meaning that $a \oplus a = a$. It is important to note that the only rings satisfying additive idempotency are trivial; that is, the only element is the additive identity. Thus,

This work was supported in part by Air Force Office of Research grant FA9550-09-1-0524.

Dr. Ben Fitzpatrick is President of Tempest Technologies and Clarence J. Wallen Professor of Mathematics at Loyola Marymount University, 1 LMU Drive, Los Angeles, CA 90045.

Dr. Li Liu is a Senior Analyst with Tempest Technologies, 8939 S. Sepulveda Blvd, Los Angeles, CA 90045.

ISBN 978-963-311-370-7

extending the semi-ring to a ring (and hence a field) is not a possibility.

From these basic operations, we can build standard linear algebraic objects, such as matrices and vectors. If we consider an $n \times n$ array, A, of elements of $\mathbf{R}^-$ and a column vector, x, of n elements of $\mathbf{R}^-$, define the max-plus matrix-vector product $y = A \otimes x$ by

$$y_i = \bigoplus_{j=1}^{n} \left(A_{i,j} \otimes x_j\right) = \max_j \left\{A_{i,j} + x_j\right\}.$$

Similarly, we may define max-plus matrix multiplication and addition:

$$\left(A \otimes B\right)_{i,j} = \max_k \left\{A_{i,k} + B_{j,k}\right\},$$

$$\left(A \oplus B\right)_{i,j} = \max\left\{A_{i,j}, B_{i,j}\right\}$$

Raising a matrix to a power, then, is repeated applications of multiplication.

To illustrate the application of max-plus algebraic structure, we consider a standard nonlinear control problem. We begin with a dynamical system under control, of the form

$$\dot{x} = f(x,u), \quad x(t_0) = x_0$$

with a control objective given by

$$J(u, x_0, t_0) = \int_{t_0}^{t_f} g(x(t), u(t))dt,$$

which is to be maximized of the set of admissible control functions, $u \in U(t_0, t_f) \subset L^2(t_0, t_f)$. For a given control function and initial state, we denote the solution of the differential equation by $x(\bullet; t_0, x_0, u)$. To apply the finite element method to dynamic programming, we first examine the deterministic Bellman equation

$$V(y,t) = \max\left\{\int_t^s g(x(\tau), u(\tau))d\tau + V(x(s;t,y,u),s)\right\}$$

We define the family of operators $S_{s,t}$ by

$$S_{s,t}(\phi) = \max\left\{\int_t^s g(x(\tau), u(\tau)d\tau + \phi(x(s;t,y,u))\right\}.$$

$$= \bigoplus_u \left\{G(u) \otimes L_u(\phi)\right\}$$

in which G denotes the running cost integral and the operator L_u is defined by

$$L_u(\phi)(y) = \phi(x(s;t,y,u)).$$

By inspection, for each u, this operator is max-plus linear. Thus, as a max-plus-linear-combination of max-plus linear operators, the dynamic programming propagation operators are max-plus linear.

The dynamic programming propagation operator in the stochastic case,

$$S_{s,t}(\phi) = \max\left\{E\left[\int_t^s g(X(\tau), u(\tau)d\tau + \phi(X(s;t,y,u))\right]\right\},$$

however, is not max-plus linear. However, we may use the distributive property of multiplication over addition, to apply max-plus effectively.

Defining $I = \{1,2\}, J_I = \{(1,1),(1,2),(2,1),(2,2)\}$, we note that

$$\left(a_{1,1} \oplus a_{1,2}\right) \otimes \left(a_{2,1} \oplus a_{2,2}\right) = \bigotimes_{i=1}^{2} \bigoplus_{j=1}^{2} \left(a_{i,j}\right)$$

$$\left(a_{1,1} \otimes a_{2,1}\right) \oplus \left(a_{1,1} \otimes a_{2,2}\right) \oplus \left(a_{1,2} \otimes a_{2,1}\right) \oplus \left(a_{1,2} \otimes a_{2,2}\right)$$

$$= \bigoplus_{(j_1, j_2) \in J_I} \bigotimes_{i=1}^{2} \left(a_{i,j_i}\right)$$

More generally, as noted by McEneaney [Mc1], we have this fact. When $N = \{1,2,\cdots,n\}$, $M = \{1,2,\cdots,n\}$, and $J_M =$ the set of all ordered n-tuples of elements of M, then

$$\bigotimes_{i=1}^{n} \bigoplus_{j=1}^{m} \left(a_{i,j}\right) = \bigoplus_{(j_1,\cdots,j_n) \in J_M} \bigotimes_{i=1}^{n} \left(a_{i,j_i}\right).$$

Note that the number of summands has increased dramatically on the right side of the equality. This distributive property is generalized further in the following result (see [12]).

Theorem 1. Suppose that W and Z are separable metric spaces. Suppose that P is a finite Borel measure on W. Suppose that h is a measureable function on $W \times Z$ satisfying the following condition: for every $\varepsilon > 0$ and every $w \in W$, there exists a $\delta > 0$ such that

$$\left|h(w,z) - h(w',z)\right| < \varepsilon, \quad \forall z \in Z \text{ and } w' \in B_\delta(w).$$

Then,

$$\int_W \sup_{z \in Z} h(w,z)\, dP(w) = \sup_{f \in M(W,Z)} \int_W h(w, f(w))\, dP(w),$$

in which $M(W,Z)$ denotes the set of Borel measureable functions mapping W to Z.

To apply this result, we develop max-plus finite element approximations to the Bellman equations.

IV. FINITE ELEMENT APPROXIMATIONS

Max-plus finite elements (see, e.g., [4,5]) provide useful approximation tools for dynamic programming. Within the context of this work, we use the finite elements in conjunction with max-plus distributivity to approximate solutions to the stochastic Bellman dynamic programming equation. In this paper, we examine the linear elements

$$\psi_i(x) = -c_i\left|x - x_i\right|,$$

in which x_i are the element nodes, and c_i are scale parameters. A max-plus approximation of a function f takes the form

$$f(x) \approx \bigoplus_{k=1}^{N} a_k \otimes \psi_k(x) = \max_k \{a_k + \psi_k(x)\}$$

in which the weights a_i are defined by (again, see [4,5])

$$a_i = -\max_x \{\psi_i(x) - f(x)\}.$$

Note that a max-plus interpolation has an interesting and perhaps unintuitive structure. Figure 1 below illustrates the projection for an example function.

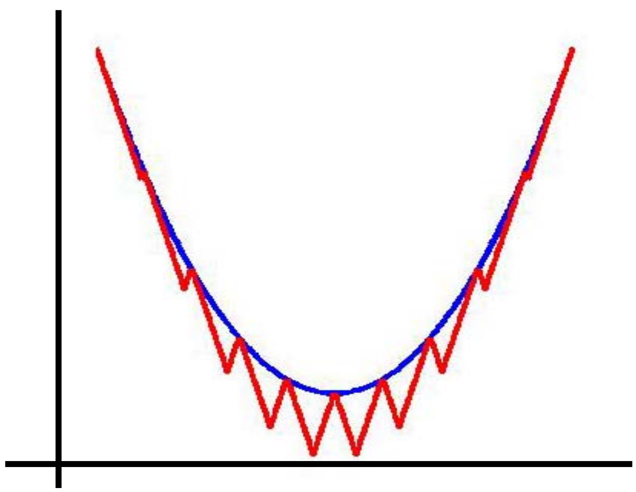

Figure 1. Blue curve is the original function; red curve is the max-plus finite element projection.

To apply the finite element method to dynamic programming, we first examine the deterministic Bellman equation

$$V(y,t) = \max\left\{ \int_t^s g(x(\tau), u(\tau))d\tau + V(x(s;t,y,u),s) \right\}$$

which is max-plus linear. We plug in the finite element expansion

$$V^N(x,t) = \bigoplus_{k=1}^{N} a_k(t) \otimes \psi_k(x) = \max_k \{a_k(t) + \psi_k(x)\}$$

so that the approximate Bellman equation becomes

$$\begin{aligned}\tilde{V}(y,t) &= \bigoplus_u \left\{ \int_t^s g(x(\tau),u(\tau))d\tau \otimes \left[\bigoplus_{i=1}^{N} a_k(s) \otimes \psi_k(x(s;t,y,u)) \right] \right\} \\ &= \bigoplus_u \bigoplus_{i=1}^{N} \left[\left(\int_t^s g(x(\tau),u(\tau))d\tau \right) \otimes a_k(s) \otimes \psi_k(x(s;t,y,u)) \right] \\ &= \bigoplus_{i=1}^{N} \left[a_k(s) \otimes \bigoplus_u \left(\int_t^s g(x(\tau),u(\tau))d\tau \otimes \psi_k(x(s;t,y,u)) \right) \right]\end{aligned}$$

in which $\tilde{V}$ is computed directly using the finite element expansion. The second step in the propagation process is to project $\tilde{V}$ onto the finite element basis:

$$a_i(t) = -\max_x \{\psi_i(x) - \tilde{V}(x,t)\}.$$

Recalling the semigroup propagation for the stochastic case, we have

$$S_{s,t}(\phi) = \max\left\{ E\left[\int_t^s g(X(\tau),u(\tau)d\tau + \phi(X(s;t,y,u)) \right] \right\}$$

into which we plug a finite element expansion

$$\begin{aligned}S_{s,t}(V^N) &= \max\left\{ E\left[\int_t^s g(X(\tau),u(\tau)d\tau \right.\right. \\ &\quad \left.\left. + \bigoplus_{i=1}^{N} a_i(s) \otimes \psi_i(X(s;t,y,u)) \right] \right\} \\ &= \bigoplus_u \left\{ \int_\Omega \bigoplus_{i=1}^{N} \left(\int_t^s g(X(\tau),u(\tau)d\tau \otimes a_i(s) \right.\right. \\ &\quad \left.\left. \otimes \psi_i(X(s;t,y,u))dP \right) \right\}\end{aligned}$$

Applying Theorem 1, we have

$$\begin{aligned}S_{s,t}(V^N)(y) &= \bigoplus_{u \in U(s,t)} \bigoplus_{Z \in M(\Omega, I_N)} \left\{ \int_\Omega \int_s^t g(X_\tau, u_\tau)d\tau \right. \\ &\quad \left. \otimes \int_\Omega a_{Z(\omega)}(t) \otimes \psi_{Z(\omega)}\big(X(t;s,y,u)\big)dP(\omega) \right\}\end{aligned}$$

in which the distributivity property of the theorem involves the set of random variables taking values in the set {1,2, …, N} for the interchange of expectation and maximization order. This propagation is then projected back onto the finite element basis through the relation as in the deterministic situation.

V. AN EXAMPLE PROBLEM: THE VAN DER POL OSCILLATOR

To illustrate the application of max-plus methods in a nonlinear control example, we consider the van der Pol oscillator [13]:

$$\ddot{x} + \mu(x^2 - 1)\dot{x} + x = u$$

Here μ is a constant, and u is the control.

The quantity we seek to minimize is

$$J(u) = \int_0^T (x^2 + \dot{x}^2 + au^2)dt$$

a standard quadratic regulator problem. We rewrite the second order equation as a first order system

$$\dot{x} = y$$

$$\dot{y} = \mu(1 - x^2) - x + u$$

Likewise, we rewrite the optimization problem as

$$\max \int_0^T -(x^2 + y^2 + au^2)dt = \min \int_0^T \ell(x,y,u)dt \,,$$

for $\ell(x,y,u) = x^2 + y^2 + au^2$.

The van der Pol system has a well known asymptotic behavior: the equilibrium point (0,0) is unstable, and there is a stable limit cycle with approximate radius 2, for small μ.

In order to illustrate the dynamic programming approach, we compare to the simple approach of linearization. That is, we linearize the system around the (0,0) equilibrium, and we apply standard linear quadratic regulator theory. We see in Figure 2 that the uncontrolled system oscillates at radius 2, while the LQR controller reduces the radius of the oscillation somewhat. The full nonlinear controller drives the system to (0,0).

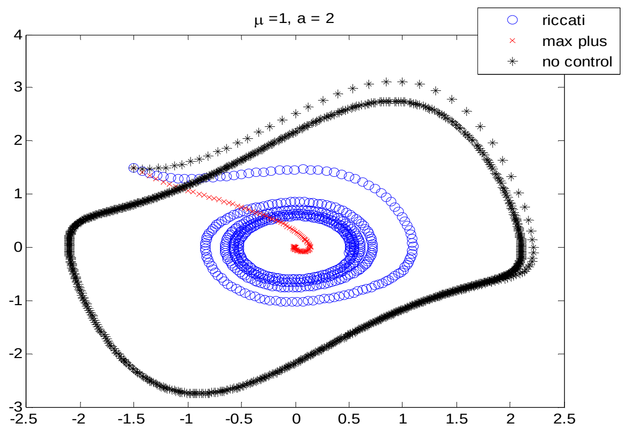

Figure 2. Uncontrolled (in black asterisk), linearized control (in blue circle), and nonlinear control (in red x) trajectories for the van der Pol system.

The controllers for the two systems have some similarities. In Figure 3, we compare the (long time asymptotic) controllers from the LQR solution and the value function computation.

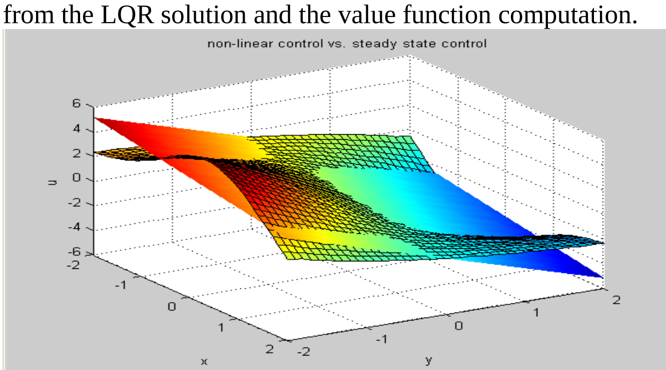

Figure 3. Linearized control (smooth surface) and nonlinear control (mesh surface) as function of the state variable (x,y).

To examine the stochastic controller, we add plant noise to the system:

$$dX = Ydt + \sigma dW_1$$

$$dY = \left(\mu(1 - X^2) - X + u\right)dt + \sigma dW_2$$

in which W_1, W_2 are standard Brownian motions. The resulting trajectories for one simulation of the uncontrolled system and the LQR controlled system are shown in Figure 4.

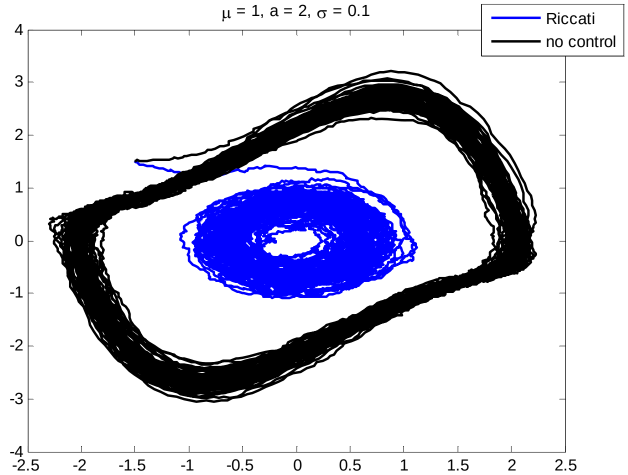

Figure 4. Uncontrolled (in black) and LQR controlled (in blue) trajectories of the van der Pol system.

The trajectories wander around the limit cycles due to the plant noise. If we add the nonlinearly controlled trajectory, the result is shown in Figure 5.

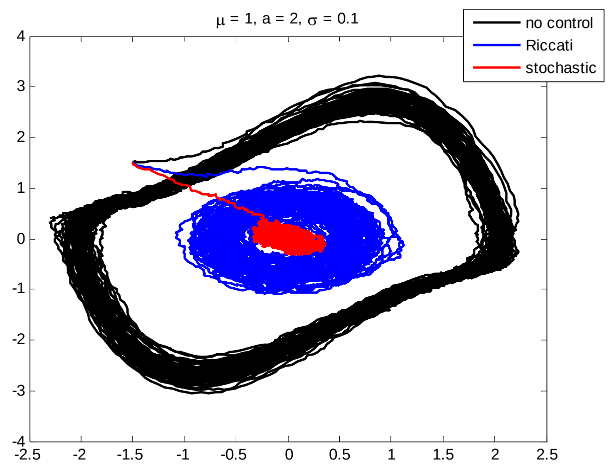

Figure 5. Uncontrolled (in black), LQR controlled (in blue), and nonlinear controlled (in red) trajectories of the van der Pol system.

The nonlinear control drives the system to (0,0) on average, with small variance, an improvement over the LQR control.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the Air Force Office of Scientific Research. This effort was supported in part by AFOSR grant FA9550-09-1-0524.

REFERENCES

[1] Barles, G., and E. R. Jakobsen, "On The Convergence Rate of Approximation Schemes for Hamilton-Jacobi-Bellman Equations," *Math. Modeling and Num. Anal.*, 36, No. 1, (2002), 33-54.

[2] Barles, G., and E. R. Jakobsen, "Error Bounds for Monotone Approximation Schemes for Hamilton-Jacobi-Bellman Equations," *Preprint*.

[3] Kushner, H., and P. DuPuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, New York (2001).

[4] Akian, M., S. Gaubert and A. Lakhoua. "The max-plus finite element method for solving deterministic optimal control problems: basic properties and convergence analysis," *SIAM J. Control and Optim.*, **47** (2008), 817–848.

[5] Akian, M., S. Gaubert and A. Lakhoua. "A max-plus finite element method for solving finite horizon determinsitic optimal control problems," *Proc. 16th International Symposium on Mathematical Theory of Networks and Systems* (2004).

[6] Baccelli, F.L. , G. Cohen, G.J. Olsder and J.-P. Quadrat, *Synchronization and Linearity*, John Wiley, New York (1992).

[7] Heidergott, B., G. Olsder, and J. van der Woude. *Max Plus at Work*, Princeton University Press, Princeton (2006).

[8] Kolokoltsov, V.N., and V.P. Maslov. *Idempotent Analysis and Its Applications*, Kluwer (1997).

[9] Litvinov, G. L. "The Maslov Dequantization, Idempotent and Tropical Mathematics: A Very Brief Introduction," ESI Preprint 1588, (2005).

[10] G.L. Litvinov, V.P. Maslov and G.B. Shpiz, "Linear Functionals on Idempotent Spaces: An Algebraic Approach," *Dolkady Math*, **58**, No. 3 (1998), 389-391.

[11] G.L. Litvinov, V.P. Maslov and G.B. Shpiz, "Idempotent Functional Analysis: An Algebraic Approach," *Math. Zametki*, **69**, No. 5 (2001), 758-797.

[12] McEneaney, W. M. "Distributed Dynamic Programming for Discrete-Time Stochastic Control and Idempotent Algorithms," to appear (2008).

[13] Arnold, V. I. *Ordinary Differential Equations*, MIT Press, Cambridge (1973).