Weakly chained matrices, policy iteration, and impulse control

SONAD 2016

May 27, 2016

Parsiad Azimzadeh

Joint work with Peter A. Forsyth

UNIVERSITY OF WATERLOO

Agenda

Motivation: combined stochastic and impulse control

Some important matrix families

Policy iteration

The nonexpansive problem

Motivation: combined stochastic and impulse control

Controlled stochastic differential equation

- $\{\tau_j\}$ stopping times with $0 = \tau_0 \leq \tau_1 \leq \tau_2 \leq \ldots \leq T$.
- $\mathfrak{W}$ denotes a standard Brownian motion.
- X solves a d-dimensional stochastic differential equation (SDE):

$$dX_t = \mu\left(t, X_t, w_t\right) dt + \sigma\left(t, X_t, w_t\right) d\mathfrak{W}_t \text{ on } \tau_j < t < \tau_{j+1}.$$

- w is a càdlàg adapted stochastic control taking values in some compact set W.
- ζ_j is a τ_j-adapted impulse taking values in the set $Z(\tau_j, X_{\tau_j -})$.

$$X_{\tau_j} = \Gamma\left(\tau_j, X_{\tau_j -}, \zeta_j\right) \text{ for } j > 0.$$

- Combined stochastic and impulse control:

$$\alpha = \left(w, \tau_1, \tau_2, \ldots, \zeta_1, \zeta_2, \ldots\right)$$

Value function

- The value under a particular control α is

$$J^{(\alpha)}(t,x) = \mathbb{E}^{(t,x)}\left[\int_t^T f(s, X_s^\alpha, w_s)\, ds + g(X_T^\alpha) + \sum_{\tau_j \leq T} K\left(\tau_j, X_{\tau_j -}^\alpha, \zeta_j\right)\right].$$

- The value function is the “best” value:

$$u(t,x) = \sup_\alpha J^{(\alpha)}(t,x).$$

David Cheriton School of Computer Science
University of Waterloo
Waterloo, Ontario

Dynamic programming equation

Using a dynamic programming argument (see [Seydel, 2009]) and a comparison principle, u can be shown[1] to be the (unique) viscosity solution of

$$\min \left(- \sup_{w \in W} \left\{ \frac{\partial u}{\partial t} + Ł^{w} u + f^{w} \right\}, u - \mathcal{M} u \right) = 0 \text{ on } [0, T) \times \mathbb{R}^{d}$$

$$\min \{ u - g, u - \mathcal{M} u \} = 0 \text{ on } \{ T \} \times \mathbb{R}^{d}$$

where $Ł^{w}$ is the infinitesimal generator of the SDE and

$$\mathcal{M} u (t, x) = \sup_{\zeta \in Z(t,x)} \{ u (t, \Gamma (t, x, \zeta)) + K (t, x, \zeta) \} .$$

[1] ...subject to some technical conditions :-)

Partial order on $\mathbb{R}^M$ and $\mathbb{R}^{M\times M}$

Definition

Let $\leq$ denote the element-wise partial order. Denote by sup the supremum induced by this order.

Example

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} < \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ but } \begin{pmatrix} 0 \\ 1 \end{pmatrix} \not\geq \not< \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Example

$$\sup_{\pi \in \mathbb{R}^2} \begin{pmatrix} -|\pi_1| \\ -(\pi_2)^2 + \pi_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1/4 \end{pmatrix}.$$

Numerical scheme

A numerical scheme for the dynamic programming equation takes the form

$$\text{find } v \in \mathbb{R}^M \text{ such that } \sup_{\pi \in \Pi} \{-A(\pi)v + b(\pi)\} = 0. \qquad \text{(Bellman problem)}$$

- Solved on a spatial grid with M nodes.
- v is the solution at each node.
- $\pi = (\pi_1, \ldots, \pi_M) \in \Pi$ is a node-wise control.
- $A(\pi) \in \mathbb{R}^{M \times M}$ (resp. $b(\pi) \in \mathbb{R}^M$) with i-th row depending only on π_i.

Classical results

Theorem

The problem

$$\text{find } v \in \mathbb{R}^M \text{ such that } \sup_{\pi \in \Pi} \{-A(\pi)v + b(\pi)\} = 0$$

(Bellman problem)

admits a unique solution whenever

- *A and b are bounded and*
- *$I - A(\pi)$ is a contraction with Lipschitz constant bounded uniformly with respect to π.*

Proof hint.

Write v as the fixed point of some equation and apply the Banach fixed point theorem. □

Classical results (detailed proof)

8 Motivation: combined stochastic and impulse control

Some important matrix families

Policy iteration

The nonexpansive problem

Proof.

Adding v to both sides of the equation yields

$$\sup_{\pi \in \Pi} \left\{ \left(I - A(\pi) \right) v + b(\pi) \right\} = v.$$

This suggests the fixed point iteration

$$v^{k+1} = f(v^k) \text{ where } f(v) := \sup_{\pi \in \Pi} \left\{ \left(I - A(\pi) \right) v^k + b(\pi) \right\}.$$

Moreover,

$$|f(v) - f(v')| \leq \sup_{\pi \in \Pi} |(I - A(\pi))(v - v')| \leq L |v - v'|$$

where $0 \leq L < 1$ is the Lipschitz constant of the previous slide. Now, apply the Banach fixed point theorem. □

David Cheriton School of Computer Science
University of Waterloo
Waterloo, Ontario

Problem!

The impulse term in

$$\min\left(-\sup_{w\in W}\left\{\frac{\partial u}{\partial t}+Ł^{w}u+f^{w}\right\},u-\mathcal{M}u\right)=0$$

causes $I-A(\pi)$ in the Bellman problem to be nonexpansive (and not necessarily contractive).

Some important matrix families

UNIVERSITY OF WATERLOO

Diagonal dominance

Definition

Row i of a complex matrix $A = (a_{ij})$ is strictly diagonally dominant (SDD) if

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|.$$

We say A is SDD if all of its rows are SDD.

Weakly diagonally dominant (WDD) is defined with $\geq$.

Example

The following matrix is SDD if $\epsilon > 0$ and WDD if $\epsilon = 0$:

$$\begin{pmatrix} 2 & & \\ -1 & 2+\epsilon & -1 \\ & -1 & 2 \end{pmatrix}$$

Matrix graph

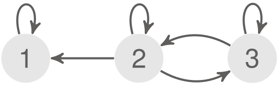
Definition

The (directed) graph of an $M \times M$ complex matrix $A = (a_{ij})$ is given by graph$(A) = (V, E)$ where

$$V = \{1, \ldots, M\} \text{ and } E = \{(i, j) \colon a_{ij} \neq 0\}.$$

Example

$$\begin{pmatrix} 2 & & \\ -1 & 2+\epsilon & -1 \\ & -1 & 2 \end{pmatrix}$$



Motivation: combined stochastic and impulse control

11 Some important matrix families

Policy iteration

The nonexpansive problem

David Cheriton School of Computer Science
University of Waterloo
Waterloo, Ontario

Weakly chained diagonally dominant matrix

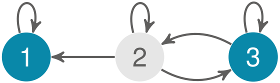
Definition

A complex square matrix A is said to be **weakly chained diagonally dominant (WCDD)** if

- A is WDD;
- each vertex in graph(A) has a path to an SDD vertex.

Example

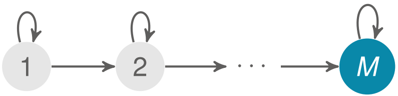
$$\begin{pmatrix} 2 & & \\ -1 & 2 & -1 \\ & -1 & 2 \end{pmatrix}$$



Another WCDD example

Example

$$\begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ & & & & 1 \end{pmatrix}$$



David Cheriton School of Computer Science
University of Waterloo
Waterloo, Ontario

WCDD nonsingularity

Lemma

A WCDD matrix is nonsingular.

Proof sketch (see [Azimzadeh and Forsyth, 2016] for a detailed proof).

- Suppose that there exists a WCDD matrix with eigenvalue $\lambda = 0$ and associated eigenvector $v \neq 0$.
- Pick i_0 such that $|v_{i_0}| = \max_i |v_i|$.
- By the definition of WCDD, there exists a path $i_0 \to i_1 \to \cdots \to i_k$ where i_k is an SDD vertex.
- Apply the Gershgorin circle theorem to establish that each vertex along this path is WDD; a contradiction.

□

Monotone matrix

Definition

A real square matrix A is monotone (in the sense of Collatz) if for all real vectors v, $Av \geq 0$ implies $v \geq 0$.

To see the motivation behind this definition...

1. Let $F \colon \mathbb{R}^M \to \mathbb{R}^M$ be a function satisfying

$$x < y \implies F(x) < F(y).$$

2. The contrapositive of the above proposition is

$$F(x) \geq F(y) \implies x \geq y.$$

3. If in addition F is linear,

$$F(x - y) \geq 0 \implies x - y \geq 0.$$

M-matrix

Definition

An M-matrix is a monotone matrix with nonpositive off-diagonals.

Theorem ([Azimzadeh and Forsyth, 2016])

The following are equivalent:

- *A is a WDD M-matrix;*
- *A is a WCDD matrix with positive diagonals and nonpositive off-diagonals.*

An M-matrix need not be WCDD:

$$\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

Policy iteration

Policy iteration

Recall the Bellman problem:

$$\text{find } v \in \mathbb{R}^M \text{ such that } \sup_{\pi \in \Pi} \{-A(\pi)v + b(\pi)\} = 0$$

POLICY-ITERATION(v^0)

1 **for** $\ell = 1, 2, \ldots$
2 Pick π^ℓ such that
$-A(\pi^\ell)v^{\ell-1} + b(\pi^\ell) = \sup_{\pi \in \Pi}\{-A(\pi)v^{\ell-1} + b(\pi)\}$
3 Solve $A(\pi^\ell)v^\ell = b(\pi^\ell)$ for v^ℓ

Idea: pick a quasi-optimal policy, improve the value, and repeat until convergence.

Convergence of policy iteration

Theorem

Suppose

1. $\pi \mapsto A(\pi)^{-1}$, *A, and b are bounded,*
2. *for all* $x \in \mathbb{R}^M$, *the supremum of* $\{-A(\pi)x + b(\pi)\}_{\pi \in \Pi}$ *is attained at an element* $\pi_x \in \Pi$, *and*
3. $A(\pi)$ *is a monotone matrix for all* $\pi \in \Pi$.

$(v^\ell)_{\ell \geq 1}$ *defined by* POLICY-ITERATION *converges from below to the unique solution v of the Bellman problem. Moreover, if* Π *is finite, convergence occurs in at most* $|\Pi|$ *iterations.*

In fact, we can even remove the second requirement by allowing the algorithm to pick the policy π^ℓ with some tolerable error ϵ^ℓ such that $\epsilon^\ell \searrow 0$ quickly enough (see appendix in [Azimzadeh and Forsyth, 2016]). In this case, convergence is only "nearly" from below.

Motivation: combined stochastic and impulse control

Some important matrix families

18 Policy iteration

The nonexpansive problem

David Cheriton School of Computer Science
University of Waterloo
Waterloo, Ontario

The nonexpansive problem

SDD/WDD splitting

Assumption

Each policy π is of the form

$$\pi = (\pi_1, \ldots, \pi_M) \text{ and } \pi_i = (\phi_i, \psi_i)$$

where $\phi = (\phi_1, \ldots, \phi_M)$ is arbitrary and $\psi_i \in \{0, 1\}$. $A(\pi)$ has nonpositive off-diagonals, and can be split as follows:

$$A(\pi) = (I - \operatorname{diag}(\psi)) \underbrace{(I - L(\phi))}_{SDD} + \operatorname{diag}(\psi) \underbrace{(I - B(\phi))}_{WDD}.$$

The SDD and WDD parts correspond to the contractive and nonexpansive parts of $A(\pi)$.

The nonpositive off-diagonals occur naturally in applications (Markov chains and monotone discretizations of PDEs).

Convergence of policy iteration in the splitting problem

Theorem ([Azimzadeh and Forsyth, 2016])

Suppose

- *A and b are bounded and*
- *for each $\pi = (\phi, \psi) \in \Pi$ and vertex i with $\psi_i = 1$, there exists a path in the graph of $N(\phi)$ from i to some vertex $j(i)$ with $\psi_{j(i)} = 0$.*

$(v^\ell)_{\ell \geq 1}$ defined by Policy-Iteration *converges from below to the unique solution v of the Bellman problem.*

Proof.

The above establishes that $A(\pi)$ is WCDD. Moreover,

$$\text{WCDD} \implies \text{M-matrix} \implies \text{monotone.}$$

Now we can apply a previous result to yield convergence. □

David Cheriton School of Computer Science
University of Waterloo
Waterloo, Ontario

Convergence under weaker conditions

The conditions of the previous theorem are too strong to cover all relevant cases.

Insight: policy iteration fails whenever $A(\pi^\ell)$ is singular.

We use the following idea:

1. establish uniqueness for the Bellman problem;
2. pick a subset $\mathit{\Pi}' \subset \mathit{\Pi}$ on which $A(\pi)$ is nonsingular and

$$\underbrace{\sup_{\pi \in \mathit{\Pi}'} \{-A(\pi)v + b(\pi)\} = 0}_{\text{modified problem}} \implies \underbrace{\sup_{\pi \in \mathit{\Pi}} \{-A(\pi)v + b(\pi)\} = 0}_{\text{original problem}}.$$

3. perform policy iteration on the modified problem.

Uniqueness under weaker conditions

Let $\mathbb{B}v = \sup_{\pi \in \Pi} \{B(\phi)v + b(\psi)\}$.

Theorem ([Azimzadeh and Forsyth, 2016])

Suppose

- *A and b are bounded and*
- *for each solution v of the Bellman problem and each vertex i, there exist integers $m(i)$ and $n(i)$ such that $0 \leq n(i) < m(i)$ and $[\mathbb{B}^{m(i)}v]_i < [\mathbb{B}^{n(i)}v]_i$.*

A solution of the Bellman problem is unique.

Proof sketch.

1. Let v be a solution with $-A(\pi^\ell)v + b(\pi^\ell) \nearrow 0$ as $\ell \nearrow \infty$.
2. Establish that $A(\pi^\ell)$ is WCDD (and hence monotone) for ℓ large enough.
3. Establish uniqueness using monotonicity.

□

Thank you for your attention :-)

UNIVERSITY OF WATERLOO

Bibliography

[Azimzadeh and Forsyth, 2016] Azimzadeh, P. and Forsyth, P. A. (2016).
Weakly chained matrices, policy iteration, and impulse control.
SIAM Journal on Numerical Analysis, 54(3):1341–1364.

[Seydel, 2009] Seydel, R. C. (2009).
Existence and uniqueness of viscosity solutions for qvi associated with impulse control of jump-diffusions.
Stochastic Processes and their Applications, 119(10):3719–3748.