New Error Bounds for Approximations from Projected Linear Equations

H. Yu* D. P. Bertsekas**

*Department of Computer Science
University of Helsinki

**Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology

European Workshop on Reinforcement Learning,
Lille, France, Jun. 30 – Jul. 4, 2008

Outline

Projected Equations and TD Type Methods

x^*: a solution of the linear fixed point equation

$$x = Ax + b$$

$\bar{x}$: the solution of the projected equation

$$x = \Pi(Ax + b)$$

Π: weighted Euclidean projection on subspace $S \subset \Re^n$, $\dim(S) << n$
Assume: $I - \Pi A$ invertible

Example: TD(λ) for approximate policy evaluation in MDP

- Solve a projected form of a multistep Bellman equation; linear function approximation of the cost function
- A: a stochastic or substochastic matrix
- ΠA is usually a contraction

Example: large linear systems of equations in general

Two Standard Error Bounds for the Contraction Case

$x^* - \bar{x}$: approximation error due to solving projected equation

Standard bound I (arbitrary norm): assume $\|\Pi A\| = \alpha < 1$, then

$$\|x^* - \bar{x}\| \leq \frac{1}{1-\alpha} \|x^* - \Pi x^*\| \tag{1}$$

Standard bound II (weighted Euclidean norm $\| \cdot \|_\xi$, use Pythagorean theorem, much sharper than I): assume $\|\Pi A\|_\xi = \alpha < 1$, then

$$\|x^* - \bar{x}\|_\xi \leq \frac{1}{\sqrt{1-\alpha^2}} \|x^* - \Pi x^*\|_\xi \tag{2}$$

- These are upper bounds on the ratios of

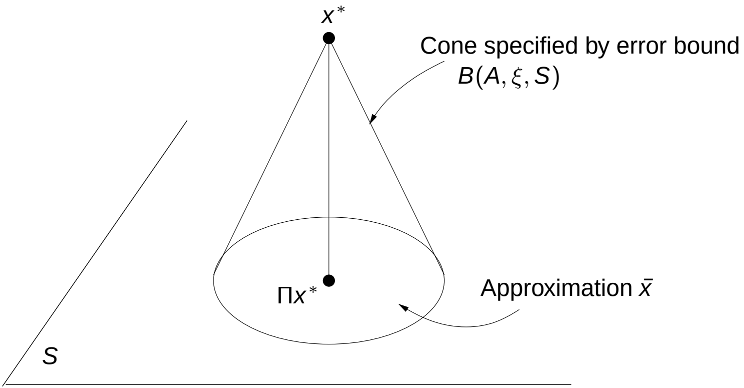
  amplification: $\dfrac{\|x^* - \bar{x}\|_\xi}{\|x^* - \Pi x^*\|_\xi}$ *bias-to-distance:* $\dfrac{\|\bar{x} - \Pi x^*\|_\xi}{\|x^* - \Pi x^*\|_\xi}$

- Our bounds will be in a similar form

$$\|x^* - \bar{x}\|_\xi \leq B(A, \xi, S)\, \|x^* - \Pi x^*\|_\xi \ ,$$

but *apply to both contraction and non-contraction cases*.

Illustration of the Form of Bounds



- $B(A, \xi, S) = 1 \;\Rightarrow\; \bar{x} = \Pi x^*$

Data-Dependent Error Analysis: Motivations

Motivation I: with or without contraction assumptions,

$$x^* - \bar{x} = (I - \Pi A)^{-1}(x^* - \Pi x^*) \tag{3}$$

How this equality is relaxed in the standard bounds:

- Standard bound I:

$$(I - \Pi A)^{-1} = I + \Pi A + (\Pi A)^2 + \cdots, \quad \|(\Pi A)^m\| \leq \alpha^m$$

- Standard bound II:

$$(I - \Pi A)^{-1} = I + \Pi A(I - \Pi A)^{-1}$$

$$\begin{aligned}\|x^* - \bar{x}\|_\xi^2 &= \|x^* - \Pi x^*\|_\xi^2 + \|\Pi A(I - \Pi A)^{-1}(x^* - \Pi x^*)\|_\xi^2 \\ &= \|x^* - \Pi x^*\|_\xi^2 + \|\Pi A(x^* - \bar{x})\|_\xi^2 \leq \|x^* - \Pi x^*\|_\xi^2 + \alpha^2\|x^* - \bar{x}\|_\xi^2\end{aligned}$$

Data-Dependent Error Analysis: Motivations

Motivation II:

$$(I - \Pi A)^{-1} = I + \Pi A (I - \Pi A)^{-1} = I + (I - \Pi A)^{-1} \Pi A$$

(i) Bound the term $(I - \Pi A)^{-1} \Pi A (x^* - \Pi x^*)$ *directly*
so that α will not be in the denominator

(ii) Seek computable bounds
with low order calculations involving small size matrices

Consider the technical side of (ii): some notation and facts

- $\Phi : n \times k$ matrix, whose columns form a basis of S; $\Xi = diag(\xi)$
- $k \times k$ matrices:

$$B = \Phi' \Xi \Phi \; , \qquad M = \Phi' \Xi A \Phi \; , \qquad F = (I - B^{-1} M)^{-1}$$

- $\Pi = \Phi(\Phi' \Xi \Phi)^{-1} \Phi' \Xi = \Phi B^{-1} \Phi' \Xi$;
the projected equation is equivalent to $\Phi r = \Phi B^{-1} \big(Mr + \Phi' \Xi b \big), r \in \Re^k$
- B and M can be computed easily by simulation.

Technical Lemmas for New Error Bounds

Lemma 1

$$(I-\Pi A)^{-1} = I + (I-\Pi A)^{-1}\Pi A = I + \Phi F B^{-1}\Phi'\Xi A \,. \tag{4}$$

Also, $I-\Pi A$ *invertible* $\iff F = (I - B^{-1}M)^{-1}$ *exists.*

Lemma 2

H and D: $n \times k$ *and* $k \times n$ *matrix, respectively. Then,*

$$\|HD\|_\xi^2 = \sigma\big((H'\Xi H)(D\Xi^{-1}D')\big)\,. \tag{5}$$

Apply the lemmas to bound $\|(I-\Pi A)^{-1}(x^* - \Pi x^*)\|_\xi$:

First bound: $(I-\Pi A)^{-1}\Pi A(x^* - \Pi x^*) \overset{\text{Lemma 1}}{=} \underbrace{\Phi F B^{-1}}_{H}\,\underbrace{\Phi'\Xi}_{D}\, A\,(x^* - \Pi x^*)$

$$\Longrightarrow \quad \|(I-\Pi A)^{-1}\Pi A(x^* - \Pi x^*)\|_\xi^2 \overset{\text{Lemma 2}}{\leq} \sigma(G_1)\|A\|_\xi^2\|(x^* - \Pi x^*)\|_\xi^2$$

where $G_1 = (H'\Xi H)(D\Xi^{-1}D') = B^{-1}F'BF$.

Main Results: First Bound

Theorem 1

$$\|x^* - \bar{x}\|_\xi \leq \sqrt{1 + \sigma(G_1)\|A\|_\xi^2}\ \|x^* - \Pi x^*\|_\xi \qquad (6)$$

where

- *G_1 is the product of $k \times k$ matrices*

$$G_1 = B^{-1}F'BF \qquad (7)$$

- *$\sigma(G_1) = \|(I - \Pi A)^{-1}\Pi\|_\xi^2$, so the bound is invariant to the choice of basis vectors of S (i.e., Φ).*

Notes:

- Thm. 1 equivalent to $\|(I - \Pi A)^{-1}\Pi A(x^* - \Pi x^*)\|_\xi \leq \|(I - \Pi A)^{-1}\Pi\|_\xi \|A\|_\xi \|x^* - \Pi x^*\|_\xi$
- Easy to compute, and better than the standard bound I
- Weaknesses: two over-relaxations; $\|A\|_\xi$ is required

Two Over-Relaxations in Theorem 1

1. $\Pi(x^* - \Pi x^*) = 0$ is not used.
   - Effect: degrade (to the standard bound I in the contraction case), if S nearly contains an eigenvector of A associated with the dominant real eigenvalue.
   - For applications in practice: orthogonalization of basis vectors w.r.t. the eigenspace to obtain sharper bounds

2. When ΠA is near zero, the bound cannot fully utilize this fact.
   - This is due to the splitting of Π and A in bounding $\|(I - \Pi A)^{-1}\Pi A\|$:

     $$\text{Thm. 1} \quad \Leftrightarrow \quad \|\Pi A + \Pi A(I - \Pi A)^{-1}\Pi A\|_\xi \leq \|\Pi + \Pi A(I - \Pi A)^{-1}\Pi\|_\xi \|A\|_\xi$$

   - Effect: when ΠA is near zero but $\|A\|_\xi = 1$, $\sigma(G_1) \approx \|\Pi\|_\xi^2 = 1$, and the bound tends to $\sqrt{2}$ instead of 1.

Apply the lemmas in a different way to sharpen the bound
$\Longrightarrow$ the second bound

Main Results: Second Bound

Use the fact $\Pi(x^* - \Pi x^*) = 0$,

$$\left\|(I - \Pi A)^{-1}\Pi A(x^* - \Pi x^*)\right\|_\xi = \left\|(I - \Pi A)^{-1}\Pi A(I - \Pi)(x^* - \Pi x^*)\right\|_\xi \leq \left\|(I - \Pi A)^{-1}\Pi A(I - \Pi)\right\|_\xi \|x^* - \Pi x^*\|_\xi$$

Relate the norm of the matrix to the spectral radius of a $k \times k$ matrix:

$$\left\|(I - \Pi A)^{-1}\Pi A(I - \Pi)\right\|_\xi^2 \overset{Lemma\ 1}{=} \Big\| \underbrace{\Phi F B^{-1}}_{H} \underbrace{\Phi' \Xi A(I - \Pi)}_{D} \Big\|_\xi^2 \overset{Lemma\ 2}{=} \sigma\big((H'\Xi H)(D\Xi^{-1}D')\big)$$

Notes:

- Incorporating the matrix $I - \Pi$ is crucial for improving the bound.
- $\|A\|_\xi$ is no longer needed.

Main Results: Second Bound

Theorem 2

$$\|x^* - \bar{x}\|_\xi \leq \sqrt{1 + \sigma(G_2)}\ \|x^* - \Pi x^*\|_\xi \qquad (8)$$

where

- *G_2 is the product of $k \times k$ matrices*

$$G_2 = B^{-1}F'BFB^{-1}(R - MB^{-1}M'), \quad R = \Phi'\Xi A \Xi^{-1}A'\Xi\Phi, \qquad (9)$$

- *$\sigma(G_2) = \|(I - \Pi A)^{-1}\Pi A(I - \Pi)\|_\xi^2$, so the bound is invariant to the choice of basis vectors of S (i.e., Φ).*

Proposition 1 (Comparison with the Standard Bound II)

Assume that $\|\Pi A\|_\xi \leq \alpha < 1$. Then, the error bound (8) is always no worse than the standard bound II, i.e., $1 + \sigma(G_2) \leq 1/(1 - \alpha^2)$.

Notes:

- The bound is tight in the worst case sense.
- Estimating R by simulation is less straightforward than estimating B and M; it is doable, except for TD(λ) with $\lambda > 0$.

MDP Applications and Numerical Comparisons of Bounds

Cost function approximation for MDP with TD(λ):

- A is defined for a pair of values (α, λ) by

$$A = P^{(\alpha,\lambda)} \overset{def}{=} (1-\lambda)\sum_{\ell=0}^{\infty} \lambda^{\ell} (\alpha P)^{\ell+1}$$

  discounted cases: $\alpha \in [0,1), \lambda \in [0,1]$
  undiscounted cases: $\alpha = 1, \lambda \in [0,1)$

Choices of the projection norm:

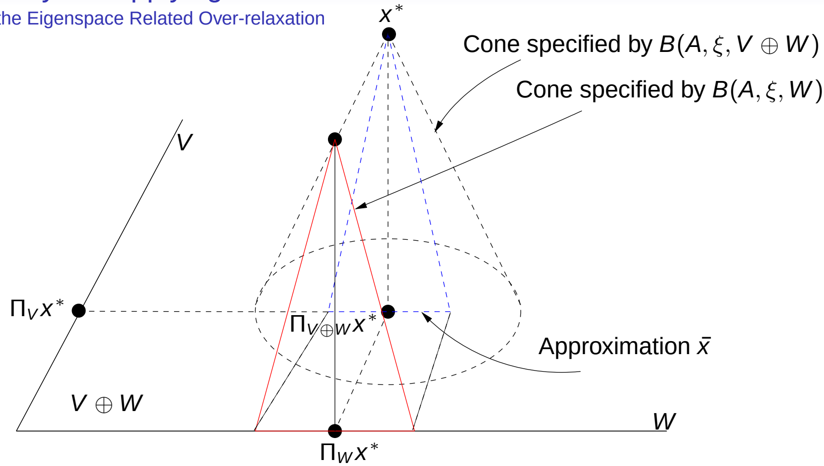
- W/o exploration: $\xi =$ invariant distribution of P; ΠA contraction
- W/ exploration: ξ determined by policies/simulations that enhance exploration; ΠA may or may not be contraction (λ needs to be chosen properly; LSTD(0) always safe to apply)

On applying Thm. 1:

- $e = [1, 1, \ldots, 1]'$: an eigenvector of A associated with the dominant eigenvalue $\frac{(1-\lambda)\alpha}{1-\alpha}$.
- To obtain a sharper bound, orthogonalize the basis vectors w.r.t. e (i.e., project them on $e^{\perp}$ – easy to do online).

Practical Ways of Applying Theorem 1

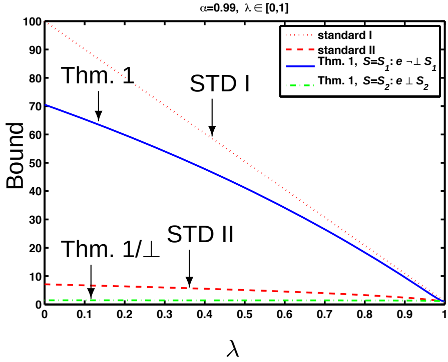
To Overcome the Eigenspace Related Over-relaxation

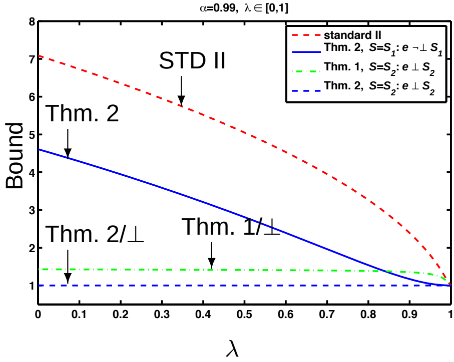


- Form the eq. satisfied by $x^* - \Pi_V x^*$ and solve its proj. eq. on W
  When V is an eigenspace of A, this is the same eq. as the original proj. eq. for x^*, and $\Pi_V x^*$ is not needed if this quantity is unimportant.
- Can replace $\Pi_V x^*$ with any vector in V (a guess of $\Pi_V x^*$).

Standard Bounds vs. Theorems 1 & 2 / Discounted

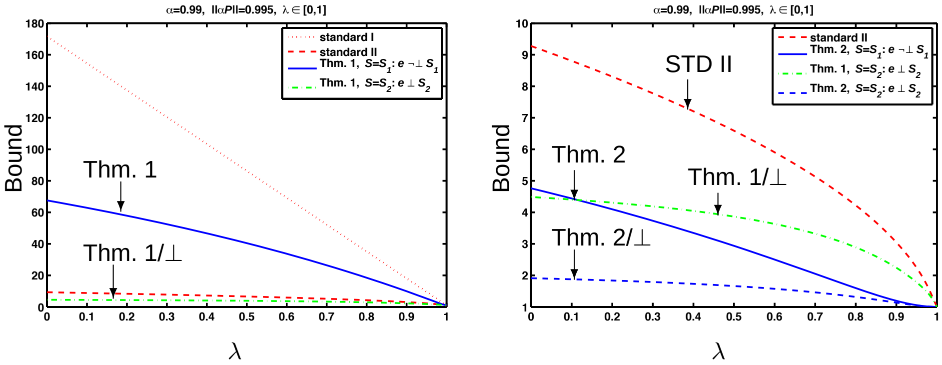
Markov chain: 200 states; $k = 50$; ξ: invariant distribution of P





Standard Bounds vs. Theorems 1 & 2 / Exploration Case

Markov chain: 200 states; $k = 50$; ξ: uniform

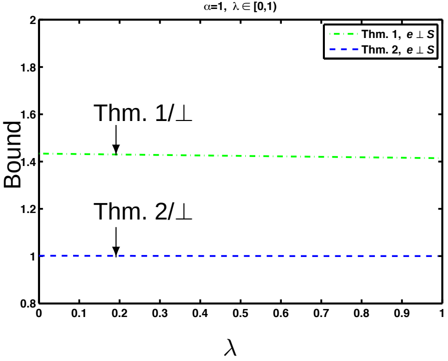


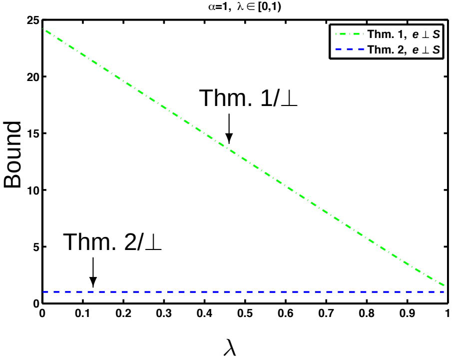
- In general, $\|\Pi A\|$ is not necessarily a contraction.
  need to choose λ properly; TD(0) can always be safely applied.
- The first bound needs $\|A\|$, so do the standard bounds for the contraction case.

Theorem 1 vs. Theorem 2 / Average Cost

Markov chains: 200 states; $k = 50$; ξ: invariant distribution of P

On the right: states of the Markov chain form two “tight clusters.”





- The standard bound II in this case is qualitative:

$$\|x^* - \bar{x}\|_\xi \le \frac{1}{\sqrt{1-\alpha_\lambda^2}} \|x^* - \Pi x^*\|_\xi \ ,$$

where $\alpha_\lambda < 1$ and $\alpha_\lambda \to 0$ as $\lambda \to 1$.

Discussion

New error bounds:

- Data dependent, w/o contraction assumptions
- Computable by simulation and low order calculations with small size matrices
- Sharper than the standard bounds (which are available only for the contraction case)
- Depend on A but not b (so they are valid for the worst case of b)
- Potential use in the MDP context:
  - Provide error bound for exploration policies
  - Aid in choosing the value of λ in TD
  - Aid in basis function evaluation and selection