On payoff games

Bruno Scherrer*

April 1, 2022

Abstract

Given a zero-sum two-player γ-discounted game with n states, ...

Consider a zero-sum two-player γ-discounted game with n states and m transitions, and its corresponding Bellman operators:

$$T_{\mu,\nu}v = r + \gamma P_{\mu,\nu}v$$
$$T_{\mu}v = \min_{\nu} T_{\mu,\nu}v$$
$$Tv = \max_{\mu} T_{\mu}v.$$

It is well-known that the optimal value v_* is the only fixed point of T and that any pair of stationary policies (μ_*, ν_*) such that $T_{\mu_*,\nu_*}v_* = v_*$ form a pair of optimal policies.

1 Facts

Let $v \geq Tv$. Compute

$$T^n v = T_{\vec{\mu},\vec{\nu}}v$$

For any policies μ and ν, there exist x, c such that $1 \leq c \leq n$ and

$$\mathbb{1}_x (P_{\mu,\nu})^c = \mathbb{1}_x.$$

Writing

$$M_{x,c}(v) = \{\ \mu\ ;\ \exists \nu \in \arg\min v_{\mu,\nu},\ \mathbb{1}_x (P_{\mu,\nu})^c = \mathbb{1}_x \}.$$

Note that

$$\bigcup_{x,c} M_{x,c} = M.$$

Take a $\mu \in M_{x,c}$. Then

$$\begin{aligned}
v_\mu(x) - v(x) &= \min_{\nu} v_{\mu,\nu}(x) - v(x) \\
&= \min_{\nu \in N_{x,c}(\mu)} v_{\mu,\nu}(x) - v(x) \\
&= \min_{\nu \in N_{x,c}(\mu)} \mathbb{1}_x (I - (\gamma P_{\mu,\nu})^c)^{-1} (T^c_{\mu,\nu}v - v) \\
&= \min_{\nu \in N_{x,c}(\mu)} \frac{1}{1-\gamma^c} \mathbb{1}_x (T^c_{\mu,\nu}v - v).
\end{aligned}$$

*INRIA, Université de Lorraine, bruno.scherrer@inria.fr

When running $\vec{\mu}$ against its adversary, there exists a x such that

$$v_{\vec{\mu}}(x) - v(x) \leq \frac{1}{n(1-\gamma)} \mathbb{1}_x (T^n v - v)$$

For all policies such $\mu_+ \in M_x(v)$,

$$\begin{aligned} v_{\mu_+}(x) - v_{\vec{\mu}}(x) &= v_{\mu_+}(x) - v(x) + v(x) - v_{\vec{\mu}}(x) \\ &\leq (1 - \frac{2}{n})(v_{\mu_+}(x) - v(x)) \end{aligned}$$