MATH/COMP 562 LECTURE NOTES
STABILITY THEORY

ADAM M. OBERMAN

1. Introduction

The idea here is to

- Rewrite the bellman equation in operator form
- understand how it is a contraction
- use the properties of contractions to show that value iteration converges
- Also understand error of policies in terms of the value of a given policy.

We state the Banach Fixed point theorem in an abstract metric space. Then we show how to estimate the error based on the residual. Finally, still in an abstract space, we show how to estimate the error for a perturbed operator. Later the perturbed operator will be a linearization of the residual.

Then we have a section on contraction in the maximum norm.

Finally, we consider operators which correspond to the Bellman operator for a Markov decision process. We make use of the results in the previous sections.

2. Application to CMP

We will use the contraction results above to show that

- Discounted Markov Decision process is a strict contraction
- The error in the value function can be estimated by the residual norm
- For an approximate value function x, the corresponding policy π satisfies $Tx = Sx$, where T is the Bellman Operator, and S is linearized Bellman operator corresponding to the policy. The value of the policy satisfies $Sy^* = y^*$
- The second theorem estimates the error between the value of the policy and the true value function x^*.

2.1. **Bellman equation in vector form.** When we derived the Bellman equation the setup was as follows:

- A finite number of states $s_1, \ldots, s_D$
- A finite number of actions, a
- The value function $U(s)$
- The transition probabilities $P(s' \mid a, a)$
- The rewards $R(s, a, s')$

Leading to

$$U(s) = \max_{a \in A(s)} \sum_{s'} P(s' \mid s, a)\left[R(s, a, s') + \gamma U(s')\right]$$

Date: April 6, 2023.

Now we convert this notation to vector form. Note that we can write the states as

- A finite number of states $i \in S = \{1, \dots, d\}$
- A finite number of actions $a \in A = \{1, \dots, J\}$
- The value function $v = (v_1, \dots, v_d), v \in \mathbb{R}^d$
- For each action, a, and state, i, a vector $w(a, i) \in \mathbb{R}^d$ corresponding to the transition probabilities.
- Simpler case: the reward is a scalar, $r(a, i) \in \mathbb{R}$
- Most general case: For each action, a, and state, i, a reward vector $r(a, i) \in \mathbb{R}^d$, so the reward depends on the current state, the action chosen, and the next state.
- A discount factor, $\gamma \in (0, 1]$.

So then, in vector form, the Bellman operator becomes

$$B(v)_i = \max_{a \in A} \{w(a, i) \cdot \gamma v + r(a, i)\}$$

In the more general rewards case, it is

$$B(v)_i = \max_{a \in A} \{w(a, i) \cdot (\gamma v + r(a, i))\}$$

We can include the first case in the second by allowing for a constant vector.

Note, the maximum occurs at each state, i, so it is not the maximum of the operators. So to write this in operator form, we need to define a policy as a function mapping states to actions, chosen from

$$\mathcal{A} = \{\pi : J \to A\}$$

We need to also define the probability matrix, and the reward matrix,

$$W(\pi)_i = w(\pi(i), i), \qquad R(\pi)_i = r(\pi(i), i)$$

whose rows are defined above.

This allows us to write

$$B(v) = \max_{\pi \in \mathcal{A}} \{W(\pi) \cdot (\gamma v + R(\pi))\}$$

$$\pi(v) = \arg\max_{\pi \in \mathcal{A}} \{W(\pi) \cdot (\gamma v + R(\pi))\}$$

The value function is given by solving the fixed point equation

$$v^* = B(v^*)$$

For a given policy, $\pi \in \mathcal{A}$, the restriction of the Bellman operator becomes the linear operator

$$L(\pi, v)_i = w(\pi(i), i) \cdot (\gamma v + r(\pi(i), i))$$

or

$$L(\pi, v) = W(\pi) \cdot (\gamma v + R(\pi))$$

So we can also write

$$B(v) = \max_{\pi \in \mathcal{A}} L(\pi, v)$$

The way to write the policy, $\pi(v)$, for a given value function, v, is to observe that the policy solves

$$B(v) = L(\pi, v)$$

For a fixed policy, π, the value function of the policy, v^π, comes from running the policy. So it is determined by the fixed point of the operator

$$v^\pi = L(\pi, v^\pi)$$

2.2. **Contractions for the bellman operators.** In order to apply the results of the previous section, we need to establish that the bellman operator is a contraction. Using the operator notation for the Bellman equation, we can establish that it is a contraction. We can write, for each π,

$$L(\pi, v) = W(\pi) \cdot (\gamma v + R(\pi))$$

Now each W has rows which are probability vectors. So by Corollary 4.6, $L(\pi, v)$ is a contraction. By Lemma 4.8, the Bellman operator with a discount is a strict contraction.

(Need to just add and subtract the discount factor) **This is a possible exercise**

3. Banach Fixed point theorem

Reference: `https://proofwiki.org/wiki/Banach_Fixed-Point_Theorem`

3.1. **Statement of the theorem.**

Definition 3.1. Let $(X, \|\cdot\|)$ be a complete metric space. The mapping $T : X \to X$ is a γ-contraction, if there is $0 \leq \gamma \leq 1$ such that

$$\|T(x) - T(y)\| \leq \gamma \|x - y\|, \qquad \forall x, y \in X \tag{1}$$

If, in addition, $\gamma < 1$, we say it is a strict contraction. We say $x^* \in X$ is a fixed point of T if

$$T(x^*) = x^*$$

Theorem 3.2 (Banach Fixed Point)**.** *Let $(X, \|\cdot\|)$ be a complete metric space. Let $T : X \to X$ be a strict contraction. Then there exists a unique fixed point of T.*

Proof. See the reference: `https://proofwiki.org/wiki/Banach_Fixed-Point_Theorem` □

Definition 3.3. Let $T : X \to X$ be a strict γ-contraction. Let x^* be the unique fixed point of T. Given x, we define the error, and the residual, by

$$e(x) = x - x^*, \qquad r(x) = T(x) - x$$

respectively,

The fixed point iteration is the sequence $x_0, x_1, \dots$ defined by

$$x^{n+1} = T(x^n), \tag{FPI}$$

3.2. **Error based on the residual.** In this section we are interested in showing that the sequence x^n defined by (FPI), converges to x^*. We are also interested in estimating the norm of the error in terms of the residual.

Theorem 3.4. *Suppose $T : X \to X$ is a strict γ-contraction with fixed point, x^*. Then for any $x \in X$,*

$$\|e(x)\| \leq \frac{\gamma}{1-\gamma} \|r(x)\|$$

in other words,

$$\|T(x) - x^*\| \leq \frac{\gamma}{1-\gamma} \|T(x) - x\| \tag{2}$$

In particular, for the (FPI),

$$\|x^{n+1} - x^n\| \leq \frac{(1-\gamma)\epsilon}{\gamma} \implies \|x^{n+1} - x^*\| \leq \epsilon$$

Proof.

$$\begin{aligned}\|T(x) - x^*\| = \|T(x) - T(x^*)\| &\leq \gamma\|x - x^*\| && \text{by } (\gamma\text{-contraction})\\ &\leq \gamma\left(\|x - T(x)\| + \|T(x) - x^*\|\right)\end{aligned}$$

Subtracting the second term on the right hand side,

$$(1-\gamma)\|T(x) - x^*\| \leq \gamma\|x - T(x)\|$$

which gives (3.4). The second result follows easily. □

3.3. **Error of a policy.** In the sequel, we will use the following result to estimate the error in the value of a policy. For now, we present the following abstract theorem.

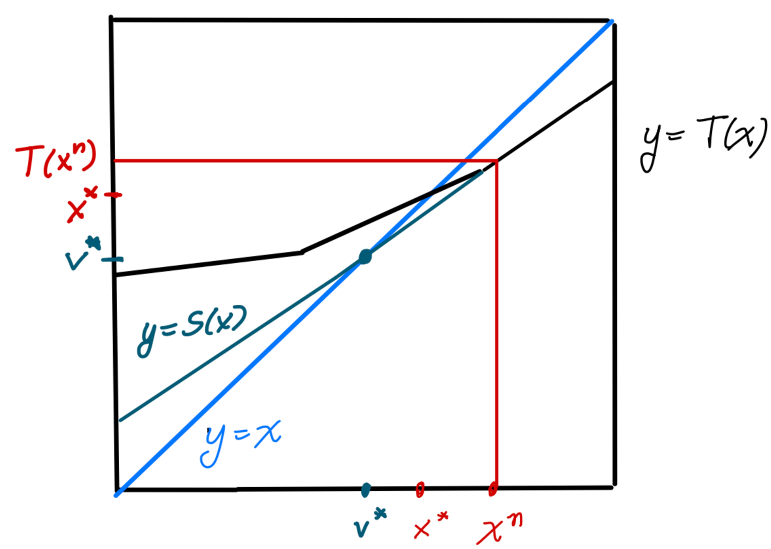

FIGURE 1. Illustration of the theorem

Definition 3.5.

Theorem 3.6. *Let $S, T : X \to X$, be strict contractions. Write,*

$$y^* = S(y^*), \quad x^* = T(x^*)$$

for the fixed points of S and T. Suppose $x \in X$ satisfies

$$T(x) = S(x),$$

Then

$$\|y^* - x^*\| \leq \frac{2\gamma}{1-\gamma}\|T(x) - x\|$$

Proof. Estimate

$$
\begin{aligned}
\|y^* - x^*\| &= \|y^* - S(x) + T(x) - x^*\| && \text{since } S(x) = T(x) \\
&\le \|y^* - S(x)\| + \|T(x) - x^*\| \\
&\le \frac{\gamma}{1-\gamma}\|S(x) - x\| + \frac{\gamma}{1-\gamma}\|T(x) - x\| && \text{by (3.4) applied to } S \text{ and to } T \\
&= \frac{2\gamma}{1-\gamma}\|T(x) - x\| && \text{since } S(x) = T(x)
\end{aligned}
$$

□

4. Contractions in Max Norm

Now we specialize to $X = \mathbb{R}^d$ with the maximum norm

$$\|x\| = \|x\|_\infty = \max_{i=1}^{d} |x_i|$$

While our final result will be for mapping $T : \mathbb{R}^d \to \mathbb{R}^d$, along the way we consider maps to $\mathbb{R}^n$ for n which is allowed to be different from d. We give a definition similar to (3.1), specialized to the maximum norm. Our definition coincides with γ-Lipschitz, but we require $\gamma \le 1$.

Definition 4.1. The function $g : \mathbb{R}^d \to \mathbb{R}^n$ is a γ-contraction (in the maximum norm) if there is $0 < \gamma \le 1$ such that

$$\|g(x) - g(y)\|_\infty \le \gamma \|x - y\|_\infty, \qquad \forall x, y \in \mathbb{R}^d \tag{γ-contraction}$$

If, in addition, $\gamma < 1$, we say it is a strict contraction.

Lemma 4.2. *Let $g : \mathbb{R}^d \to \mathbb{R}^n$, $g(x) = (g_1(x), \dots, g_d(x))$. Suppose each g_i is a γ_i-contraction. Then g is a γ-contraction, with*

$$\gamma = \max_i \gamma_i$$

Proof. We want to establish (γ-contraction). Given any x, y, let $j = j(x, y)$ be such that

$$\|g(x) - g(y)\|_\infty = |g(x) - g(y)|_j = |g_j(x) - g_j(y)|$$

Then

$$
\begin{aligned}
|g_j(x) - g_j(y)| &\le \gamma_j \|x - y\|_\infty && \text{by assumption on } g_j \\
&\le \gamma \|x - y\|_\infty && \text{by defn of } \gamma
\end{aligned}
$$

Since the choice of x, y was arbitrary, the result holds. □

4.1. **Monotone maps.**

Definition 4.3. Write $x \wedge y$, $x \vee y$ for the componentwise minimum, and maximum, respectively.

$$(x \wedge y)_i = \min(x_i, y_i), \qquad (x \vee y)_i = \max(x_i, y_i), \quad , i = 1, \dots, k$$

Given vectors $x, y \in \mathbb{R}^d$, write $x \le y$ for

$$x_i \le y_i, \quad i = 1, \dots, d$$

Redundantly, but for emphasis, for $g, h \in \mathbb{R}^n$, write $g \le h$ for

$$g_i \le h_i, \quad i = 1, \dots, n$$

Also write $\mathbf{1} = \mathbf{1}_n \in \mathbb{R}^n$ for the vector $\mathbf{1} = (1, \dots, 1)$.

By definition we have

$$x_1 \wedge x_2 \le x_i \le x_1 \vee x_2, \quad i = 1, 2$$

Definition 4.4. If $g : \mathbb{R}^d \to \mathbb{R}^n$ satisfies

$$x \le y \implies g(x) \le g(y), \quad \forall x, y \in \mathbb{R}^d$$

we say g is *monotone increasing* (or monotone when the meaning is clear).

Lemma 4.5. *If $g : \mathbb{R}^d \to \mathbb{R}^n$ is a γ-contraction, and if g is monotone increasing, then g satisfies the monotone-contraction inequality,*

$$\gamma \min_i (x_i - y_i, 0)\mathbf{1} \le g(x) - g(y) \le \gamma \max_i (x_i - y_i, 0)\mathbf{1} \tag{MC}$$

Proof. For each i

$$\begin{aligned} g_i(x) - g_i(y) &\le g_i(x \vee y) - g_i(y) && \text{since } g_i \text{ is monotone} \\ &\le \gamma \|x \vee y - y\|_\infty && \text{since } g_i \text{ is a } \gamma\text{-contraction} \\ &= \gamma \max_i (x_i - y_i, 0) \end{aligned}$$

So $g(x) - g(y) \le \gamma \max_i (x_i - y_i, 0)\mathbf{1}$.

The proof of the other inequality can be established in a similar way. □

Note, (MC) implies (γ-contraction), but not all contractions are monotone. For example, $g(x) = -x$ is monotone decreasing, and $g(x) = |x-1|/2$ is a contraction which is not monotone.

Corollary 4.6. *The following functions $g : \mathbb{R}^d \to \mathbb{R}$ all satisfy* (MC) *with $\gamma = 1$.*

- $g(x) = \max(x)$,
- $g(x) = \min(x)$
- $g(x) = w \cdot x + b$, *for* $w \ge 0$, $\sum_i w_i \le 1$.

The following functions $g : \mathbb{R}^d \to \mathbb{R}^d$ all satisfy (MC) *with $\gamma = 1$.*

- *The function* $g(x) = x$
- $g(x) = W \cdot (x + B)$, *where B is a matrix, and each row of W is a probability vector,* $w_j \ge 0$, $\sum_i w_{ji} = 1$.

Proof. It's easy to check that these functions are 1-contractions and monotone. □

The family of functions satisfying (MC) is closed under composition. We won't need the full generality, but we establish the following result.

Lemma 4.7. *Suppose $g_i : \mathbb{R}^d \to \mathbb{R}^n$ satisfies* (MC), *for $i = 1, \ \ldots, k$. Define $g(x) : \mathbb{R}^d \to \mathbb{R}^n$ componentwise, by $g(x)_j = \max_i (g_i(x))_j$. Then g also satisfies* (MC).

Proof. First note that (MC) can be established for each component. So it is enough to consider the case $n = 1$. Given any x, y, define

$$i(x) = \arg\max_i g_i(x),$$

Then

$$\begin{aligned} g(x) - g(y) &= g_{i(x)}(x) - g_{i(y)}(y) && \text{by definition of index} \\ &\le g_{i(x)}(x) - g_{i(x)}(y) && \text{since we have decreased } g(y) \\ &\le \gamma \max_i (x_i - y_i, 0) && \text{by assumption on } g_i \end{aligned}$$

The other inequality is established in a similar way. Since the choice of x, y was arbitrary, the result holds.

□

Lemma 4.8. *Suppose* $g : \mathbb{R}^d \to \mathbb{R}^d$ *satisfies* (MC) *with* $\gamma = 1$, *then the function* $h(x) = g(x) - \delta x$ *is a strict contraction, it satisfies* (γ-contraction), *with constant* $\gamma = 1 - \delta$.

Proof. From Corr 4.6, the function δx satisfies (MC) with $\gamma = \delta$, write the inequalities with x, y reversed,

$$\delta \min_i (y_i - x_i, 0)\mathbf{1} \leq \delta(y - x) \leq \delta \max_i (y_i - x_i, 0)\mathbf{1}$$

By assumption, g satisfies

$$\min_i (x_i - y_i, 0)\mathbf{1} \leq g(x) - g(y) \leq \max_i (x_i - y_i, 0)\mathbf{1}$$

Adding the two inequalities,

$$\delta \min_i (y_i - x_i, 0)\mathbf{1} + \min_i (x_i - y_i, 0)\mathbf{1} \leq h(x) - h(y) \leq \max_i (x_i - y_i, 0)\mathbf{1} + \delta \max_i (y_i - x_i, 0)\mathbf{1}$$

Calculating directly, or applying (MC) to the functions $\min, \max$ we obtain

$$(1 - \delta) \min_i (x_i - y_i, 0)\mathbf{1} \leq h(x) - h(y) \leq (1 - \delta) \max_i (x_i - y_i, 0)\mathbf{1}$$

which gives the (γ-contraction) inequality. □

Conclusion: the Bellman operator is a strict contraction.

Note: in the case where there is no discounting, without extra assumptions may not have contractions. (For example when $W = I$).