On Convergence and Contraction Mapping (By Mohammad Pezeshki)

Planning: The environment is known. (MDP is already given.)

On Convergence and Contraction Mapping

Planning: The environment is known. (MDP is already given.)

- Prediction: Given a policy, compute the value function
- Control: Find the best policy

On Convergence and Contraction Mapping

Planning: The environment is known. (MDP is already given.)

- Prediction: Given a policy, compute the value function
- Control: Find the best policy

Prediction:

- Iterative Policy Evaluation

Control:

- Policy Iteration
- Value Iteration

On Convergence and Contraction Mapping

Planning: The environment is known. (MDP is already given.)

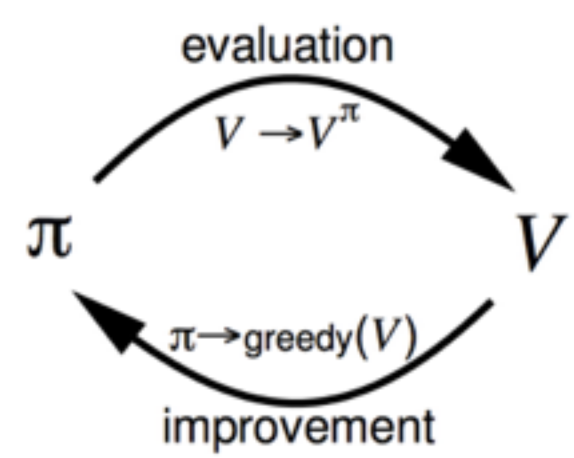
- → Prediction: Given a policy, compute the value function
- → Control: Find the best policy

Prediction:

- → Iterative Policy Evaluation → Bellman Expectation Equation

Control:

- → Policy Iteration → Bellman Expectation Equation, Greedy Policy Improvement
- → Value Iteration → Bellman Optimality Equation



On Convergence and Contraction Mapping

Iterative Policy Evaluation $\longrightarrow$ Bellman Expectation Equation

Policy Iteration $\longrightarrow$ Bellman Expectation Equation
Greedy Policy Improvement

Value Iteration $\longrightarrow$ Bellman Optimality Equation

Bellman Expectation Equation

Bellman Optimality Equation

On Convergence and Contraction Mapping

Iterative Policy Evaluation $\longrightarrow$ Bellman Expectation Equation

Policy Iteration $\longrightarrow$ Bellman Expectation Equation
Greedy Policy Improvement

Value Iteration $\longrightarrow$ Bellman Optimality Equation

Bellman Expectation Equation

$$v_{k+1}(s) = \sum_{a \in \mathcal{A}} \pi(a|s) \left(\mathcal{R}_s^a + \gamma \sum_{s' \in \mathcal{S}} \mathcal{P}_{ss'}^a v_k(s') \right)$$

$$\mathbf{v}^{k+1} = \boldsymbol{\mathcal{R}}^{\boldsymbol{\pi}} + \gamma \boldsymbol{\mathcal{P}}^{\boldsymbol{\pi}} \mathbf{v}^k$$

Bellman Optimality Equation

On Convergence and Contraction Mapping

Iterative Policy Evaluation $\longrightarrow$ Bellman Expectation Equation

Policy Iteration $\longrightarrow$ Bellman Expectation Equation
Greedy Policy Improvement

Value Iteration $\longrightarrow$ Bellman Optimality Equation

Bellman Expectation Equation

$$v_{k+1}(s) = \sum_{a \in \mathcal{A}} \pi(a|s) \left(\mathcal{R}_s^a + \gamma \sum_{s' \in \mathcal{S}} \mathcal{P}_{ss'}^a v_k(s') \right)$$

$$\mathbf{v}^{k+1} = \boldsymbol{\mathcal{R}}^{\boldsymbol{\pi}} + \gamma \boldsymbol{\mathcal{P}}^{\boldsymbol{\pi}} \mathbf{v}^k$$

Bellman Optimality Equation

$$v_{k+1}(s) = \max_{a \in \mathcal{A}} \left(\mathcal{R}_s^a + \gamma \sum_{s' \in \mathcal{S}} \mathcal{P}_{ss'}^a v_k(s') \right)$$

$$\mathbf{v}_{k+1} = \max_{a \in \mathcal{A}} \boldsymbol{\mathcal{R}}^{\boldsymbol{a}} + \gamma \boldsymbol{\mathcal{P}}^{\boldsymbol{a}} \mathbf{v}_k$$

On Convergence and Contraction Mapping

Iterative Policy Evaluation $\longrightarrow$ Bellman Expectation Equation

Policy Iteration $\longrightarrow$ Bellman Expectation Equation
Greedy Policy Improvement

Value Iteration $\longrightarrow$ Bellman Optimality Equation

Policy evaluation converges to V_π

Policy iteration converges to V_*

Value iteration converges to V_*

$\rightarrow$ ***Contraction Mapping Theorem***

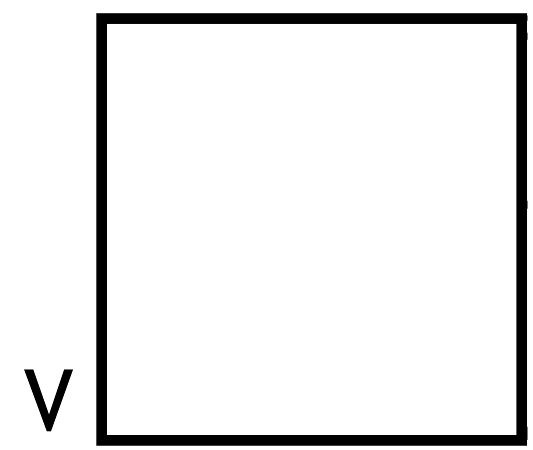
On Convergence and Contraction Mapping

Contraction Mapping

Space of all value functions: V

Operator: T: V → V

→ Shrinks the space

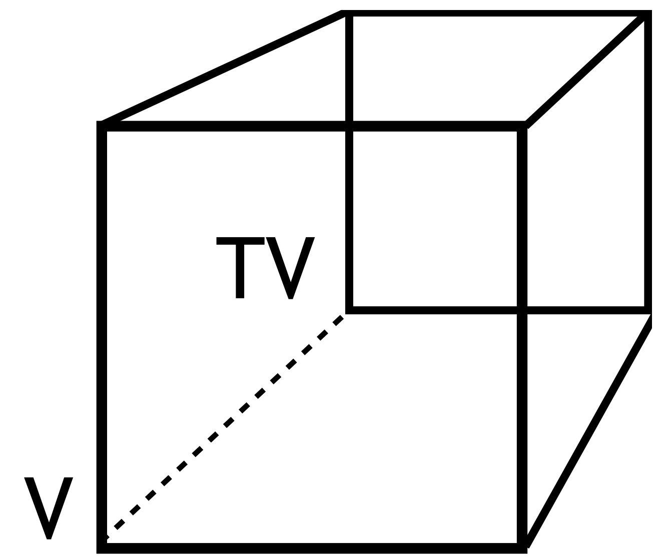


Contraction Mapping

Space of all value functions: V

Operator: T: V → V

→ Shrinks the space



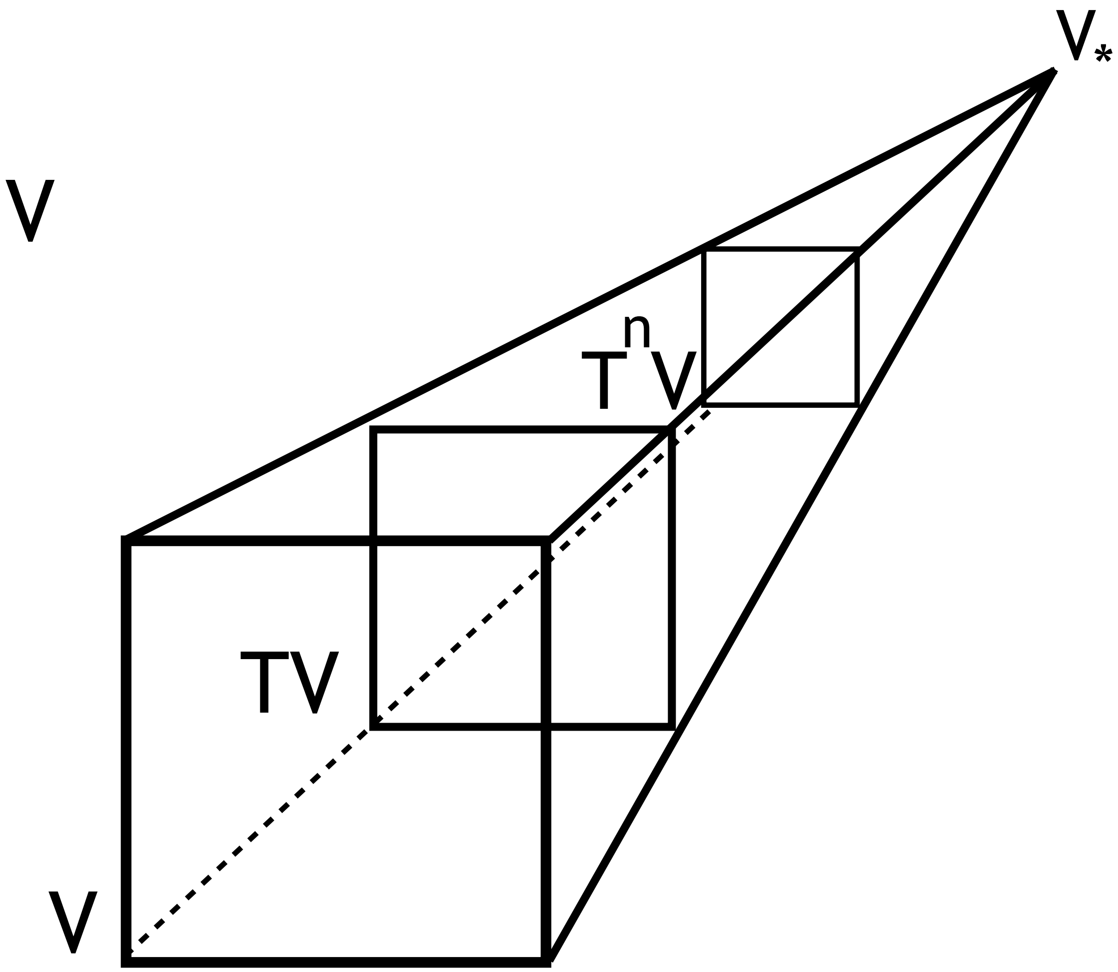
On Convergence and Contraction Mapping

Contraction Mapping

Space of all value functions: V

Operator: T: V $\rightarrow$ V

↳ Shrinks the space



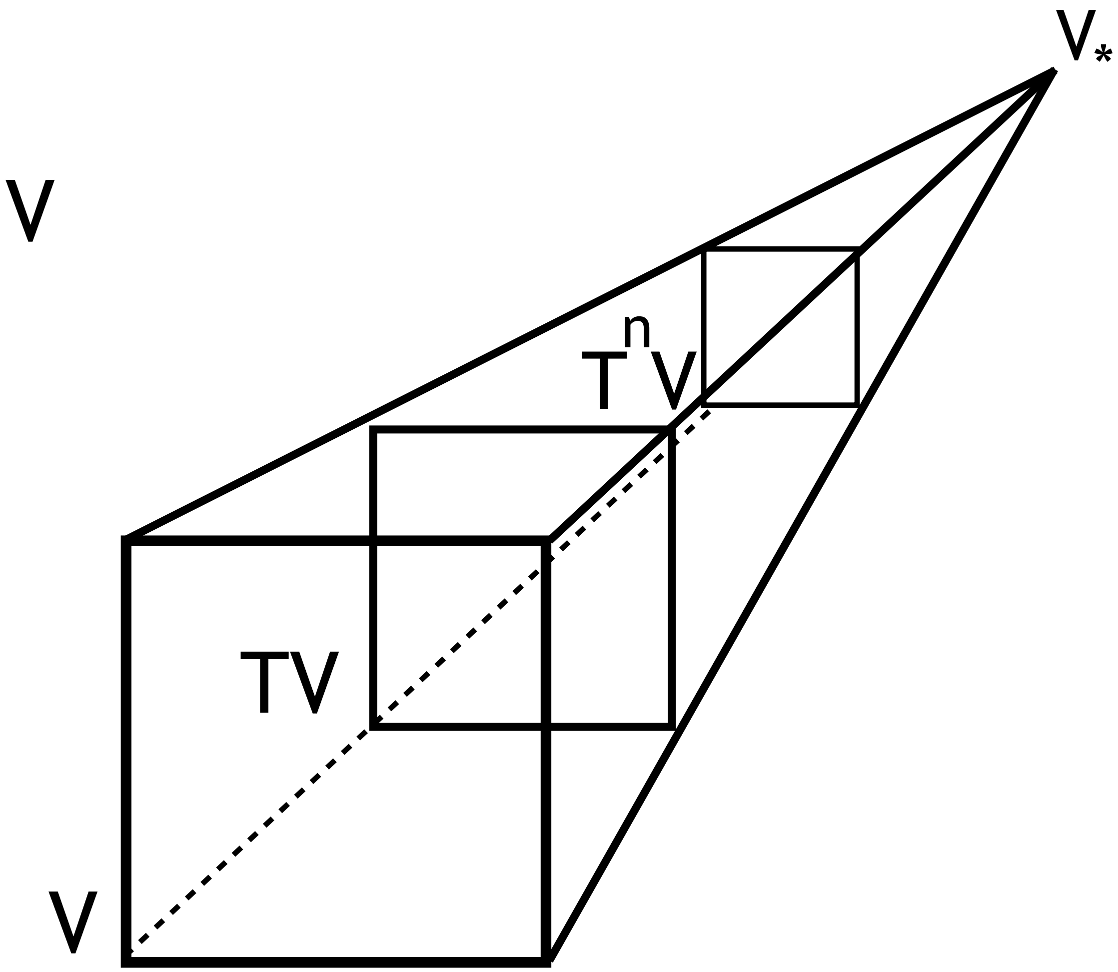
On Convergence and Contraction Mapping

Contraction Mapping

Space of all value functions: V

Operator: T: V $\rightarrow$ V

↳ Shrinks the space



a. It converges to unique V_*
b. V_* is fixed-point meaning $T V_* = V_*$

On Convergence and Contraction Mapping

Both "Bellman Expectation Equation"

and "Bellman Optimality Equation"

are γ-contractions.

→ They make value functions closer by at least a factor of γ.

Going to MATH:

Why **Bellman Equations** are contraction mapping?

Why **Contraction Mapping Theorem** holds?

On Convergence and Contraction Mapping

From Puter 1994: Why **Bellman Equations** are contraction mapping?

Proposition 6.2.4. Suppose that $0 \leq \lambda < 1$; then L and $\mathscr{L}$ are contraction mappings on V.

On Convergence and Contraction Mapping

From Puter 1994: Why **Bellman Equations** are contraction mapping?

Proposition 6.2.4. Suppose that $0 \leq \lambda < 1$; then L and $\mathscr{L}$ are contraction mappings on V.

Proof. Since S is discrete, L maps V into V. Let u and v be in V, fix $s \in S$, assume that $Lv(s) \geq Lu(s)$, and let

$$a_s^* \in \arg\max_{a \in A_s} \left\{ r(s,a) + \sum_{j \in S} \lambda p(j|s,a) v(j) \right\}.$$

Then

$$\begin{aligned} 0 \leq Lv(s) - Lu(s) &\leq r(s,a_s^*) + \sum_{j \in S} \lambda p(j|s,a_s^*) v(j) \\ &\qquad - r(s,a_s^*) - \sum_{j \in S} \lambda p(j|s,a_s^*) u(j) \\ &= \lambda \sum_{j \in S} p(j|s,a_s^*)[v(j) - u(j)] \leq \lambda \sum_{j \in S} p(j|s,a_s^*) \|v - u\| = \lambda \|v - u\|. \end{aligned}$$

Repeating this argument in the case that $Lu(s) \geq Lv(s)$ implies that

$$|Lv(s) - Lu(s)| \leq \lambda \|v - u\|$$

for all $s \in S$. Taking the supremum over s in the above expression gives the result. □

Shrinks the space by a factor of gamma (lambda) (discount factor)

On Convergence and Contraction Mapping

From Puter 1994: Why **Contraction Mapping Theorem** holds?

Theorem 6.2.3. (Banach Fixed-Point Theorem) Suppose U is a Banach space and $T: U \to U$ is a contraction mapping. Then

a. there exists a unique v^* in U such that $Tv^* = v^*$; and
b. for arbitrary v^0 in U, the sequence $\{v^n\}$ defined by

$$v^{n+1} = Tv^n = T^{n+1}v^0 \tag{6.2.10}$$

converges to v^*.

On Convergence and Contraction Mapping

From Puter 1994: Why **Contraction Mapping Theorem** holds?

Theorem 6.2.3. (Banach Fixed-Point Theorem) Suppose U is a Banach space and $T: U \rightarrow U$ is a contraction mapping. Then

a. there exists a unique v^* in U such that $Tv^* = v^*$; and
b. for arbitrary v^0 in U, the sequence $\{v^n\}$ defined by

$$v^{n+1} = Tv^n = T^{n+1}v^0 \tag{6.2.10}$$

converges to v^*.

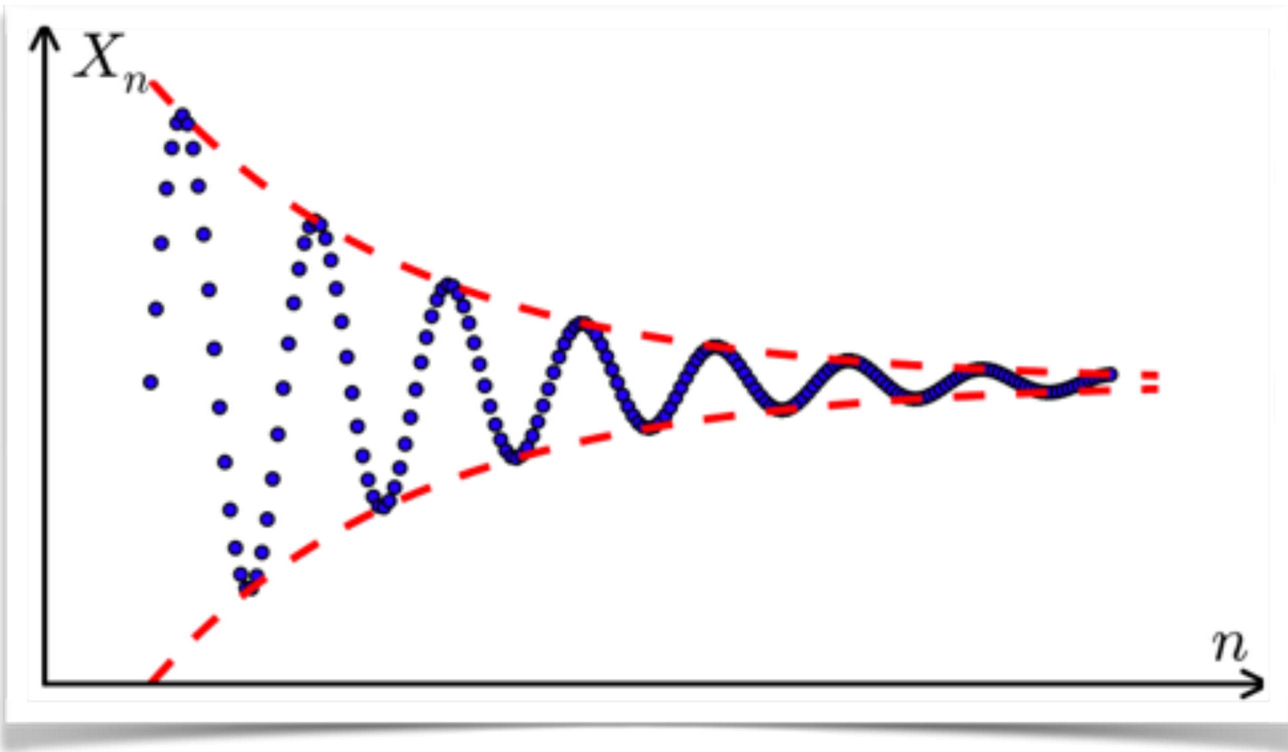
Banach space is a vector space with a metric that allows the computation of vector length and distance between vectors and is complete in the sense that a Cauchy sequence of vectors always <u>converges</u> to a well defined limit that is within the space.

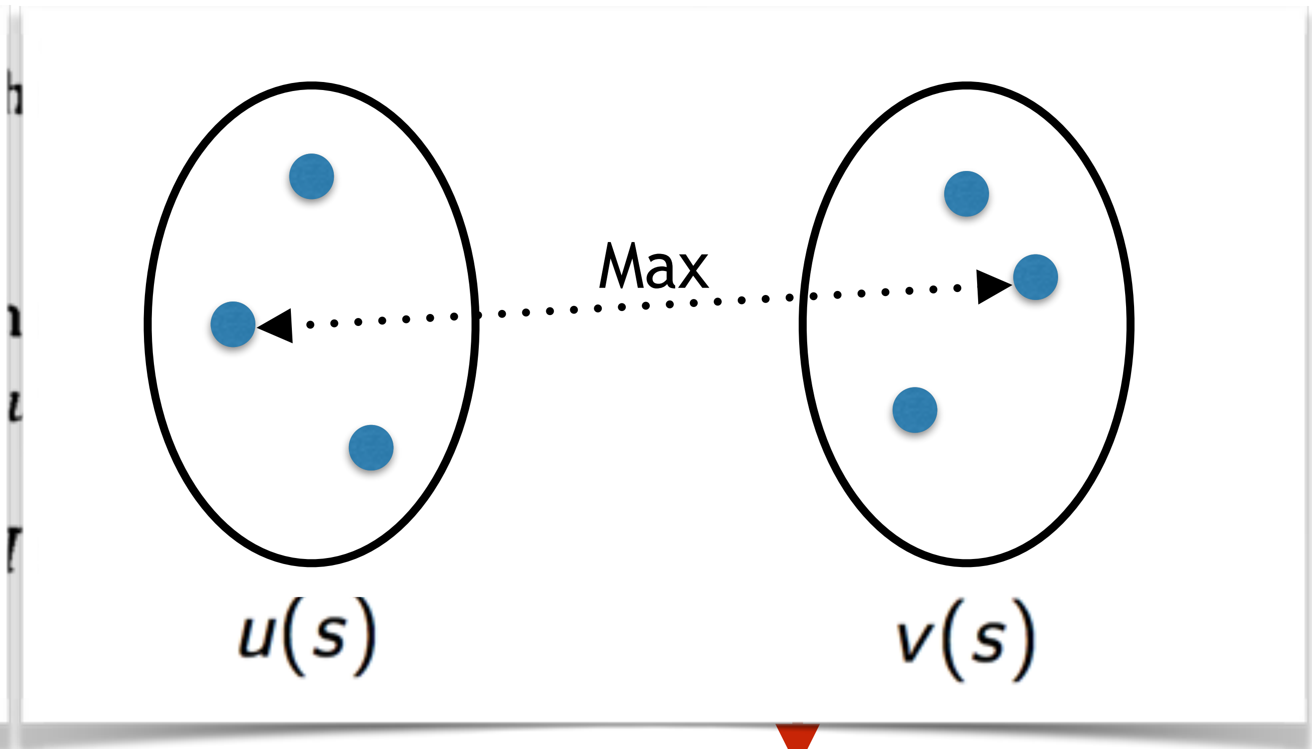
$$||u - v||_\infty = \max_{s \in \mathcal{S}} |u(s) - v(s)|$$

Cauchy sequence is a sequence whose elements become arbitrarily close to each other as the sequence progresses

On Convergence and Contraction Mapping

From Puter 1994: Why **Contraction Mapping Theorem** holds?





Banach space is a vector space with a metric that allows the computation of vector length and distance between vectors and is complete in the sense that a Cauchy sequence of vectors always converges to a well defined limit that is within the space.

$$||u - v||_\infty = \max_{s \in \mathcal{S}} |u(s) - v(s)|$$

Cauchy sequence is a sequence whose elements become arbitrarily close to each other as the sequence progresses

On Convergence and Contraction Mapping

From Puter 1994: Why **Contraction Mapping Theorem** holds?

Proof. Let $\{v^n\}$ be defined by (6.2.10). Then, for any $m \geq 1$,

$$\|v^{n+m} - v^n\| \leq \sum_{k=0}^{m-1} \|v^{n+k+1} - v^{n+k}\| = \sum_{k=0}^{m-1} \|T^{n+k}v^1 - T^{n+k}v^0\|$$

$$\leq \sum_{k=0}^{m-1} \lambda^{n+k}\|v^1 - v^0\| = \frac{\lambda^n(1-\lambda^m)}{(1-\lambda)}\|v^1 - v^0\|. \tag{6.2.11}$$

Since $0 \leq \lambda < 1$, it follows from (6.2.11) that $\{v^n\}$ is a Cauchy sequence; that is, for n sufficiently large, $\|v^{n+m} - v^n\|$ can be made arbitrarily small. From the completeness of U, it follows that $\{v^n\}$ has a limit $v^* \in U$.

We now show that v^* is a fixed point of T. Using properties of norms and contraction mappings, it follows that

$$0 \leq \|Tv^* - v^*\| \leq \|Tv^* - v^n\| + \|v^n - v^*\|$$
$$= \|Tv^* - Tv^{n-1}\| + \|v^n - v^*\| \leq \lambda\|v^* - v^{n-1}\| + \|v^n - v^*\|.$$

Since $\lim_{n\to\infty} \|v^n - v^*\| = 0$, both quantities on the right-hand side of the above inequality can be made arbitrarily small by choosing n large enough. Consequently, $\|Tv^* - v^*\| = 0$, from which we conclude that $Tv^* = v^*$.

We leave the proof of uniqueness of v^* as an exercise. □