A Guided tour of 'Metrics for MDPs with Infinite State Spaces'

Srinivas Venkattaramanujam

January 31, 2019

1 Introduction

Dynamic Programming problems are typically solved using methods such as value iteration or policy iteration. However, these methods are not suited for tasks involving large or continuous state space. One approach to solve the dynamic programming task in such settings is to reduce the state space by aggregation. The conventional approach to aggregate states has been to use heuristics without any theoretical guarantees but in [2] an approach to define the distance between states that is meaningful to aggregate states is proposed. In order to develop the distance function, tools from analysis, measure theory, optimal transport and the language of topology are employed. The objective of this course project is to gain familiarity with those tools and understand the arguments made.

2 Background

2.1 Measure Theory

The need for measure theory arised due to the Banach-Tarski paradox[9]. Banach-Tarski paradox states that a solid ball in a 3-dimensional space can be decomposed into a finite number of subsets and assembled in a different way to obtain two identical copies of the original solid ball. This does not match our intuition and hence some amendments have to made to the theory, the result of which is measure theory.

An algebra on a set S is just a collection of subsets of S with certain properties (axioms of algebra). Specifically,

Definition 1. *A collection Σ_0 of subsets in S is called an algebra on S if*

(i) $\Omega \in \Sigma_0$

(ii) $F \in \Sigma_0 \implies F^c \in \Sigma_0$

(iii) $F, G \in \Sigma_0 \implies F \cup G \in \Sigma_0$

An algebra is closed under finite unions and finite intersections because of the properties (ii) and (iii) along with de-morgan's law.

The structure imposed by an algebra is insufficient to study many questions of practical interest. For example, consider a random experiment of tossing a coin until the occurrence of a head. To answer the question whether the number of tosses is even, one needs to consider a set of countable elements. Such a set cannot be contained in the algebra (because the algebra is only closed under finite set operations). This motivates a structure called $\sigma-$algebra.

Definition 2. *A collection Σ of subsets of S is called a $\sigma-$algebra on S if Σ is an algebra on S and for $F_1, F_2 \cdots \in \Sigma$,*

$$\cup_{i\in\mathbb{N}} F_i \in \Sigma$$

Therefore, the $\sigma-$algebra is closed under countable unions and countable intersections.

Definition 3. *A pair (S, Σ) is called a measurable space. An element of Σ is called a Σ-measurable subset of S.*

Definition 4. *A measure is a function μ: $\Sigma \to [0, \infty]$ such that:*

(i) $f(\phi) = 0$

(ii) If $A_1, A_2, \cdots$ is a countable collection of disjoint $\Sigma-$measurable sets, then $\mu(\cup_{i\in\mathbb{N}} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$

The definition of measure enforces that the a measure is only countably additive and most importantly nothing is specified about uncountable addition.

Definition 5. *A triple (S, Σ, μ) is called a measure space.*

μ is called a finite measure if $\mu(S) < \infty$. Otherwise it is called a infinite measure.

A measure μ such that $\mu(S) = 1$ is called a probability measure.

Definition 6. *Let C be a class of subsets of S. Then $\sigma(C)$, the $\sigma-$algebra generated by C, is the smallest $\Sigma-$algebra that contains C.*

This definition is useful because it lets us start with some sets that are of interest (i.e to which we want to assign a measure) and construct a $\sigma-$algebra accordingly. This is non-empty because the power set is a $\sigma-$algebra. The smallest Σ-algebra containing C can be constructed by taking intersections of all the $\sigma-$algebras containing C. This argument utilizes the fact that the arbitrary intersections of $\sigma-$algebras is a $\sigma-$algebra, which is true because the intersection includes the entire set S, the complement of every intersecting element and also the countable unions of the intersecting elements.

An important application of the generating $\sigma-$algebra is the following:

Definition 7. *$\mathcal{B}(S)$, the Borel $\sigma-$algebra on S, is the $\sigma-$algebra generated by the family of open subsets of S.*

Definition 8. *(X, Σ) and (Y, T) be measurable spaces. A function $f : X \to Y$ is said to be measurable if*

$$f^{-1}(E) = \{x \in X | f(x) \in E\} \in \Sigma, \forall E \in T$$

That is, a measurable function takes back $T-$measurable sets in Y to $\Sigma-$measurable sets in X.

Definition 9. *A function $f : X \rightarrow Y$ is called Borel measurable if any open for any open $A \in Y$, $f^{-1}(A)$ is a Borel set.*

2.2 MDPs

Markov decision process is a mathematical model for sequential decision making under uncertainty.

Formally, an MDP is a 4-tuple (S, A, P, r) where S is the set of feasible states, A the set of feasible actions, P is the transition kernel with the Markov property and $r(s, a)$ is the reward function. The expected cumulative sum of discounted rewards starting from a state $s \in S$ is known the value of the state s. The objective of solving an MDP is to obtain a strategy, known as a policy π, to select an action $a \in A$ in every $s \in S$ such that the expected rewards is maximized.

For the purpose of [2], S is a complete separable metric space with the Borel sigma algebra Σ, A is a finite set of actions, $r : S \times A \rightarrow \mathcal{R}$ is a measurable function. The transition kernel $P(s, a, .) : \Sigma \rightarrow [0, 1]$ is a probability measure $\forall s \in S, a \in A$.

In [2] the following assumptions are made:

(i) The reward function is bounded i.e $B := sup_{s,s',a}|r_s^a - r_{s'}^a| < \infty$

(ii) For each $a \in A$, $r(., a)$ is continuous on S

(iii) For each $a \in A$, P_s^a is weakly continuous on S

These assumptions provide a foundation on which the theoretical arguments are made in order to construct the 'distance' function. We will see the usefulness of these assumptions in the later sections.

2.3 Partitions and Equivalence Relations

The definitions and tools introduced in this section are required to understand the definition of bisimulation and how the problem of state aggregation is reformulated in terms of bisimulation.

Definition 10. *A partition of a non-empty set X is a disjoint class $\{X_i\}$ of non-empty subsets of X whose union is X. X_i are called partition sets.*

Definition 11. *A binary relation R on a set S classifies whether two elements, x, $y \in S$ are related or not.*

Example 1. *$<$ defined in the usual sense on $\mathcal{R}$.*

An alternate characterization of the binary relation is defined as follows:

Definition 12. *A binary relation R is on a set X is a subset of $X \times X$. x, $y \in X$ are related if $(x, y) \in R$.*

Definition 13. *An equivalence relation, denoted by $\sim$, on a set X is any binary relation that satisfies the following properties:*

(i) $x \sim x \; \forall \; x \in X$ *(reflexivity)*

(ii) $x \sim y \implies y \sim x$ *(symmetry)*

(iii) $x \sim y$ *and* $y \sim z \implies x \sim z$ *(transitivity)*

For our purposes it is important to note that a partition induces an equivalence relation and vice-versa. Let $\{X_i\}$ be a a partition of X and let a binary relation $\sim$ be defined as follows: $x \sim y$ if both x and y belong to the same partition set X_i. $\sim$ defined as above satisfies all the properties of an equivalence relation and thus $\sim$ is an equivalence relation. To see that an equivalence relation induces a partition, let $\sim$ be an equivalence relation and for $x \in X$, let $\{x\} = \{y : y \sim x\}$. Because of the transitivity of $\sim$, it is easy to see that for *any* $z \in X$, $z \sim x \implies \{z\} = \{x\}$. Therefore, the $\{x\}$ and $\{z\}$ are either the same or disjoint. The collection of distinct $\{x\}$, thus forms a disjoint class $\mathbf{C}$ and each $\{x\}$ is non-empty because $x \sim x$. Thus, the C is a partition of X.

We have the tools necessary to introduce the notion of bisimulation, an equivalence relation that determines whether two states of an MDP are equivalent or not. Bisimulation can be equivalently seen as a notion to aggregate states thereby creating a partition of the state space where two states lie in the same partition only if they are equivalent.

Definition 14. *Let (S, A, P, r) be an MDP. An equivalence relation R is a bisimulation relation if the following properties are satisfied:*

$$sRs^{'} \iff \text{for every } a \in A, r_s^a = r_{s^{'}}^a \text{ and for every } X \in \Sigma(R), P_s^a(X) = P_{s^{'}}^a(X)$$

The largest of the bisimulation relation is called bisimilarity.

To understand the term *largest* in the definition of bisimilarity it is helpful to see bisimulation relations as a subset of $S \times S$.

2.4 Partial Ordered Sets and Lattices

The second type of relations that can be defined on a set that is useful for us is the order relation. Specifically we consider the partial order relation and concepts that arises as a consequence of partial orders. The term *partial* suggests that the order relation may be undefined for some pairs.

Definition 15. *A partial order is a relation $\leq$ and has the following properties:*

(i) $x \leq x$ *for* $\forall x \in S$

(ii) $x \leq y$ *and* $y \leq x \implies x = y$

(iii) $x \leq y$ *and* $y \leq z \implies x \leq z$

Definition 16. *Any non-empty set S endowed with a partial order is known as a partially ordered set.*

If for $x, y \in S$, if $x \leq y$ or $y \leq x$, then x and y are said to be comparable.

Definition 17. *If a relation P is a partial order and further possesses the property that any two elements $x, y \in S$ are comparable, then P is called a total order.*

Definition 18. *Let A be a non-empty subset of a partially ordered set P. An element $x \in P$ is called a lower bound of A if $x \leq a \forall a \in A$. A lower bound of A is greatest lower bound of A if it is greater than or equal to every lower bound of A.*

Greatest lower bound is unique if it exists. To see why, let $a1$ and $a2$ be both greatest lower bounds. By the definition of greatest lower bound, $a1 \leq a2$ and $a2 \leq a1$ and by property (2) of the partial order, $a1 = a2$.

An upper bound and the least upper bound of a non-empty subset A of a set X is defined similarly.

Definition 19. *A Lattice is a partially ordered set L in which every pair of elements have the least upper bound and the greatest lower bound.*

If $x, y \in S$, then the *least upper bound* and *greatest lower bound* are denoted by $x \vee y$ and $x \wedge y$ and known as *meet* and *join* respectively.

Definition 20. *A complete lattice is a lattice with the property that for every non-empty subset has a greatest lower bound and a least upper bound.*

Definition 21. *Let $(L, \preceq)$ be a partial order. A function $f : L \to L$ is monotone if $x \preceq x^{'} \implies f(x) \preceq f(x^{'})$. A point $x \in X$ is said to be a prefixed point if $f(x) \leq x$, a fixed point if $f(x) = x$ and a post-fixed point if $f(x) \geq x$.*

We have all the tools to understand *section* 2.2 of [2]. In order to get an intuition of *Theorem* 2.1, it is helpful to visualize a monotonically increasing function whose domain and range are the same on a closed interval $[a, b]$ in $\mathbb{R}$. If $f(a) = a$, then a is the least prefixed point and also the least fixed point, otherwise $f(a) > a$, let's say $f(a) = x$. Then, the prefixed point cannot occur in $[a, x)$. By applying the same reasoning in $[x, b]$ and by noticing that this process has to stop since $f(b) \leq b$, we see that the least prefixed point exists and the least prefixed point is the least fixed point. By similar argument, we also see that the postfixed point exists and greatest postfixed point is the greatest fixed point.

The definition 14 characterizes *bisimulation* in terms of equivalence relation whereas the *definition* 2.2 in [2] characterizes *bisimulation* in terms of fixed point theory. Let R_{rst} be the reflexive, symmteric and transitive closure eof R and REL be the lattice of binary relations without subset ordering. REL is closed because the greatest lower bound of any class C of binary relations can be obtained by taking the intersection of the sets in C and the least upper bound can be obtained by taking the intersection of all the upper bounds. The definition of *bisimulation* is as follows:

Definition 22. *Define $\mathcal{F} : REL \to REL$ by*

$$s\mathcal{F}(R)s^{'} \implies \forall a \in A, r_s^a = r_{s^{'}}^a \text{ and } \forall X \in \Sigma(R_{rst}), P_s^a(X) = P_{s^{'}}^a(X)$$

The greatest fixed point is bisimulation.

The function $\mathcal{F}$ is is monotonic because the partial order of REL is the subset ordering. Therefore, *Theorem* 2.1 guarantees the existence of the greatest fixed point, the *bisimulation*.

2.5 Metric Space, Open Sets, Closed Sets, Convergence and Continuity

The motivation to study metric spaces is to generalize the notion of continuity and convergence in $\mathbb{R}^n$ to arbitrary spaces. The definitions of continuity and convergence rely on a notion of distance and if we were to generalize the definition of convergence and continuity, then we need to generalize the notion of distance to the arbitrary spaces. The desired properties of the distance function are enforced in the definition of metric as given below.

Definition 23. *Let X be a non-empty set. A metric on X is a real function d of ordered pairs of elements of X which satisfies:*

(i) $d(x, y) \geq 0$ *and* $d(x, y) = 0 \iff x = y$ *(non-negativity)*

(ii) $d(x, y) = d(y, x)$ *(symmetry)*

(iii) $d(x, y) \leq d(x, z) + d(z, y)$ *(triangle inequality)*

Triangle inequality intuitively means that shortest path from x to y must not include visiting any point that is not the straight line from x to y. If $d(x, y) = 0$ does not imply $x = y$ but all other properties of the metric holds, then d is called a *semi-metric*.

Definition 24. *A metric space is a non-empty set X endowed with a metric d.*

Every non-empty set induces a metric space with the metric being $d(x, y) = 0$ if $x = y$ and 1 *otherwise*. One could also generalize the definition of =, for example as done by bisimulation. Let us denote the distance induced on the non-empty state space S of an MDP with = being the bisimulation as $\mathbb{1}_{\not\sim}$.

Definition 25. *Let X be a metric space with metric d. If $x_0 \in X$ and $r \in \mathcal{R}^+$ the open sphere $S_r(x_0)$ with center x_0 and radius r is the subset of X defined by*

$$S_r(x_0) = \{x : d(x, x_0) < r\}$$

An open sphere is non-empty since it always contains its center.

Definition 26. *A subset G of a metric space X is called an open set if for every $x \in G$, there exists an open sphere $S_r(x)$ centered on x and contained entirely in G.*

Any set G is open or not only in the context of a universal set X and is not intrinsic to the set G itself. For example $[0, 1]$ is not a open set in $\mathbb{R}$ but $[0, 1]$ is closed with respect to $[0, 1]$. The implication of this is that the full set X is open and trivially an empty set is open.

Theorem 1. *In any metric space X, each open sphere $S_r(x_0)$ is an open set.*

For any $x \in S_r(x_0)$, $S_{r-d(x,x_0)}(x)$ is an open sphere contained entirely in $S_r(x_0)$. Therefore, $S_r(x_0)$ is an open set.

Theorem 2. *Let X be a metric space. A subset G of X is open $\iff$ it is a union of open spheres.*

The proof of $\implies$ can be seen as follows: a set X is open implies that for each $x \in X \exists$ a open sphere centered on x and contained in X and the union of all such open spheres is X. The proof of $\impliedby$ is seen using *Theorem* 1.

Theorem 3. *Let X be a metric space. Then*

(i) arbitrary union of open sets in X is open

(ii) finite intersection of open sets in X is open

The proof of (1) involves decomposing each open set as a union of open spheres and using *Theorem* 2. For the proof (2), for each x in the intersection, there exists a open sphere (since we are taking intersections of open sets) and hence the intersection is open.

In order to see why the restriction of finite intersections is required, let us consider the following case. $\mathcal{Q}$ is countable and hence it can be listed. Let Q_i be the i^{th} rational and K_i be $(-\infty, Q_i) \cup (Q_i, +\infty)$. Then, $K = \cap_i K_i$ is $\mathbb{R} - \mathcal{Q}$, the set of irrationals, which is not open.

Definition 27. *Let X be a metric space with metric d and $A \subset X$. A point $x \in X$ is called a limit point of A if each open sphere centered on x contains at least one point in A other than x itself.*

A useful characterization of the limit point in terms of converging sequences will be given later.

Definition 28. *A subset F of a metric space X is a closed set if it contains each of its limit points.*

From the definition, it can be seen that a set is closed or not only in the context of some superset and is not intrinsic to the set itself. Also, the full set and the empty set are closed.

Theorem 4. *Let X be a metric space. A subset F is closed $\iff$ $F^{'}$ is open.*

To see $\implies$, we will assume $F^{'}$ is not open and show it leads to a contradiction. Since $F^{'}$ is not open, there exists a $x \in F^{'}$ s.t there does not exist a open sphere centered on x contained entirely in $F^{'}$. This implies that each open sphere centered on x contains a point in F. Therefore, x is a limit point of F and $x \notin F$. Since F is closed by assumption, it is a contradiction and $F^{'}$ is open.

To see $\impliedby$, observe that a limit point of F must contain at least one point in F different from itself for every open sphere centered on it. Therefore, it cannot be in $F^{'}$ and hence F is closed.

Definition 29. *Let X be a metric space with metric d and let*

$$\{x_n\} = \{x_1, x_2, \cdots x_n, \cdots\}$$

be a sequence of points in X. $\{x_n\}$ is convergent if $\exists x \in X$ s.t

(i) for each $\epsilon > 0$, $\exists$ a positive integer n_0 s.t $n \geq n_0 \implies d(x_n, x) < \epsilon$ or equivalently

(ii) for each open sphere $S_\epsilon(x)$ centered on x, $\exists$ a positive integer n_0 s.t $x_n \in S_\epsilon(x) \forall n \geq n_0$

A Cauchy sequence is a sequence $\{x_n\}$ that has the following property: for each $\epsilon > 0 \exists$ a positive integer n_0 s.t $m, n \geq 0 \implies d(x_m, x_n) < \epsilon$

Every convergent sequence is a Cauchy sequence but the converse is not true.

Definition 30. *A complete metric space is a metric space in which every Cauchy sequence is convergent.*

Whether a sequence is convergent or not also depends on the space in which it lies. Hence, any metric space that is not complete can be made complete by adjoining the 'missing' points.

An alternate characterization of the limit point in terms of sequences is as follows (Caution: I made up this definition):

Definition 31. *Let X be a metric space. A point $x \in X$ is a limit point of $A \subset X \implies \exists$ a sequence in A that converges to x.*

The converse is true only if the set of all points in the sequence is infinite as shown in the following theorem:

Theorem 5. *If a convergent sequence in a metric space has infinitely many distinct points, then its limit is a limit point of the set of points of the sequence.*

Finite number of points are problematic because after some n_0, the values are repeated and hence there will be no distinct points other than the convergent point itself when $n > n_0$ but limit point requires points other than itself to be contained in the set under consideration for every open ball centered on the limit point.

Definition 32. *Let X be a metric space and let Y be a subspace of X. Then Y is complete iff Y is closed.*

$\implies$:Since any limit point of Y implies that a convergent sequence in Y and since Y is complete, the limit point is in Y. $\impliedby$: Let $\{a_n\}$ be a Cauchy sequence in Y. If $\{a_n\}$ is finite, then the limit is in Y and hence Y is complete. Otherwise, the limit of $\{a_n\}$ is the limit point of the set of points and since Y is closed, is contained in Y. Therefore, Y is closed.

Definition 33. *Let X and Y be metric spaces with metrics $d1$ and $d2$ and let f be a mapping of X into Y. f is said to be continuous at a point x_0 in X if either of the following conditions hold:*

(i) for each $\epsilon > 0$ $\exists \delta > 0$ s.t $d_1(x, x_0) < \delta \implies d_2(f(x), f(x_0)) < \epsilon$ or

(ii) for each open sphere $S_\epsilon(f(x_0))$ centered on $f(x_0) \exists$ an open sphere $S_\delta(x_0)$ centered on x_0 s.t $f(S_\delta(x_0)) \subset S_\epsilon(f(x_0))$

Theorem 6. *The set of bounded real valued functions $\mathcal{B}(X)$ on a set X under uniform norm is a complete metric space.*

2.6 Lower Semi-Continuity

Definition 34. *Let (X, d) be a metric space and $f : X \to \mathbb{R}$. The lower contour set is corresponding to y is*

$$L(y) = f^{-1}((-\infty, y]) = \{x \in X | \}$$

Definition 35.

$$\liminf_{n\to\infty}\{x_n\} := sup\{inf\{x_m : m \geq n\} : n \geq 0\}$$

Definition 36. *A function f is lower semi-continuous at x_0 if for $\epsilon > 0$, $\exists$ a neighbourhood of x_0 s.t $f(x) \geq f(x_0) - \epsilon \forall x \in U$. Equivalently $\liminf_{x\to x_0} f(x) \geq f(x_0)$*

My intuition for lower semi-continuous function is that whenever there is a jump discontinuity at some point x_0, the function cannot jump upward at x_0.

The following three statements about an lsc function are equivalent:

(i) For any $y \in \mathbb{R}$, $L(y)$ is closed.

(ii) For any $y \in \mathbb{R}$, $L(y)^c$ is open.

(iii) For any $x \in X$, if the sequence $\{x_t\}$ in X converges to x then for any $\epsilon > 0$ $\exists$ T s.t $\forall t > T$, $f(x) \leq f(x_t) + \epsilon$ (the definition of lower semi-continuous function).

(1) $\implies$ (2) by 4. To see that (2) $\implies$ (3), let's prove using contraposition. Let $\{x_t\}$ in X converges to x and for some $\epsilon > 0$, $f(x) > f(x_t) + \epsilon$ i.e $f(x_t) < f(x) - \epsilon$. Therefore, $x_t \forall t \in L(f(x) - \epsilon)$ but $x \notin L(f(x) - \epsilon)$. Hence, $L(y)$ is not closed and $L(y)^c$ is not open which is a contradiction.

The idea of the proof is that if there can't be any gap between $f(x_t)$ and $f(x)$. (3) $\implies$ (1): Let $lim_{n\to\infty} x_n = x$ be a convergent sequence and let $x_i \in L(y)$ for some $y \in \mathbb{R}$. By (3), we have $f(x) \leq f(x_t) + \epsilon$. Since, $x_i \in L(y)$, $x_i < y \implies f(x) \leq y + \epsilon \implies f(x) \leq y \implies x \in L(y) \implies L(y)$ is closed.

Now we have all the tools to understand section 2.2 of [2].$\mathcal{M}$ be the set of semi-metrics which are lower semi-continuous on $S \times S$ and uniformly bounded endowed with a uniform norm. $\mathcal{M}$ is a complete lattice as explained in [2].

2.7 Optimal Transport

The motivation of topological spaces is to be able to study continuity without requiring a metric.

Definition 37. *Let X be a set. A class $\mathbb{T}$ of subsets of X is called a topology on X if it satisfies the following conditions:*

(i) The empty set ϕ and X are $\in \mathbb{T}$

(ii) Arbitrary union of $T in \mathbb{T}$ is in $\mathbb{T}$

(iii) Finite intersection of $T in \mathbb{T}$ is in $\mathbb{T}$

The properties of a topology $\mathbb{T}$ resembles the properties of open sets. This is because continuity can be characterized using open sets and thus the requirement of a metric for characterizing continuity is only through their use in open sets. By defining a topology in this manner, we assume that there some open sets given $\mathbb{T}$ with respect to which we want to characterize continuity. In fact, $T in \mathbb{T}$ are called open sets.

Definition 38. *A subset A of a topological space X is dense if for every $x \in X$, $x \in A$ or x is a limit point of A. Equivalently, A is dense of if the smallest closure of A is X.*

Definition 39. *A topological space is called separable if it contains a countable dense subset.*

Definition 40. *Polish space is topological space that is metrizable, complete and separable.*

Definition 41. *Let X, Y be Polish Spaces and $\mu \in \mathcal{P}(X)$ and $T : X \to Y$ a Borel function. The push forward measure $T_*\mu \in \mathcal{P}(Y)$ is defined by*

$$T_*\mu(A) = \mu(T^{-1}(A)) \text{ for every Borel set} A \subset Y$$

$\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$ and $c : X \times Y \to \mathbb{R}$ be a cost function which is non-negative and continuous. The formulation of optimal transport problem is as follows:

$$\textit{minimize} \int c(x, T(x))d\mu(x) \text{ s.t } T_*\mu = \nu$$

T is called the transport map. The restriction $T_*\mu = \nu$ can be interpreted as the total mass assigned to a set $A \in Y$ must be equal to the cumulative mass of points that were transferred from $T^{-1}(A) \in X$. The cost term in the integral considers the cost of moving all the mass in x to some $T(x)$ (hence the term 'map' in Transport map). This is problematic because when μ and ν do not agree, there is no solution to the optimization problem. Hence, the case of allowing a part of $x \in X$ to be moved to $y \in Y$ was introduced by Kantorovich.

Definition 42. *A measure $\gamma \in \mathcal{P}(X, Y)$ is an optimal plan from μ to ν if*

$$\pi^1_*\gamma = \mu$$
$$\pi^2_*\gamma = \nu$$

The problem of finding the optimal plan then is formulated as:

$$\int c(x, y)d\gamma(x, y) \text{ s.t } \gamma \text{ is transport plan}$$

Definition 43. *That is, the transport plan is optimal if and only if there does not exist another transport map under which the cost is lower.*

Definition 44. *The dual problem of optimal transport is given as:*

A plan γ is optimal iff for any $n \in \mathcal{N}$, a permutation of $\sigma of \{1, \cdots, n\}$ and any $\{(x_k, y_k)\}_{k=1\cdots n} \subset support(\gamma)$, then it holds that

$$\sum_k c(x_k, y_k) \le \sum_k c(x_k, y_{\sigma(k)})$$

Kantorvich's formulation of the optimal transport problem guarantees the existence of an optimal transport plan γ. The proof is given in *Thm: 4.1* in [10] which I don't follow.

The dual formulation of the optimal transport problem is given by:

$$\textit{maximize} \int \Psi(x)d\mu(x) - \int \psi(y)d\nu(y) \text{ s.t } \Psi(x) - \psi(y) \le c(x, y)$$

Definition 45. *The Kantorovich distance is defined as*

$$T_K(h)(P,Q) = sup_f(\int fd\mu - \int fd\nu)$$

where f is bounded measurable function satisfying the Lipschitz condition $f(x) - f(y) \leq h(x,y)$

I didn't follow many of the ideas used in proofs in Optimal Transport, so I take them for granted.

3 The Tour

Lemma 3.1 of [2] states that Kantorovich distance between two probability distributions P and Q if 0 if and only if both P and Q have the same distribution over Σ-measurable sets induced by the binary relation R which relates two states $s, s^{'}$ if $h(s, s^{'}) = 0$.

Lemma 1. *Let $h \in \mathcal{M}$. Then $T_K(h)(P,Q) = 0 \Leftrightarrow P(X) = Q(X), \forall X \in \Sigma(\text{Rel}(h))$*

I did not follow the proof because of the approximation argument and argument regarding mentioning P being tight and its implications thereof.

In section 4 of [2], the upper bound α of $\mathcal{M}$ is set to $\alpha = \frac{B}{1-c}$, where B is upper bound of the reward function.

If k is a semi-metric on $S \times S$, $Rel(k)$ is a binary relation induced by k i.e two states s and $s^{'}$ are related if $k = 0$. $Rel(k)$ is in fact an equivalence relation because $k(s,s) = 0 \implies (s,s) \in Rel(k)$, $k(s,s^{'}) \implies k(s^{'},s) = 0$ and $k(s,s^{'}) = 0$ and $k(s^{'},s^{''}) \implies k(s,s^{''}) = 0$ by triangle inequality.

Theorem 4.1: For the purpose of *Theorem* 4.1: *Let $c \in (0,1)$. $F^c : \mathcal{M} \to \mathcal{M}$ is defined by*

$$F^c(s,s^{'}) = max_{a \in A}(|r_s^a - r_{s^{'}}^a| + cT_k(h)(P_s^a, P_{s^{'}}^a))$$

F^c is monotone on $\mathcal{M}$ because $T_k(h)(P_s^a, P_{s^{'}}^a)$ is monotone with respect to h and $|r_s^a - r_{s^{'}}^a|$ does not depend on h. Therefore, because of *Knaester-Tarski* theorem, the least fixed point d_{fix}^c of F^c exists. *Lemma* 4.3 in [10] mentions that $P(f)$ is lower semi-continuous and supremum over lower semi-continuous functions is a lower semi-continuous function and hence $T_k(h)(P_s^a, P_{s^{'}}^a)$ is lower semi-continuous. Therefore, $F^c(h)$ is lower semi-continuous.

$Rel(d_{fix}^c)$ is therefore an equivalence relation and the fact it is the largest equivalence relation corresponding to a fixed point follows from d_{fix}^c being the least fixed point i.e a semi-metric that assigns the least distance between between $(s,s^{'})$ and hence assigns 0 to as many pairs of $(s,s^{'})$ as possible while also being a fixed point of F^c.

To see formally, we have to show that bisimulation is contained in $Rel(d_{fix}^c)$ and $Rel(d_{fix}^c)$ is contained in bisimulation. When $F^c(h) = 0$, $r_s^a = r_{s^{'a}}$ and $T_k(h)(P_s^a, P_{s^{'}}^a) = 0 \implies P(X) = Q(X)\ \forall X$ in $\Sigma(Rel(h))$ by *Lemma 1* i.e for the Σ-measurable equivalence classes induced by $Rel(h)$ both P and Q assigns the same probability. Therefore, $Rel(F^c(h))$ satisfies the definition of bisimulation relation given in *definition* 14 and hence $Rel(F^c(h)) = \mathcal{F}(Rel(h))$. Therefore, $Rel(d^c_{fix}) = \mathcal{F}(Rel(d^c_{fix}))$ and is a fixed point and hence contained in bisimulation. To show that bisimulation is contained in the fixed point, the idea is to show $d^c_{fix} \leq \mathbb{1}_{\not\sim}$. The motivation is that when the distance $\leq \mathbb{1}_{\not\sim}$, at least as many state pairs have 0 distance as $\mathbb{1}_{\not\sim}$ and $Rel(\mathbb{1}_{\not\sim}) = \sim \subset Rel(d^c_{fix})$. Since, F^c is a contraction mapping (as shown in the following theorem), $F^c(\mathbb{1}_{\not\sim}) \leq \mathbb{1}_{\not\sim}$. Hence $d^c_{fix} \leq F^c(\mathbb{1}_{\not\sim}) \leq \mathbb{1}_{\not\sim}$. Thus, bisimilarity is contained in $Rel(d^c_{fix})$. Therefore, $Rel(d^c_{fix})$ is the bisimilarity.

To assert the existence of d^c_{fix}, Banach fixed point theorem is used. Banach fixed point theorem asserts that existence of the fixed point of a function f in a complete metric space X whenever f is a contraction.

Theorem 7. *(Banach fixed point theorem)Let (X, d) be a complete metric space and $T : X \to X$ be a contraction mapping i.e for some $c \in (0, 1)$*

$$d(Tx, Tx^{'}) \leq cd(x, x^{'})$$

$\forall x, x^{'} \in X$. Then,

(i) T has a unique fixed point x^* and

(ii) for any $x_0 \in X$, $d(x^*, T^n(x_0)) \leq \frac{c^n}{1-c} d(T(x_0), x_0)$ i.e $lim_{n\to+\infty} T^n(x_0) = x^*$

Proposition 4.2 : To prove the existence and uniqueness of d^c_{fix}, it is sufficient to show F^c is a contraction mapping. Let $h, h^{'} \in \mathcal{M}$ and $\forall\ s, s^{'} \in S$.

$$\begin{aligned}
F^c(h)(s, s^{'}) - F^c(h^{'})(s, s^{'}) &= max_{a\in A}(|r^a_s - r^a_{s^{'}}| + cT_k(h)(P^a_s, P^a_{s^{'}}))- \\
&\quad max_{a\in A}(|r^a_s - r^a_{s^{'}}| + cT_k(h^{'})(P^a_s, P^a_{s^{'}})) \\
&\leq max_{a\in A}(|r^a_s - r^a_{s^{'}}| + cT_k(h)(P^a_s, P^a_{s^{'}})- \\
&\quad |r^a_s - r^a_{s^{'}}| - cT_k(h^{'})(P^a_s, P^a_{s^{'}})) \\
&= max_{a\in A}(cT_k(h)(P^a_s, P^a_{s^{'}}) - cT_k(h^{'})(P^a_s, P^a_{s^{'}})) \\
&= max_{a\in A}(cT_k(h - h^{'} + h^{'})(P^a_s, P^a_{s^{'}}) - cT_k(h^{'})(P^a_s, P^a_{s^{'}})) \\
&\textit{(In the paper above step has an inequality but I don't see why)} \\
&\leq max_{a\in A}(cT_k(||h - h^{'}|| + h^{'})(P^a_s, P^a_{s^{'}}) - cT_k(h^{'})(P^a_s, P^a_{s^{'}})) \\
&\textit{(since } ||.|| \textit{ is the sup norm and hence } (h - h^{'}) \leq ||h - h^{'}||\forall(s, s^{'})) \\
&\leq max_{a\in A}(c||h - h^{'}|| + cT_k(h^{'})(P^a_s, P^a_{s^{'}}) - cT_k(h^{'})(P^a_s, P^a_{s^{'}})) \\
&= max_{a\in A}(c||h - h^{'}||) \\
&= c||h - h^{'}||
\end{aligned}$$

Therefore, F^c is a contraction mapping. Hence, F^c has a unique fixed point.
Lower Semi-continuity of $\mathbb{1}_{\not\sim}$: Let $\mathbb{1}_{\not\sim}$ be a semimetric as defined in 2.5. To see why this is a lower semi-continuous function, let us prove $L(y)$ *wrt* $\mathbb{1}_{\not\sim}$ is closed $\forall y \in \mathbb{R}$ and then by *Definition* 36 $\mathbb{1}_{\not\sim}$ is lower semi-continuous. y can take values 0 or 1 and when $y = 1$, $L(y) = S \times S$ and hence it is closed. So it is sufficient to prove $L(y)$ is closed when $y = 0$. Since $F^c(h)$ is lower semi-continuous, $L(y)$ corresponding to $y = 0$ is closed. Since $L(y)$ for $y = 0$ for

$\mathbb{1}_{\not\sim} = F^c(h)$ (the bisimilarity), $L(y)$ for $y = 0$ corresponding to $\mathbb{1}_{\not\sim}$ is closed and hence $\mathbb{1}_{\not\sim}$ is lower semi-continuous.

***Claim**:Bisimulation is a closed subset of $S \times S$*
Since, $\mathbb{1}_{\not\sim}$ is lower semi-continuous, $F^c(h)$ is a closed set $\forall\ y \in \mathcal{R}$. Since bisimulation corresponds to $y = 0$, bisimulation is a closed subset of $S \times S$.

Proposition** 4.3**: provides a bound on bisimulation metric when the parameters of the MDP have been perturbed. The algebra is straightforward and in the last step we notice that

$$d_1(x,y) - d_2(x,y) \leq 2max_a(||r_1^a - r_2^a||) + c||d_1 - d_2|| + 2c\frac{B}{1-c}sup_{a,s}d_{TV}(P_{1,s}^a, P_{2,s}^a)$$

holds $\forall x, y \in S$. Therefore, the following inequality also holds

$$||d_1 - d_2|| \leq 2max_a(||r_1^a - r_2^a||) + c||d_1 - d_2|| + 2c\frac{B}{1-c}sup_{a,s}d_{TV}(P_{1,s}^a, P_{2,s}^a)$$

and thus by rearranging the terms, we get the proposition.

Theorem 5.1 ***Value function Bounds:***
To Prove: $|V^*(s) - V^*(s^{'})| \leq d_{fix}^c(s, s^{'})$
Proof:
Let $h \in \mathcal{M}$ and V^0 be initialized to all zeros.
base case(i=1):

$$\begin{aligned}
V^1(s) - V^1(s^{'}) &= max_a(r_s^a + \gamma P_s^a(V^0)) - max_a(r_{s'}^a + \gamma P_{s'}^a(V^0)) \\
&\leq max_a(|r_s^a - r_{s'}^a| + \gamma(P_s^a(V^0) - P_{s'}^a(V^0))) \\
&\leq max_a(|r_s^a - r_{s'}^a| + \gamma sup_f(P_s^a(f) - P_{s'}^a(f))) \quad (f \text{ s.t } f(x) - f(y) \leq h(x,y)) \\
&= max_a(|r_s^a - r_{s'}^a| + \gamma T_k(h)(P_s^a, P_{s'}^a)) \\
&= F^c(h)(s, s')
\end{aligned}$$

Since this hold $\forall s, s^{'}$, we see that $|V^1(s) - V^1(s^{'})| \leq F^c(h)(s, s')$. *induction from i to $i+1$:*
Let us assume that $\forall s, s^{'}$, we see that $|V^i(s) - V^i(s^{'})| \leq (F^c)^i(h)(s, s')$ and prove that $|V^{i+1}(s) - V^{i+1}(s^{'})| \leq (F^c)^{i+1}(h)(s, s')$.

$$\begin{aligned}
V^{i+1}(s) - V^{i+1}(s^{'}) &= max_a(r_s^a + \gamma P_s^a(V^i)) - max_a(r_{s'}^a + \gamma P_{s'}^a(V^i)) \\
&\leq max_a(|r_s^a - r_{s'}^a| + \gamma(P_s^a(V^i) - P_{s'}^a(V^i)))
\end{aligned}$$

(By induction hypothesis $|V^i(s) - V^i(s^{'})| \leq (F^c)^i(h)(s, s')$ and hence)

$$\begin{aligned}
&\leq max_a(|r_s^a - r_{s'}^a| + \gamma sup_f(P_s^a(f) - P_{s'}^a(f))) \quad (f \text{ s.t } f(x) - f(y) \leq (F^c)^i(x,y)) \\
&= max_a(|r_s^a - r_{s'}^a| + \gamma T_k((F^c)^i)(P_s^a, P_{s'}^a)) \\
&= F^c((F^c)^i)(s, s') \\
&= (F^c)^{i+1}(s, s')
\end{aligned}$$

Therefore, $|V^{i+1}(s) - V^{i+1}(s^{'})| \leq (F^c)^{i+1}(s, s')$. Since each V^i is continuous, by taking limit $i \to \infty$, we have $|V^*(s) - V^*(s^{'})| \leq d_{fix}^c(s, s')$. The requirement $\gamma \leq c$ is to ensure that the semimetrics $(F^c)^i$ are bounded by $\frac{B}{1-c}$.

4 Conclusion

We have presented the tools required to understand [2] and explained the arguments presented. The tools from analysis were easy to pickup and form an intuition whereas measure theory and optimal transport were used mechanically. In the future, we would like to work on filling the gaps to learn both these areas. The drawbacks of the *pseudometric* introduced in [2] are the requirement of finite action space and not considering systems that evolve similarly but under different actions. The methods that take into account the equivalence among actions are studied in [**?**] and [**?**] which are possible directions of study for us.

5 Acknowledgements

In addition to the cited materials, we used the following: [8] for Topology and Analysis, [1] and [3] for clarifications on notation and arguments and online resources of [4], [5], [7] and [6].

References

[1] N. Ferns. *Metrics for Markov Decision Processes.* 2003.

[2] N. Ferns, P. Panangaden, and D. Precup. Metrics for markov decision processes with infinite state spaces. In *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence*, UAI'05, pages 201–208, Arlington, Virginia, United States, 2005. AUAI Press.

[3] N. Ferns, P. Panangaden, and D. Precup. Bisimulation metrics for continuous markov decision processes. *SIAM Journal on Computing*, 40(6):1662–1714, 2011.

[4] K. Jagannathan. Measure theory notes (`http://www.ee.iitm.ac.in/~krishnaj/ee5110notes.htm`).

[5] J.-M. F. Liou. Proof of bounded real functions with uniform norm is banach (`http://www.math.ncku.edu.tw/~fjmliou/Complex/spbounded.pdf`).

[6] J. Nachbar. Properties of semi continuous functions (`https://pages.wustl.edu/files/pages/imce/nachbar/semi-continuity.pdf`).

[7] A. Nedich. Properties of semi continuous functions (`http://www.ifp.illinois.edu/~angelia/L4_closedfunc.pdf`).

[8] G. Simmons. *Introduction to Topology and Modern Analyis.* 1963.

[9] T. Tao. *An Introduction to Measure Theory.* 2011.

[10] C. Villani. *Optimal Transport: Old and New.* 2016.