EE 736: Reading Assignment (Paper Review)

Kalpesh Patil (130040019)[1]

[1]Electrical Engineering, IIT Bombay

Abstract

This report broadly contains review of a major paper by Smith and McCardle [2002] titled "Structural Properties of Stochastic Dynamic Programs". The major contribution of this paper is to prove theorems of the form, *If reward function satisfies property P and transition probability matrix satisfies stochastic version of P, value function will also satisfy property P.* Apart from this, some other papers which were studied to elaborate ideas and necessary explanations are also cited as and when required. This report will outline the results proposed in the above mentioned paper and justify intuition behind them based on mathematical framework used by authors. It will also motivate some of the practical applications where this theory will be useful. At the end I would mention research papers which were motivated by this paper in this domain and final concluding remarks.

Introduction

Stochastic dynamic programming is a technique for modelling and solving problems of decision making under uncertainty. Markov decision processes provide a mathematical framework for modeling decision making in situations where outcomes are partly random and partly under the control of a decision maker. They are useful for studying a wide range of optimization problems solved via dynamic programming. They find tremendous applications in various fields like finance, queueing systems, epidemic processes, robotics, population processes, reinforcement learning etc. Therefore it is important to study properties of value functions of a Markov decision process.

Certain properties of value function in Markov decsion processes are of interest. These properties may include monotonicity, convexity supermodularity etc. and their combinations. This is crucial in understanding how the results of the models will change on changing its parameters (sensitivity). The major contribution of this paper is to prove theorems of the form, *If reward function satisfies property P and transition probability matrix satisfies stochastic version of P, value function will also satisfy property P.* Particularly authors have crafted out some meta-theorems that state conditions for value function to satisfy certain properties (few of the properties are mentioned earlier). There are three major results stated in the paper as follows:

- Each of the properties in C3 (Closed Convex Cone) and several other properties can be represented as a system of inequalities having specific form.
- Conditional expectation will satisfy given C3 property if random variables satisfy its stochastic version, which is defined by similar set of inequalities of earlier result (scalar inequalities are replaced by stochastic dominance inequalities).
- Application of these results to Markov reward processes and Markov decision processes. The proposition states, if property P is satisfied by reward function and transition probabilities satisfy stochastic version of it, value function also satisfies that property.

This report will describe prior literature in this domain. Then it would walk through some preliminary definitions and concepts. All three results mentioned above will be explained in subsequent sections. Later it would describe some real life problems of MDP and their solutions obtained using these techniques. Finally some concluding remarks are made.

Prior Literature

The prior literature mentioned in the paper mainly talks about one or two specific properties of dynamic programs. There hasn't been much work done on bringing out connection between general properties of value functions and reward functions & transition probabilities. Monotonicity has been studied by Derman [1963] and Miiller [1997]. Convexity properties have been studied by Hinderer [1982] and Merton [1973]. Submodularity properties were studied by Topkis [2011]. Authors have tried to unite all these properties together and metatheorems for a general class of properties rather specific ones. Athey [1998] has tried to characterize Closed Convex Cone properties containing constants (coined as C5 by authors) and applied results to in static (non-dynamic) problems. Authors have generalized them further for Convex Cone properties (C3 properties) and applied them to stochastic dynamic programs.

Preliminaries

This section will describe definitions of few newer concepts which will be used later. Some of the definitions are directly taken from paper while others have been gathered from various external sources. Consider value function $v_k^*(x_k)$ of an MDP in k^{th} period. Note that we are counting k from backwards i.e. $k = 0$ implies terminal state.

$$v_k^*(x_k) = \underset{a_k \in A_k}{sup} \Big\{ r_k(a_k, x_k) + \delta_k E[v_{k-1}^*(\tilde{x}_{k-1}(a_k, x_k))] \Big\} \qquad for \;\; k > 0$$
$$v_0^*(x_0) = 0$$

Here a_k denotes action selected in k^{th} period (from set A_k). $r_k(a_k, x_k)$ is the reward earned in state k when action a_k is selected. $\tilde{x}_{k-1}(a_k, x_k)$ is the random variable for next period state. The goal of the paper is to find properties of these value functions v_k^* and limiting forms of these functions i.e. $v^*(x) = \underset{k\to\infty}{lim} v_k^*(x_k)$. Proofs can be carried out by induction on k and proving for ground case at $k = 0$. Properties for the limiting form v^* can be established by showing that the set of functions satisfying this property is closed and therefore contains the limiting function.

Definition 1 *A set of functions $\mathcal{F}$ forms a convex cone, if for any $f_1, f_2 \in \mathcal{F}$ and scalars $a, b \geq 0$, $af_1 + bf_2 \in \mathcal{F}$*

Definition 2 *Topology of Pointwise Convergence (Munkres [2000]): Given a point x of the set X and an open set U of the space Y, let $S(x, U) = \{f | f \in Y^X and f(x) \in U\}$. The sets $S(x, U)$ are a subbasis for a topology on Y^X, which is called the topolgoy of pointwise convergence.*

Definition 3 *P is a Closed Convex Cone (C3) property if set of functions satisfying property P form a convex cone in the topology of pointwise convergence.*

Definition 4 *P is Closed Convex Cone Containing Constant (C5) property if it is a C3 property and constant functions satisfy P.*

Definition 5 *Subadditive: A function f is subadditive on Θ, if $f(\theta_1 + \theta_2) < f(\theta_1) + f(\theta_2)$. Here we assume that if $\theta_1 + \theta_2 \in \Theta$ whenever $\theta_1 \in \Theta$ and $\theta_2 \in \Theta$*

Definition 6 *Supermodular: A function f is supermodular if $f(\theta_1) + f(\theta_2) < f(\theta_1 \wedge \theta_2) + f(\theta_1 \vee \theta_2)$ for all $\theta_1, \theta_2 \in \Theta$. Here $\wedge$ and $\vee$ denote componentwise minimization and maximization respectively; Θ is a lattice if $\theta_1 \wedge \theta_2$ and $\theta_1 \vee \theta_2$ are in Θ whenever $\theta_1, \theta_2 \in \Theta$.*

Table 1 shows list of some C5 and C3 properties. Counter examples are provided for few properties which are C5 properties but not C3 properties. Note that intersection of C5 properties is a C5 property and intersection of C3 properties is a C3 property. Therefore if reward function (r) and $E[v_{k-1}^*]$ satisfy property P, then current v_k^* will also satisfy P because the given set of function is a convex cone.

Table 1: C3 and C5 properties

Property	C3	C5	Counter example for C5
Constant	Yes	Yes	N.A.
Nonnegative	Yes	No	$f(\theta) = c$, where c is negative
Increasing	Yes	Yes	N.A.
Convex	Yes	Yes	N.A.
Subadditive	Yes	No	$f(\theta) = c$, where c is a constant
Supermodular	Yes	Yes	N.A.

Table 2: Inequality Representation of C3 Properties

Property	ordering (α)	B_α	Γ_α	Weights (λ)	Details
Increasing	(θ_1, θ_2) s.t. $\theta_1 < \theta_2$	$\theta_\beta = \theta_1$	$\theta_\gamma = \theta_2$	$\lambda_\beta = \lambda_\gamma = 1$	$f(\theta_1) \leq f(\theta_2)$
Convex	$(\phi, \theta_1, \theta_2)$ s.t. $\theta_1, \theta_2 \in \Theta$ and $0 \leq \phi \leq 1$	$\theta_\beta = \phi\theta_1 + (1-\phi)\theta_2$	$\theta_{\gamma^1} = \theta_1$ and $\theta_{\gamma^2} = \theta_2$	$\lambda_\beta = 1, \lambda_{\gamma^1} = \phi, \lambda_{\gamma^2} = 1 - \phi$	$f(\phi\theta_1 + (1-\phi)\theta_2) \leq \phi f(\theta_1) + (1-\phi) f(\theta_2)$
Constant	$\theta^* \in \Theta$ and $\theta \in \Theta$				$f(\theta) \leq f(\theta^{*)}$ and $f(\theta^*) \leq f(\theta$ for all $\theta \in \Theta$
Linear					Can be established by assuming linear as convex as well as concave i.e. union of conditions for convex and concave

Inequality Representation of C3 Properties

The major result of this section is to prove that C3 or C5 properties can be written as comparison between finite weighted sums of function evaluations at specific values of θ.

Proposition 1 *: A property P is a C3 property if and only if there exists a collection, indexed by α in $\mathscr{A}$, of finite sets of points $\{\theta_\beta\}_{\beta \in B_\alpha}$, $\{\theta_\gamma\}_{\gamma \in \Gamma_\alpha}$ and positive weights $\{\lambda_\beta\}_{\beta \in B_\alpha}$, $\{\lambda_\gamma\}_{\gamma \in \Gamma_\alpha}$ that define a test of satisfaction of the form f satisfies P if and only if*

$$\sum_{\beta \in B_\alpha} \lambda_\beta f(\theta_\beta) \leq \sum_{\gamma \in \Gamma_\alpha} \lambda_\gamma f(\theta_\gamma) \qquad \forall \alpha \in \mathscr{A}$$

Furthermore, if P is a C5 property for each α in $\mathscr{A}$, we can normalize the weights to sum to one.

Each inequality is a weighted sum of function evaluation at particular points. Authors have provided proof of this preposition in the paper. Table 2 summarizes various C3 properties and their corresponding inequalities.

Also if a series of functions $\{f_0, f_1, f_2, ...\}$ each satisfying above proposition converges pointwise to some function f, then limiting function will also satisfy above proposition. Note that there exist few properties like boundedness, integrable, differentiable etc. such that the functions satisfying these properties form a convex cone but that may or may not be closed in the topology of pointwise convergence.

Table 3: Properties for conditional expectations

Property (P)	Type of functions U	Condition	Order of dominance
Increasing	Increasing	$\mu(\theta_1) \preceq_U \mu(\theta_2)$ for all $\theta_1 \leq \theta_2$	First
Concave	Increasing	$\phi\mu(\theta_1) + (1-\phi)\mu(\theta_2) \preceq_U \mu(\phi\theta_1 + (1-\phi)\theta_2)$ for all θ_1, θ_2 and $0 \leq \phi \leq 1$	First
Concave	Concave	$\phi\mu(\theta_1) + (1-\phi)\mu(\theta_2) \preceq_U \mu(\phi\theta_1 + (1-\phi)\theta_2)$ for all θ_1, θ_2 and $0 \leq \phi \leq 1$	Second
Super-modular	Increasing	$\mu(\theta_1) + \mu_(\theta_2) \preceq_U \mu(\theta_1 \wedge \theta_2) + \mu(\theta_1 \vee \theta_2)$ for all θ_1, θ_2.	First

Properties of Conditional Expectation

The major result for this section is to find necessary and sufficient conditions for $E[u(\tilde{x}(\theta))]$ to satisfy property P in θ for all functions u in set U.

Definition 7 *Stochastic Dominance: $\tilde{x}(\theta_1)$ dominates $\tilde{x}(\theta_2)$ on U if $E[u(\tilde{x}(\theta_1))] > E[u(\tilde{x}(\theta_2))]$ for all $u \in U$. It is denoted by $\tilde{x}(\theta_1) \succeq_U \tilde{x}(\theta_2)$.*

The above definition generalizes notion of dominance. According to first order dominance A dominating B means that $F_A(x) \leq F_B(x)$ for all x, with strict inequality at some x, where F_A and F_B are respective CDFs of A and B. According to second order dominance A dominating B means $\int_{-\infty}^{x} [F_B(t) - F_A(t)]\, dt \geq 0$ with strict inequality at some x. Roughly speaking second order dominance says A involves less risk and has at least as high a mean as B. (Wikipedia). Authors have proved this inequality result for stochastic dominance for various U (type of functions) like *increasing*, *concave*, *increasing and concave* and *supermodular*.

For conditional expectations $E[u(x(\theta))]$ to satisfy property P in θ for u in set of functions U, it requires to show that the family of measures $\mu(\theta)$ defining $\tilde{x}(\theta)$ to satisfy the system of inequalities defining property P. The inequalities now will be in the form of dominance relation ($\succeq_U$) for the given set of functions U rather than simple arithmetic inequalities.

Proposition 2 *: Let P be a C3 property represented as in Proposition 1. The $E[\tilde{x}(\theta)]$ satisfies P on Θ for all u in U if and only if measures $\mu(\theta)$ satisfy*

$$\sum_{\beta \in B_\alpha} \lambda_\beta \mu(\theta_\beta) \preceq_U \sum_{\gamma \in \Gamma_\alpha} \lambda_\gamma \mu(\theta_\gamma) \qquad \forall \alpha \in \mathscr{A}$$

Proposition 2 is similar to Proposition 1 except the inequality is replaced by stochastic dominance relation and f is replaced by probability measure μ. A similar table (table 3) as defined earlier can be crafted out for this relation as well. The table can be read as, $E[u(\theta)]$ satisfies property P for all u in type of functions U, if certain conditions are satisfied and $\tilde{x}(\theta)$ is satisfying P in the sense of first/second order of dominance.

Characterization of Value Functions for Markov Reward Processes

Consider simplified version of markov decision process, a markov reward process as follows:

$$v_k(x_k) = r_k(x_k) + \delta_k E[v_{K-1}(\tilde{x}_{k-1}(x_k)] \qquad for \;\; k > 0$$
$$v_0(x_0) = 0$$

Let P be a C3 property. If $r_k(x_k)$ satisfies P for all k and we can ensure that $E[v_{k-1}(\tilde{x}_{k-1}(x_k)]$ also satisfies P, whenever $v_{k-1}(x_{k-1})$ satisfies P, then we can show that v_k also satisfies P using induction. Since set of functions satisfying P is closed under pointwise limits, $\underset{k\to\infty}{lim}\, v_k$, if exists, also satisfies P.

Proposition 3 *: Let U be a set of functions on X satisfying a C3 property P. If for all k,*
(a) the reward functions $r_k(x_k)$ satisfy P and
(b) the transitions $x_{k_1}(x_k)$ satisfy $P(\succeq_U)$
then each v_k satisfies P and $\lim_{k\to\infty} v_k$ if exists, also satisfies P.

Note that, due to recursive nature of this formulation, unlike previous section, here we must have property P same as property of the class of functions U.

Characterization of Value Functions for Markov Decision Processes

As described earlier markov decision process is defined as follows;

$$v_k^*(x_k) = \sup_{a_k \in A_k} \left\{ r_k(a_k, x_k) + \delta_k E[v_{k-1}^*(x_{k-1}(\tilde{a_k}, x_k))] \right\} \qquad for \;\; k > 0$$
$$v_0^*(x_0) = 0$$

Before understanding properties of a full stochastic dynamical programs, we must have to study about which properties are preserved under maximization. Let $g(\theta) = \sup_{a \in \mathcal{A}}\{f(a,\theta)\}$. The general intuition that if a property P is satisfied by $f(a,\theta)$ for all a, it will also be satisfied for optimal a^*, is not always true. But if the underlined property is a *single-point property*, then it is true. A single-point property will have single term on LHS of the inequality of Proposition 1. i.e. $g(\theta)$ satisfies P if and only if $g(\theta_\alpha) \leq \sum_{\gamma \in \Gamma_\alpha} \lambda_\gamma g(\theta_\gamma)$. C3 properties like increasing, decreasing, convex, and subadditive are single-point properties, while concave isn't a single point property. To generalize this notion further we consider *joint extension* of P defined as follows. Given a C3 property P on Θ represented as in proposition 2, we say that a C3 property P^* on $\mathcal{A} \times \Theta$ is a joint extension of P, if for any α in $\mathscr{A}$ in proposition 2 and any set of actions $\{a_\beta\}_{\beta \in B_\alpha}$ there exists a set of actions $\{a_\gamma\}_{\gamma \in \Gamma_\alpha}$ such that $\sum_{\beta \in B_\alpha} \lambda_\beta f(a_\beta, \theta_\beta) \leq \sum_{\gamma \in \Gamma_\alpha} \lambda_\gamma f(a_\gamma, \theta_\gamma)$. Roughly speaking above expression can be thought as an extension of a given property in θ domain to $a \times \theta$ domain. Then author have proved that if $f(a,\theta)$ satisfies P^*, then $g(\theta)$ satisfies P. This is an extremely useful result in the sense that it allows transfer of a property through maximization, which is very crucial in MDP. This result holds true for single point properties, joint concavity (joint in a and θ) and joint supermodularity. It doesn't hold true for joint submodularity.

Having developed theory for property conservation in maximization, it can be applied to MDP. Let function to be maximized be $f(a_k, x_k) = r_k(a_k, x_k) + \delta E[v_{k-1}^* \tilde{x}_{k-1}(a_k, x_k)]$. Combining above result with proposition 3, we obtain

Proposition 4 *: Let U be a set of functions on X satisfying a C3 property P. and let P^* be joint extension of P on $A \times \Theta$. If for all k,*
(a) the reward functions $r_k(x_k)$ satisfy P^* and
(b) the transitions $\tilde{x}_{k_1}(x_k)$ satisfy $P^*(\succeq_U)$
then each v_k^* satisfies P and $\lim_{k\to\infty} v_k^*$ if exists, also satisfies P.

This is the most important result from the paper and allows one to talk about various structural properties of MDP. Few common C3 properties like increasing, increasing & convex, increasing & concave (jointly concave) satisfy this proposition.

Applications

Propositions made in previous two sections were used to solve some well-known problems. This sections will briefly describe each of the problem statement and obtained results.

1. Copper Mine Model

A firm operating a copper mine has to decide whether to operate mine, close temporarily or abandon altogether every period. There's cost associated with each of these options and also firm earns certain revenue depending upon price of copper when mine is open. Let $x_k = (s_k, p_k)$ be the state where s_k is the status of the mine and p_k be log(price of copper) in k^{th} period. $\tilde{p}_{k-1}$ be random variable associated with price in next period, which is assumed to be distributed normally with mean $p_k + \mu$ and standard deviation σ. Update for the value function in MDP is given as follows:

$$\begin{aligned} v_k^*(s_k, p_k) = max\big\{ & -c(abandoned, s_k) + \delta E[v_{k-1}^*(abandoned, \tilde{p}_{k-1}(p_k))], \\ & -c(closed, s_k) + \delta E[v_{k-1}^*(closed, \tilde{p}_{k-1}(p_k))] \\ & -c(open, s_k) + \gamma exp(p_k) + \delta E[v_{k-1}^*(open, \tilde{p}_{k-1}(p_k))]\big\} \end{aligned}$$

Results:

1. Value function is increasing and convex in p_k for each s_k

2. increasing σ (which implies increase in 'risk') also increases value function. This results seems intuitively true as well, because one expects higher reward when underlined risk is higher.

2. Technology Adaptation

A firm wants to decide if it should adapt a new technology whose value is uncertain and denoted by μ. In period k it normally distributed with mean m_k and precision (inverse of variance) s_k. Firm can reject technology (gains zero), adapt technology (gains $Am_k - K$) or investigate further by paying cost $c > 0$. Investigation implies normally distributed noisy observation of the value of technology with mean μ and precision t is obtained. This will allow firm to update its estimate for m_k. Next state estimate (m_{k-1}) is normally distribute with mean m_k and precision $\frac{s_k(s_k+t)}{t}$. Next state precision is $s_k + t$. These are calculated using simple Bayes' rule. We assume $x_k = (m_k, s_k)$ be the state of the system. Therefore update for the value function in the given MDP is given as follows:

$$v_k^*(m_k, s_k) = max\big\{0, Am_k - K, -c + E[v_{k-1}^*(\tilde{m}_{k-1}^*(m_k, s_k), s_k + t)]\big\}$$

Results:

1. Value function is increasing and convex in m_k

2. Value function is decreasing in s_k, which implies less value if prior precision is more (uncertainty is less).

3. Value function is increasing in t, which implies more value if observations are more precise. Note that this contradicts Lippman and McCardle [1987], which stated that value function is decreasing in t. The small error was made in calculation of precision of mean of the next state as pointed out by authors.

4. Mixing property: It captures the idea that earlier resolution of uncertainty about the value of the technology is preferred to later. Mathematically, $0.5v_k^*(m_k + \delta, s_k + \Delta) + 0.5v_k^*(m_k - \delta, s_k + \Delta) \geq v_k^*(m_k, s_k)$ for any $\Delta \geq 0, s_k \geq 0, \delta^2 = \frac{\Delta}{s_k(s_k+\Delta)}$ and m_k.

3. Stochastic Inventory Model

Assume n dimensional vector representing quantity of n products. Let x_k be quantities present in inventory in k^{th} period, a_k be quantities ordered and $\tilde{z}_k$ be random demand in k^{th} period. Therefore next period state is $\tilde{x}_{k-1} = x_k + a_k - \tilde{z}_k$ (note that it can be negative as well, which implies unfulfilled demand is carried forward). Let $c(a_k)$ be cost of ordering a_k, $l(x_{k-1})$ be convex loss function of combined cost of holding unsold inventory and penalty of unmet demands. Therefore value function is given as follows,

$$v_k^*(x_k) = \min_{a_k \geq 0} \{c(a_k) + E[l(\tilde{x}_{k-1}(a_k, x_k)) + \delta v_{k-1}^*(\tilde{x}_{k-1}(a_k, x_k))]\}$$

Results:
It was shown that v_k^* is jointly convex in the vector of inventory levels x_k.

Future Work

I have briefly gone through some other papers which were published in future and established on top of the work done in this paper. These properties have been widely used in the areas like communication, inventory models, pricing models etc. Some of the notable works are listed below.

- Djonin and Krishnamurthy [2005] have addressed the problem of finding optimal transmission adaptation policies for a single user data stream, their structure and related optimal costs. In this they have used result established here about the properties of value function that are preserved with the iterations of the dynamic programming.
- Papadaki and Powell [2007] have studied monotonicity for Multiproduct Batch Dispatch problem. MBD is a problem of multiple (N) types of products manufactured at the suppliers side and waiting to be dispatched in batches to the retailer by a vehicle with finite capacity K. The holding cost of products differs across product classes. Given the waiting queue of products, a decision is taken about whether to dispatch the vehicle. Further if the vehicle is dispatched, a decision is taken on the distribution of product types to be dispatched.
- Djonin and Krishnamurthy [2007] have used properties of stochastic dominance of random variables described here while studying structural results on the optimal transmission policies.
- Popescu and Wu [2007] have studied Dynamic Pricing Strategies with Reference Effects. They have considered dynamic pricing problem of a monopolist firm in a market with repeated interactions, where demand is sensitive to the firms pricing history. While proving Monotone Price Policy lemma in the paper,they have particularly used the result established here about the covexity of value function of MDP.
- Goumagias et al. [2012] describe a dynamic, Markov-based decision support model, aimed at predicting the behavior of a risk-neutral enterprise in Greece and at evaluating tax policies before they are implemented. They have used results related to closed convex set for predicting long-term rewards.

Concluding Remarks

Authors have developed a set of metatheorems that describe how properties of value functions are preserved and propagated through Markov reward and decision processes. The metatheorems provide specific conditions that can be checked in given applications to establish specific properties of the value functions. Main result of the paper can be stated as, if reward functions satisfy certain property P and transition probabilities satisfy stochastic version of P, then value function satisfies P. For Markov reward processes this result can be directly applied for C3 properties, while for Markov decision processes we need additional condition of being closed under maximization in choice of actions for that property.

These results are very crucial in characterizing structural properties of stochastic dynamic programs. Properties like convexity, monotonicity or concavity are useful for ensuring convergence of dynamic programs generally used in MDPs. The applications of such results can be found in various domains like queuing systems, communication, economics, game theory etc.

References

S. Athey. Characterizing properties of stochastic objective functions. 1998.

C. Derman. On optimal replacement rules when changes of state are markovian. *Mathematical optimization techniques*, 396, 1963.

D. V. Djonin and V. Krishnamurthy. Structural results on the optimal transmission scheduling policies and costs for correlated sources and channels. In *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on*, pages 3231–3236. IEEE, 2005.

D. V. Djonin and V. Krishnamurthy. Structural results on optimal transmission scheduling over dynamical fading channels: A constrained markov decision process approach. In *Wireless Communications*, pages 75–98. Springer, 2007.

N. Goumagias, D. Hristu-Varsakelis, and A. Saraidaris. A decision support model for tax revenue collection in greece. *Decision Support Systems*, 53(1):76–96, 2012.

K. Hinderer. *On the structure of solutions of stochastic dynamic programs*. Univ., Fak. für Mathematik, 1982.

S. A. Lippman and K. F. McCardle. Does cheaper, faster, or better imply sooner in the timing of innovation decisions? *Management Science*, 33(8):1058–1064, 1987.

R. C. Merton. Theory of rational option pricing. *The Bell Journal of economics and management science*, pages 141–183, 1973.

A. Miiller. How does the value function of a markov decision process depend on the transition probability? *Math. Oper. Res*, 22:872–885, 1997.

J. R. Munkres. *Topology*. Prentice Hall, 2000.

K. Papadaki and W. B. Powell. Monotonicity in multidimensional markov decision processes for the batch dispatch problem. *Operations research letters*, 35(2):267–272, 2007.

I. Popescu and Y. Wu. Dynamic pricing strategies with reference effects. *Operations Research*, 55(3):413–429, 2007.

J. E. Smith and K. F. McCardle. Structural properties of stochastic dynamic programs. *Operations Research*, 50(5):796–809, 2002.

D. M. Topkis. *Supermodularity and complementarity*. Princeton university press, 2011.