Dynamic Programming

VOLUME II: GENERAL STATES

Thomas J. Sargent and John Stachurski

April 8, 2025

Contents

Preface **v**

Common Symbols **vii**

1 Abstract Dynamic Programs **1**

1.1 Prelude: Optimal Savings 2

1.1.1 Policies and Decisions 2

1.1.2 Optimality 7

1.2 Introducing ADPs 11

1.2.1 Definition and Properties 11

1.2.2 Algorithms 20

1.2.3 Optimality: Sufficient Conditions for Well-Posed ADPs 23

1.2.4 Example: Back to Optimal Savings 26

1.3 ADPs in Topological Space 27

1.3.1 ADPs in Partially Ordered Space 27

1.3.2 Example: Optimal Savings with Markov Income 30

1.3.3 ADPs in Banach Lattices 31

1.4 Additional Topics 35

1.4.1 Nonstationary Policies 36

1.4.2 Minimization Problems 38

1.5 Chapter Notes 41

2 Classes of Dynamic Programs **43**
2.1 Markov Decision Processes 43
2.1.1 Finite State MDPs 43
2.1.2 Finite MDPs in Continuous Time 47
2.1.3 Weak Feller MDPs 50
2.1.4 Strong Feller MDPs 54
2.2 Recursive Decision Processes 56
2.2.1 Definition and Examples 56
2.2.2 RDPs vs ADPs 62
2.2.3 Existence of Greedy Policies 65
2.2.4 Weighted Contractions 65
2.2.5 Properties of Solutions 70
2.3 Beyond RDPs 73
2.3.1 Q-Factors 73
2.3.2 Structural Estimation 75
2.3.3 LQ Control 80
2.4 Chapter Notes 86

3 Transformations **87**
3.1 Isomorphic Dynamic Programs 87
3.1.1 Isomorphisms and Anti-Isomorphisms 87
3.1.2 Applications 97
3.2 Factored Dynamic Programs 101
3.2.1 Definition and Properties 101
3.2.2 Optimality 106
3.2.3 Applications 109

4 Discrete Choice **114**
4.1 Optimal Stopping . . . 114
4.1.1 Job Search . . . 114
4.1.2 Rearranging the Bellman Equation . . . 118
4.1.3 Job Search with Correlated Wage Draws . . . 124
4.1.4 Firm Exit . . . 131
4.2 Extensions . . . 133
4.2.1 Nonlinear Discounting . . . 133
4.2.2 Nonlinear Expectations . . . 135
4.2.3 Job Search with Learning . . . 137
4.2.4 Renewal Problems . . . 142
4.3 Chapter Notes . . . 147

Appendices **148**

Appendix A Mathematical Background **149**
A.1 Foundations . . . 149
A.1.1 Properties of the Real Line . . . 149
A.1.2 Partial Orders . . . 152
A.1.3 Metric Space . . . 161
A.2 Topology . . . 166
A.2.1 Topological Space . . . 167
A.2.2 Stability and Contractions . . . 172
A.3 Measure and Integration . . . 174
A.3.1 Measure Theory . . . 176
A.3.2 Integration . . . 181
A.3.3 Product Spaces . . . 186
A.3.4 Conditioning . . . 187
A.4 Vector Spaces and Norms . . . 188

A.4.1 Vector Space 189
A.4.2 Normed Vector Space 193
A.4.3 Bounded Linear Operators 199
A.5 Order 201
A.5.1 Order Continuity and Order Completeness 202
A.5.2 Ordered Vector Space 206
A.5.3 Topology and Order 213
A.5.4 Markov Models 219
A.5.5 Orders over Distributions 224
A.6 Chapter Notes 227

Appendix B Solutions **228**

Preface

This book is the second of our two-volume sequence on theory and applications of dynamic programming. While Volume 1 (Sargent and Stachurski, 2025a) focused on models with finite state and action spaces, this volume treats general state space and action spaces. This extension is valuable for two reasons. One is that many models are more easily expressed as optimization problems over infinite state and action spaces. The second reason is that working with infinite states and choices provides access to new tools, mainly involving calculus and gradients.

Following on from Chapters 8 and 9 of Sargent and Stachurski (2025a), we work within an abstract setting that builds on the framework in Bertsekas (2022). This setting admits standard dynamic programming problems discussed in Bellman (1957), Stokey and Lucas (1989), Rust (1996), Puterman (2005), and Bertsekas (2012), as well as recursive preference models, nonlinear discounting, robust control problems, and other preference specifications adopted within economics, finance, operations research, and artificial intelligence in recent years.

We recommend that most readers begin with Volume 1, which illustrates the key ideas from dynamic programming in a relatively simple environment and introduces a large amount of notation and terminology. In what follows we assume that readers have some familiarity with notation and core concepts from Volume I.

One point to note is that the mathematical prerequisites for this book are significantly higher than those for Volume 1. Readers will need to be at least somewhat familiar with key concepts from functional analysis, measure theory, and order theory. For convenience, we have provided a relatively extensive appendix that reviews these topics. Some readers will benefit from reviewing the appendix before starting to read the main content.

An additional point worth mentioning is that while Volume I starts with specific models and gradually builds towards general theory – an approach would believe will be relatively accessible for readers attempting to learn the core concepts of dynamic programming for the first time – this volume takes a more mathematical approach, beginning with general theory and then specializing to particular applications. This path will suit readers with a relatively high degree of mathematical maturity.

We emphasize that our work on dynamic programming has been deeply influenced by the elegant and insightful book *Abstract Dynamic Programming* by Dimitri Bertsekas ((Bertsekas, 2022)), now in its third edition. We were fascinated by the first volume and, while our research papers and the content of this volume now appear rather different to his work, the entire framework is inspired by Bertsekas' monograph.

Many friends, students and colleagues have helped with preparation of this book, either by directly reading and commenting on the book or by research collaboration. Extra thanks are due to Shu Hu, Yuchao Li, Nisha Peng, Longye Tian, Jingni Yang, Ziyue (Humphrey) Yang, Junnan Zhang, and Sylvia Zhao. Jingni Yang proved a number of valuable results used throughout the book.

Common Symbols and Terminology

Mathematical Notation

$\mathbb{1}\{P\}$	indicator function, (1 if P is true and 0 otherwise)
$\alpha := 1$	α is defined as equal to 1
$f \equiv 1$	function f is everywhere equal to 1
$\bigwedge$ and $\bigvee$	supremum and infimum (see §A.1.2.4)
$\wp(A)$	the power set of A; that is, the set of all subsets of given set A
$[n]$	$\{1, \ldots, n\}$
$\mathbb{C}$	the complex numbers
$\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{R}$	the natural numbers, integers and real numbers respectively
$\mathbb{Z}_+$, $\mathbb{R}_+$, etc.	the nonnegative elements of $\mathbb{Z}$, $\mathbb{R}$, etc.
$\|x\|$ for $x \in \mathbb{R}$	the absolute value of x
$\|\lambda\|$ for $\lambda \in \mathbb{C}$	the modulus of λ
$\|B\|$ for set B	the cardinality of B
$\mathbb{R}^n$	all n-tuples of real numbers
$\mathbb{R}^{m \times n}$	all $m \times n$ real matrices
$x \leqslant y \quad (x, y \in \mathbb{R}^n)$	$x_i \leqslant y_i$ for $i = 1, \ldots n$ (pointwise partial order)
$\mathbb{R}^{\mathsf{X}}$	all functions from X to $\mathbb{R}$
$m\mathbb{R}^{\mathsf{X}}$	the set of real-valued Borel measurable functions on $(\mathsf{X}, \mathscr{A})$
$b\mathsf{X}$	the set of bounded functions from X to $\mathbb{R}$
$bm\mathsf{X}$	the set of Borel measurable functions in $b\mathsf{X}$
$bc\mathsf{X}$	the set of continuous functions in $b\mathsf{X}$
$ibc\mathsf{X}$	the set of increasing functions in $bc\mathsf{X}$

$\mathscr{D}(\mathsf{X})$	the set of distributions (Borel probability measures) on X
$\mathcal{L}(U, V)$	the set of bounded linear operators from U to V
$\langle a, b \rangle$	the inner product of a and b
$v_n \uparrow v$	(v_n) is increasing and $\bigvee_n v = v$ (see §A.1.2.6)
IID	independent and identically distributed
$X \stackrel{d}{=} Y$	X and Y have the same distribution
$X \sim F$	X has distribution F
$F \preceq_{\mathrm{F}} G$	F first order stochastically dominates G

Dynamic Programming Notation and Terminology

$(V, \mathbb{T})$	an ADP with value space V and policy operators $T_\sigma \in \mathbb{T}$
v_σ	a σ-value function (fixed point of T_σ)
T	the Belman operator, defined by $Tv = \bigvee_\sigma T_\sigma v$
H	the Howard operator, defined by $Hv = v_\sigma$ where σ is v-greedy
W	the optimistic policy operator (see (1.19))
V_G	all $v \in V$ with at least one v-greedy policy
V_U	all $v \in V$ such that $v \preceq Tv$
V_Σ	the set of fixed points of the policy operators
v^*	the value function (greatest element of V_Σ)
VFI	value function iteration
OPI	optimistic policy iteration
HPI	Howard policy iteration

Chapter 1

Abstract Dynamic Programs

Dynamic programming is a technique for solving optimization problems using recursive methods (see, e.g., Bellman (1957) or Bertsekas (2012)). While initially developed for intertemporal problems (inventory management, investment planning, optimal savings and consumption, etc.), it has since been realized that the method can also be applied to a great variety of atemporal problems, ranging from genome sequencing and matrix multiplication to the structure of production chains. Many recent applications of dynamic programming are connected to machine learning and artificial intelligence.

One challenge for any modern theory of dynamic programming is the sheer breadth of applications currently being tackled by this class of techniques. As well as the vast range of concrete problems faced in applied settings, researchers are constantly creating new dynamic programs from existing ones by injecting novel features. These features include time-varying discount rates, nonlinear discounting, risk-sensitive control, ambiguity aversion, nonlinear time aggregation, and so on. Such novel features allow researchers to pose more realistic frameworks and bring previously studied models closer to the data.

Handling this ever-expanding range of applications from a single theoretical framework requires considerable abstraction. This requirement has precipitated the development of *abstract dynamic programming* (Bertsekas, 2022; Sargent and Stachurski, 2025b). Volume 1 of this book, published as (Sargent and Stachurski, 2025a), covers abstract dynamic programming under the assumption that states and actions are finite (see, in particular, Chapters 8 and 9). The present volume drops these assumptions and studies the general case.

To set the scene, we begin by examining an optimal savings problem that is a building block for many economic models. It features a basic intertemporal trade-off from consuming now or later. This trade-off can be solved by dynamic programming. After reminded ourselves of key concepts in the context of optimal savings, we pull back the lense, introducing "abstract dynamic programs" and constructing a theory around them. The optimal savings model is one of many special cases.

1.1 Prelude: Optimal Savings

This section presents an optimal savings model. While some proofs are omitted, please rest assured that we will prove far more general results very shortly.

1.1.1 Policies and Decisions

In an optimal savings problem (sometimes called an "income fluctuation problem"), a household seeks to maximize

$$\mathbb{E} \sum_{t=0}^{\infty} \beta^t u(C_t) \quad \text{s.t.} \quad W_{t+1} = R(W_t - C_t) + Y_{t+1} \quad \text{and} \quad 0 \leqslant C_t \leqslant W_t. \tag{1.1}$$

The constraints in (1.1) are required to hold for all $t \geqslant 0$, and an initial condition w_0 is taken as given. The **utility function** function $u \colon \mathbb{R}_+ \to \mathbb{R}$ maps current consumption C_t into a utility value (loosely speaking, a measure of satisfaction), $\beta \in (0, 1)$ is a **discount factor** indicating impatience, and $R > 0$ is a gross rate of return on assets. The variable W_t represents wealth at time t, while Y_t is labor income. To keep the model simple, we assume (Y_t) is IID with common distribution $\varphi \in \mathscr{D}(\mathbb{R}_+)$, the set of distributions (i.e., Borel probability measures) on $\mathbb{R}_+$.

(We study more general settings later.)

The variable W_t is the **state** of the dynamic program, while C_t is the **action** or **control**. In general, the space in which the state (resp., action) takes values is called the **state space** (resp., **action space**). For this optimal savings model, state and action spaces both equal $\mathbb{R}_+$.

Figure 1.1 shows the timing for the optimal savings problem. After observing W_t, the household chooses C_t and hence savings $W_t - C_t$. Then labor income Y_{t+1} is realized and the state updates to W_{t+1}. The process then repeats.

In maximization problem (1.1) there is another constraint: C_t can depend only on information available at time t. Formally, current consumption C_t must be a (deterministic) Borel measurable function of shocks, states, and actions observed up to and

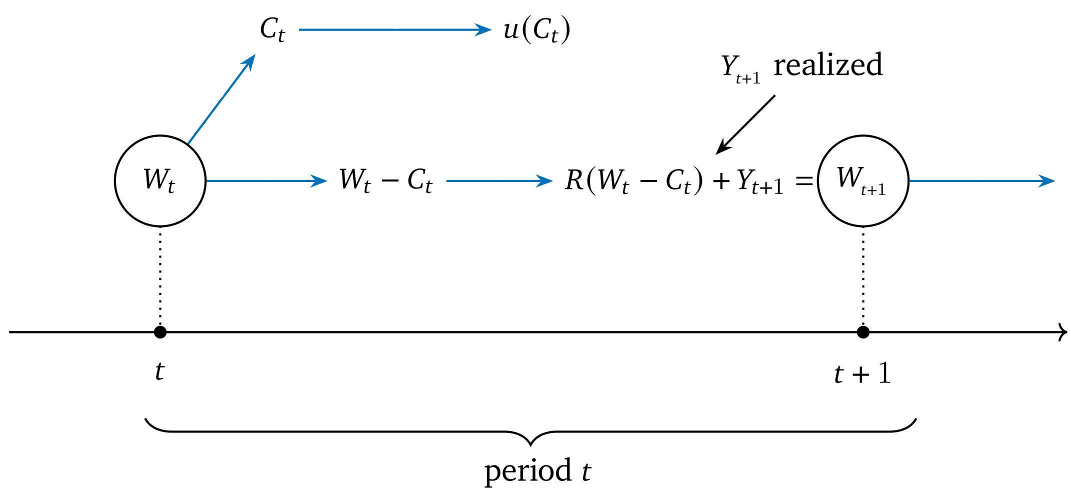


Figure 1.1: Timing for the optimal savings problem

including time t. Thus, the current action cannot depend on future values such as Y_{t+1} or W_{t+1}. A mapping from the history of the state and the shocks into current action is called a **policy function**.[1]

The infinite horizon, IID (Y_t)-process, time-invariant structure of the optimal savings problem lets us focus on policies that make current consumption C_t be a deterministic function σ of the current state W_t. (We will prove this later and discover how it depends on the IID-nature of the (Y_t) process.)

We impose the following simplifying conditions:

Assumption 1.1.1. The function u is continuous and bounded on $\mathbb{R}_+$ and the distribution of labor income can be represented by a continuous density φ on $\mathbb{R}_+$.

In a slight abuse of notation, we use φ to represent the density of labor income as well as the corresponding distribution (i.e., Borel probability measure on $\mathbb{R}_+$). Thus, in the integrals below, $\varphi(\mathrm{d}y)$ and $\varphi(y)\,\mathrm{d}y$ have the same meaning.

1.1.1.1 Lifetime Value

For the remainder of this section, a **stationary Markov policy** is a Borel measurable map σ from $\mathbb{R}_+$ to itself. We often refer to stationary Markov policies more simply

[1] In engineering it is sometimes called a *closed loop control* to emphasize that the control must be a measurable function of an observed history and *not* depend on as yet unrealized random variables.

as **policies**. We call a policy σ **feasible** if $0 \leqslant \sigma(w) \leqslant w$ for all $w \in \mathbb{R}_+$, so that the consumption response $c = \sigma(w)$ obeys the inequalities in (1.1). Let Σ denote the set of all feasible policies. We seek $\sigma \in \Sigma$ that maximizes expected lifetime value. For given σ and initial condition $w = w_0$, expected lifetime value is

$$v_\sigma(w) = \mathbb{E} \sum_{t \geqslant 0} \beta^t u(\sigma(W_t)) \quad \text{when } \; W_{t+1} = R(W_t - \sigma(W_t)) + Y_{t+1} \tag{1.2}$$

for all $t \geqslant 0$ and $(W_t)_{t \geqslant 0}$ starts at w. Below, we refer to v_σ as the **σ-value function**.

It is helpful to express v_σ as the fixed point of an operator. To this end we introduce the **policy operator** corresponding to $\sigma \in \Sigma$ via

$$(T_\sigma v)(w) = u(\sigma(w)) + \beta \int v(R(w - \sigma(w)) + y)\varphi(\mathrm{d}y) \qquad (w \in \mathbb{R}_+). \tag{1.3}$$

We assume that v is an element of

$$V := bm\mathbb{R}_+ := \text{all bounded Borel measurable functions from } \mathbb{R}_+ \text{ to } \mathbb{R}.$$

We recall that V is a Banach space (see §A.4.2) with supremum norm $\|v\| := \sup_x |v(x)|$.

EXERCISE 1.1.1. Show that $T_\sigma V \subset V$ when Assumption 1.1.1 holds.

Policy operators are useful because $v \in V$ is a fixed point of T_σ if and only if it equals the σ-value function. Thus, the fixed point of T_σ characterizes the lifetime value of σ. This is a consequence of the following lemma.

Lemma 1.1.1. *If Assumption 1.1.1 holds, then every policy operator T_σ is globally stable on V. Moreover, the unique fixed point of T_σ in V is the function v_σ defined in (1.2).*

Proof. Fix $\sigma \in \Sigma$ and set $r_\sigma := u \circ \sigma$. Let P_σ be the Markov operator (see §A.5.4.2) defined at $v \in V$ by

$$(P_\sigma v)(w) := \int v(R(w - \sigma(w)) + y)\varphi(\mathrm{d}y) \qquad (w \in \mathbb{R}_+).$$

Using this notation, we can write

$$T_\sigma v = r_\sigma + \beta P_\sigma v. \tag{1.4}$$

In §A.5.4 and Corollary A.4.11 we show that P_σ is a bounded linear operator from V to itself and, using $\beta \in (0, 1)$, that T_σ is globally stable on V with unique fixed point

$v_\sigma \in V$ obeying

$$v_\sigma = (I - \beta P_\sigma)^{-1} r_\sigma = \sum_{t \geqslant 0} (\beta P_\sigma)^t \, r_\sigma. \tag{1.5}$$

(Here I is the identity map on V and the second equality follows from the Neumann series lemma.) It remains only to show that v_σ in (1.5) agrees with v_σ defined in (1.2). To obtain this we use the fact that, when $W_{t+1} = R(W_t - \sigma(W_t)) + Y_{t+1}$ for all t and $W_0 = w$,

$$\left(P_\sigma^t \, r_\sigma\right)(w) = \mathbb{E}\left[r_\sigma(W_t) \mid W_0 = w\right] = \mathbb{E}\left[u(\sigma(W_t)) \mid W_0 = w\right]. \tag{1.6}$$

(The first equality also uses results in §A.5.4.) Combining this with the last expression in (1.5), we see that v_σ in (1.5) and (1.2) are identical. □

Incidently, one can alternatively use the law of iterated expectations to prove that the σ-value function v_σ is a fixed point of T_σ. To do so, one first writes

$$v_\sigma(w) = u(\sigma(w)) + \mathbb{E} \sum_{t \geqslant 1} \beta^t u(\sigma(W_t)).$$

Letting $\mathbb{E}_1$ be a conditional expectation conditional on W_1, applying the law of iterated expectations implies

$$v_\sigma(w) = u(\sigma(w)) + \beta \mathbb{E}\left[\mathbb{E}_1 \sum_{t \geqslant 1} \beta^{t-1} u(\sigma(W_t))\right] = u(\sigma(w)) + \beta \mathbb{E}\, v_\sigma(W_1).$$

Expanding the last expression yields

$$v_\sigma(w) = u(\sigma(w)) + \beta \int v_\sigma(R(w - \sigma(w)) + y) \varphi(\mathrm{d}y). \tag{1.7}$$

Thus, v_σ is a fixed point of T_σ.

1.1.1.2 Lifetime Values as Limits

In the previous section we learned that fixed points of policy operators represent lifetime value. What do finite iterates of policy operators represent? Fixing σ and inspecting the definition of T_σ (see (1.3)) indicates that $(T_\sigma \, v)(w)$ represents the reward received from using policy σ for one period, when w is initial wealth and the function v is used to evaluate the reward from wealth in the second period.

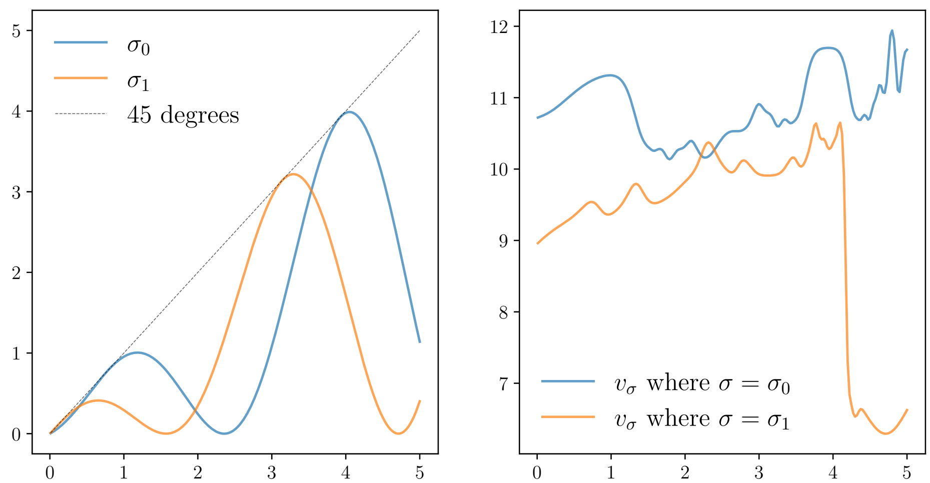


Figure 1.2: Randomly chosen policies and their lifetime values

We can lengthen the horizon by iterating with T_σ while keeping the terminal value function v fixed. Choosing $k \in \mathbb{N}$ and using the expression for T_σ in (1.4), we get

$$T_\sigma^k v = r_\sigma + \beta P_\sigma r_\sigma + \cdots + (\beta P_\sigma)^{k-1} r_\sigma + (\beta P_\sigma)^k v \tag{1.8}$$

The expression on the right is the value of following policy σ for k periods and then receiving a reward for terminal wealth determined by the function v. In other words, it is the finite horizon value of following σ under this terminal condition.

It seems plausible that the infinite-horizon lifetime value of a policy σ will equal the limit of finite horizon values, so that

$$v_\sigma = \lim_{k \to \infty} T_\sigma^k v. \tag{1.9}$$

Lemma 1.1.1 assures us that this is true: since T_σ is globally stable on V with unique fixed point v_σ, the limit in (1.9) exists and equals v_σ, independent of the terminal condition $v \in V$.

Figure 1.2 shows two arbitrarily chosen feasible policies and their lifetime values when $R = 1.04$, $\beta = 9.4$, $u(c) = 1 - \exp(c)$, and $Y_t = \exp(\nu Z_t)$ when $\nu = 0.1$ and Z_t is standard normal. The lifetime values were computed via (1.9).

1.1.2 Optimality

The **value function** for the optimal savings model is

$$v_{\top}(w) := \sup_{\sigma \in \Sigma} v_{\sigma}(w) \qquad (w \in \mathbb{R}_{+}). \tag{1.10}$$

Under Assumption 1.1.1 the supremum is always well defined in $\mathbb{R}$, since u and hence r_{σ} is bounded by some constant M, implying that, for any $w \in \mathbb{R}_{+}$ and $\sigma \in \Sigma$,

$$v_{\sigma}(w) \leqslant \sum_{t \geqslant 0} \beta^{t} P^{t} M = \frac{M}{1-\beta}.$$

A policy is called **optimal** if $v_{\sigma} = v_{\top}$; that is, if following the policy from every initial state w leads to the largest possible lifetime value attainable from w.

The set of feasible policies lies in an infinite-dimensional function space that we must search over to find an optimal policy. We want a systematic and efficient search procedure. Below we will show that an optimal policy can be obtained by first analyzing a functional equation called the Bellman equation, named after the mathematician Richard Bellman (1920–1984).

1.1.2.1 Bellman's Method

Fix $v \in V$. A policy $\sigma \in \Sigma$ is called **v-greedy** if

$$\sigma(w) \in \operatorname*{argmax}_{0 \leqslant c \leqslant w} \left\{ u(c) + \beta \int v(R(w-c)+y) \varphi(\mathrm{d}y) \right\} \quad \text{for all } w \geqslant 0. \tag{1.11}$$

In essence, a v-greedy policy uses v to value next-period states and then chooses consumption optimally to trade off current utility against expected discounted future value associated with the implied level of savings. The following statements are both true:

(i) Computing v-greedy policies is typically much easier than computing optimal policies, since we are only solving a two-period problem.

(ii) Computing v-greedy policies is can be equivalent to computing optimal policies, given the right choice of v.

What is the right choice of v? A natural candidate is the value function, since the value function tells us the maximal reward from alternative states. We explain this in more detail in §1.1.2.2. In that same section, we will also use the fact that that the value function satisfies an important functional equation, which we now describe.

We say that $v \in V$ satisfies the **Bellman equation** for the optimal savings problem if

$$v(w) = \max_{0 \leqslant c \leqslant w} \left\{ u(c) + \beta \int v(R(w-c)+y)\varphi(\mathrm{d}y) \right\} \quad \text{for all } w \geqslant 0. \tag{1.12}$$

Stating that v solves the Bellman equation is equivalent to stating that v is a fixed point of the **Bellman operator** T that maps a value function $v(w)$ into a value function $(Tv)(w)$ defined by

$$(Tv)(w) = \max_{0 \leqslant c \leqslant w} \left\{ u(c) + \beta \int v(R(w-c)+y)\varphi(\mathrm{d}y) \right\} \qquad (w \geqslant 0). \tag{1.13}$$

The next lemma discusses properties of greedy policies and the Bellman operator.

Lemma 1.1.2. *If Assumption 1.1.1 holds, then the function f defined by*

$$f(c,w) := u(c) + \beta \int v(R(w-c)+y)\varphi(\mathrm{d}y) \qquad (0 \leqslant c \leqslant w)$$

is continuous for all $v \in V$. In addition,

(i) *there exists at least one v-greedy policy for each $v \in V$ and*

(ii) *Tv is continuous and bounded whenever $v \in V$.*

Proof. The first claim follows from Assumption 1.1.1 and Lemma A.5.30 on page 223 (see, in particular, Example A.5.20). The second and third claims then follow from Theorem A.3.3 on page 179. □

Exercise 1.1.2. Prove that T is a contraction (see §A.2.2.2 for the definition) on $(V, \|\cdot\|)$. [Hint: Apply the sup inequality from Lemma A.1.2 on page 151.]

Since $(V, \|\cdot\|)$ is a Banach space, the contraction property in Exercise 1.1.2 implies that T is globally stable on V. (See §A.2.2.2 for details).

1.1.2.2 DP Results for Optimal Savings

Dynamic programming theory tells us that, under Assumption 1.1.1,

(S1) at least one optimal policy exists, (S2) the value function $v_\top$ is the unique solution to the Bellman equation in V, and

(S3) a policy $\sigma \in \Sigma$ is optimal if and only if it is $v_\top$-greedy.

A direct proof of (S1)–(S3) can be found in Stokey and Lucas (1989), Stachurski (2022) and numerous other sources. The proofs exploit the fact that the Bellman operator is a contraction mapping (as discussed in Exercise 1.1.2). Here, rather than providing a direct exposition of this theory, we develop more general results.

(As discussed above, and also in Sargent and Stachurski (2025b), we are motivated to provide general results because dynamic programming has become extremely diverse over the past few decades, and a large range of problems do not fit the simple format described above. At the same time, the optimal savings problem will be a useful reference point, providing a concrete example for the abstract definitions that begin in §1.2.1.)

1.1.2.3 Algorithms

The three most important algorithms for solving dynamic programming problems are value function iteration (VFI), Howard policy iteration, and optimistic policy iteration (OPI). In the present setting, they can be expressed as in Algorithms 1.1–1.3 respectively.

Algorithm 1.1: Value function iteration for the savings model

---

input $v \in \mathbb{R}^{\mathsf{X}}$, an initial guess of $v_\top$
input τ, a tolerance level for error
$\varepsilon \leftarrow \tau + 1$
while $\varepsilon > \tau$ **do**
  $v' \leftarrow Tv$
  $\varepsilon \leftarrow \|v - v'\|$
  $v \leftarrow v'$
end
return a v-greedy policy σ

VFI amounts to iterating k times with T from some initial condition $v \in V$ (where k is determined by a fixed tolerance level for error), producing an approximation $v_k := T^k v$ to $v_\top$, and then computing a v_k-greedy policy σ. This idea is natural, given that $v_\top$-greedy policies are optimal, since T is a contraction mapping and $v_\top$ is the unique fixed point (so that v_k is close to $v_\top$).

In HPI, one begins with a guess σ of the optimal policy and then iterates between computing the lifetime value of that policy (as given in (1.5)) and the corresponding

Algorithm 1.2: Howard policy iteration for the savings model

input $\sigma \in \Sigma$ and tolerance $\tau > 0$
$v \leftarrow v_\sigma$
repeat
  $\sigma \leftarrow$ a v-greedy policy
  $v' \leftarrow (I - \beta P_\sigma)^{-1} r_\sigma$
  if $\|v - v'\| < \tau$ **then** break
  $v \leftarrow v'$
return σ

Algorithm 1.3: Optimistic policy iteration for the savings model

input $m \in \mathbb{N}$ and tolerance $\tau > 0$
input $\sigma \in \Sigma$ and set $v \leftarrow (I - \beta P_\sigma)^{-1} r_\sigma$
repeat
  $\sigma \leftarrow$ a v-greedy policy
  $v' \leftarrow T_\sigma^m v$
  if $\|v - v'\| < \tau$ **then** break
  $v \leftarrow v'$
return σ

greedy policy. It can be shown that HPI is equivalent to Newton fixed point iteration applied to the Bellman operator (see, e.g., Chapter 5 of Sargent and Stachurski (2025a)).

Optimistic policy iteration can be thought of as a "convex combination" of VFI and HPI. Instead of computing the lifetime value $v_\sigma = (I - \beta P_\sigma)^{-1} r_\sigma$ of current policy σ, one computes instead $T_\sigma^m v$, which is an approximation to v_σ, as shown in (1.9). On one hand, if m is large, this approximation is tight and OPI is close to HPI. If $m = 1$, however, OPI reduces to VFI. OPI often outperforms both VFI and HPI for some intermediate values of m.

We discuss convergence of these algorithms in a general setting in §1.2.2.

Figure 1.3 shows an approximation of the optimal policy σ^* and the value function $v_\top$, both computed by OPI, for the same version of the optimal savings problem used in Figure 1.2. In this case we set $m = 20$.

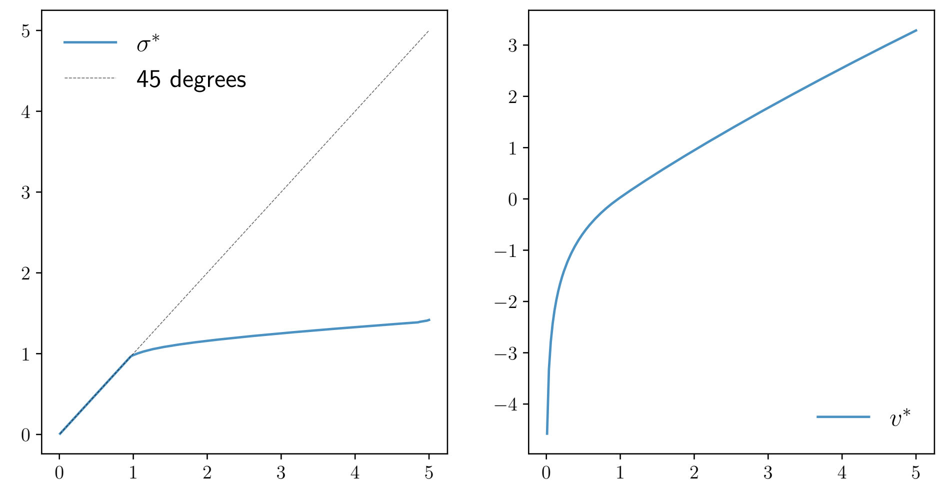


Figure 1.3: Approximating the optimal policy and value fuction via OPI

1.2 Introducing ADPs

Having reviewed the optimal savings problem, a relatively standard and familiar dynamic program, we now change course with a view to presenting a general theory of dynamic programming that includes optimal savings as one of many special cases. We begin with definitions and basic properties. Then we consider optimality.

1.2.1 Definition and Properties

In this section we define abstract dynamic programs and list some simple properties.

1.2.1.1 Definition

We define an **abstract dynamic program** (ADP) to be a pair $(V, \mathbb{T})$, where

(i) $V = (V, \preceq)$ is a partially ordered set and

(ii) $\mathbb{T} = \{T_\sigma : \sigma \in \Sigma\}$ is a nonempty family of order preserving self-maps on V.

In what follows,

- V is called the **value space**.

- Each operator T_σ in $\mathbb{T}$ is called a **policy operator**.
- Σ is an arbitrary index set and elements of Σ will be referred to as **policies**.

In applications we either seek or impose conditions under which each T_σ has a unique fixed point. In these settings, the significance of T_σ is that its fixed point represents the lifetime value of following policy σ. When it exists, we denote the unique fixed point of T_σ in V by v_σ and call it the **σ-value function**.

Remark 1.2.1. In many applications we will see that V is a function space and $\preceq$ is the pointwise partial order on V. It is, however, helpful to adopted a more general setting, where V is any poset and $\preceq$ is any partial order. For example, when we study LQ problems, we will find it convenient to take V to be a space of matrices and $\preceq$ to be the Loewner partial order. When we work in spaces of integrable functions, V will be a subset of L_p and $\preceq$ will be the almost everywhere partial order.

1.2.1.2 Greedy Policies

We saw in §1.1.2.2 that, for the optimal savings problem, the concept of greedy policies (see (1.11))) plays an key role in optimality theory. The same will be true in the ADP framework. To operationalize this idea, let $(V, \mathbb{T})$ be an ADP with policy set Σ. Given $v \in V$, we say that

$$\sigma \in \Sigma \text{ is } \boldsymbol{v}\textbf{-greedy} \text{ if } T_\tau\, v \preceq T_\sigma\, v \text{ for all } \tau \in \Sigma. \tag{1.14}$$

In other words, σ is v-greedy iff $T_\sigma\, v$ is a greatest element of $\{T_\tau\, v : T_\tau \in \mathbb{T}\}$.

The definition just given generalizes the notion of greedy policies for the optimal savings problem, as will be proved below (in §1.2.1.4).

Throughout the book we let

$$V_G := \{v \in V : \text{ at least one } v\text{-greedy policy exists}\}. \tag{1.15}$$

In applications, we typically find that solving for a greedy policy is much easier than solving the entire dynamic program directly. At the same time, there exist conditions under which solving the overall problem reduces to solving for a v-greedy policy with the “right” choice of v. We pursue this idea below.

1.2.1.3 The Bellman Equation

In §1.1.2.1 we saw how the Bellman equation plays a central role in optimality theory for the optimal savings problem. The same will be true for ADPs. Here we define the ADP Bellman equation and note some preliminary observations. Throughout this section, $(V, \mathbb{T})$ is an ADP with policy set Σ.

We say that $v \in V$ satisfies the **Bellman equation** if

$$v := \bigvee_{\sigma} T_{\sigma} v \qquad (v \in V). \tag{1.16}$$

In (1.16), the supremum is taken over all $\sigma \in \Sigma$. There is, in general, no guarantee that the supremum exists.

We define the **Bellman operator** generated by $(V, \mathbb{T})$ via

$$Tv := \bigvee_{\sigma} T_{\sigma} v \quad \text{whenever the supremum exists.} \tag{1.17}$$

We sometimes compress the definition of T to $T := \bigvee_{\sigma} T_{\sigma}$, with the understanding that the full definition is as given in (1.17). Evidently $v \in V$ satisfies the Bellman equation if and only if Tv exists and $Tv = v$.

The next lemma provides some essential facts about T on V_G (as defined in (1.15)).

Lemma 1.2.1. *The Bellman operator T has the following properties:*

(i) *T is well-defined and order preserving on V_G.*

(ii) *For $v \in V_G$ we have*

  (a) *$T_{\sigma} v \preceq Tv$ for all $\sigma \in \Sigma$ and*

  (b) *$T_{\sigma} v = Tv$ if and only if σ is v-greedy.*

Proof. We begin with part (ii). Fix $v \in V_G$ and let σ be v-greedy. Then, by definition, $T_{\sigma} v$ is the greatest element of $\{T_{\tau} v\}_{\tau \in \Sigma}$. A greatest element is also a supremum, so we have $Tv := \bigvee_{\tau \in \Sigma} T_{\tau} v = T_{\sigma} v$. This gives both (a) and $\Leftarrow$ in (b) of part (ii). For $\Rightarrow$ of (b), if $Tv = T_{\sigma} v$, then $T_{\tau} v \preceq T_{\sigma} v$ for all $\tau \in \Sigma$. In particular, σ is v-greedy.

Next we prove (i). For $v \in V_G$, a v-greedy policy exists, so Tv is well-defined by (b) of part (ii). Regarding the order preserving claim, fix $v, w \in V_G$ with $v \preceq w$. Let $\sigma \in \Sigma$ be v-greedy. Since T_{σ} is order preserving, we have $Tv = T_{\sigma} v \preceq T_{\sigma} w \preceq Tw$. □

1.2.1.4 Example: Optimal Savings as an ADP

Consider the optimal savings model from §1.1. We can represent this model as an ADP by taking $V := bm\mathbb{R}_+$ as the value space, paired with the pointwise order $\leqslant$, letting Σ be the set of (Borel measurable) feasible policies, as defined in §1.1.1.1, and setting $\mathbb{T}_{\text{OS}} := \{T_\sigma : \sigma \in \Sigma\}$, where each policy operator T_σ is as given in (1.3). It is straightforward to verify that each $T_\sigma \in \mathbb{T}$ is order preserving under $\leqslant$, and Exercise 1.1.1 confirms that T_σ maps V to itself. Hence $(V, \mathbb{T}_{\text{OS}})$ is an ADP.

By Lemma 1.1.1, each policy operator T_σ has a unique fixed point (i.e., σ-value function) v_σ. Consistent with the discussion in §1.2.1.1, the real number $v_\sigma(w)$ represents the lifetime value of policy σ, conditional on initial wealth state $W_0 = w$.

In (1.11) we defined the concept of a v-greedy policy for the optimal savings model. Later, in (1.14), we introduce the notion of a v-greedy policy for an arbitrary ADP. The second definition is a generalization of the first: if σ obeys the optimal savings greedy condition (1.11) and τ is any other feasible policy, then

$$u(\tau(w)) + \beta \int v(R(w - \tau(w)) + y)\varphi(\mathrm{d}y) \leqslant u(\sigma(w)) + \beta \int v(R(w - \sigma(w)) + y)\varphi(\mathrm{d}y)$$

for all $w \in \mathbb{R}_+$. This is equivalent to the statement $T_\tau\, v \leqslant T_\sigma\, v$ for all $\tau \in \Sigma$, which in turn is the ADP definition of v-greedy (when the value space is $V = bm\mathbb{R}_+$ and the partial order is $\leqslant$).

In addition, the ADP Bellman operator for $(V, \mathbb{T}_{\text{OS}})$, as defined in (1.17), is a generalization of the optimal savings Bellman operator given in (1.13). To see this, let $T = \bigvee_\sigma T_\sigma$ be the ADP Bellman operator and fix $v \in V$. By (ii) of Lemma 1.2.1, we have $Tv = T_\sigma\, v$ whenever σ is v-greedy. We just agreed that, in the setting of $(V, \mathbb{T}_{\text{OS}})$, this v-greedy property means that

$$\sigma(w) \in \operatorname*{argmax}_{0 \leqslant c \leqslant w} \left\{ u(c) + \beta \int v(R(w - c) + y)\varphi(\mathrm{d}y) \right\} \quad \text{for all } w \geqslant 0.$$

Taking such a policy (which exists by 1.1.2), fixing $w \in \mathbb{R}_+$ and combining these facts, we get

$$(Tv)(w) = (T_\sigma v)(w) = \max_{0 \leqslant c \leqslant w} \left\{ u(c) + \beta \int v(R(w - c) + y)\varphi(\mathrm{d}y) \right\}.$$

This confirms the claim that, in the setting of $(V, \mathbb{T}_{\text{OS}})$, the ADP Bellman operator reduces to the optimal savings Bellman operator in (1.13).

1.2.1.5 Properties

Below we list properties that will be useful for ADP optimality theory. We call $(V, \mathbb{T})$

- **well-posed** if each $T_\sigma \in \mathbb{T}$ has a unique fixed point in V,
- **regular** if $V_G = V$; that is, if a v-greedy policy exists for all $v \in V$,
- **bounded above** if there exists a $u \in V$ with $T_\sigma u \preceq u$ for all $T_\sigma \in \mathbb{T}$,
- **downward stable** if each $T_\sigma \in \mathbb{T}$ is downward stable (see §A.1.2.8) on V,
- **upward stable** if each $T_\sigma \in \mathbb{T}$ is upward stable (see §A.1.2.8) on V,
- **order stable** if each $T_\sigma \in \mathbb{T}$ is order stable (see §A.1.2.8) on V,
- **strongly order stable** if each $T_\sigma \in \mathbb{T}$ is strongly order stable on V, and
- **order continuous** if each $T_\sigma \in \mathbb{T}$ is order continuous (see §A.5.1.3) on V.

Well-posedness is an essential condition because we need well-defined lifetime values in order to maximize them, and maximizing lifetime values (or, equivalently, minimizing lifetime costs) is the objective of all dynamic programming problems. Regularity is desirable but does not always hold. Since $V = V_G$ under regularity, Lemma 1.2.1 implies that T is well-defined on all of V whenever this property holds.

Order stability is a relatively weak condition that generalizes conditions such as contractivity. Order continuity will hold automatically in many applications and is useful for studying optimality. These ideas will become clearer in the applications below.

Example 1.2.1. Consider the optimal savings ADP $(V, \mathbb{T}_{\mathrm{OS}})$ from §1.2.1.4, where $V = bm\mathbb{R}_+$ and each $T_\sigma \in \mathbb{T}_{\mathrm{OS}}$ has the form $T_\sigma v = r_\sigma + \beta P_\sigma v$ with $r_\sigma \in V$ and P_σ being a Markov operator sending V to itself. In view of Lemma 1.1.2, a v-greedy policy exists for every $v \in V$. Hence $(V, \mathbb{T}_{\mathrm{OS}})$ is regular. In §1.2.1.4 we discussed the fact that each $T_\sigma \in \mathbb{T}_{\mathrm{OS}}$ has a unique fixed point, so $(V, \mathbb{T}_{\mathrm{OS}})$ is well-posed. Lemma 1.1.1 on page 4 implies that each policy operator in $\mathbb{T}_{\mathrm{OS}}$ is globally stable. Lemma A.5.17 now implies that $(V, \mathbb{T}_{\mathrm{OS}})$ is strongly order stable.

Exercise 1.2.1. Show that the ADP $(V, \mathbb{T}_{\mathrm{OS}})$ in Example 1.2.1 is bounded above and order continuous.

Exercise 1.2.2. Let $(V, \mathbb{T})$ be a regular and well-posed ADP. Show that if $(V, \mathbb{T})$ is downward stable and bounded above, then the set V_Σ has an upper bound in V.

The next lemma shows that order continuity is sometimes passed from the policy operators to the Bellman operator. The order-theoretic completeness concept used in the lemma is introduced in §A.5.1.

Lemma 1.2.2. *Let $(V, \mathbb{T})$ be regular. If $(V, \mathbb{T})$ is order continuous and V is σ-chain complete, then the Bellman operator is order continuous on V.*

Proof. Let $(V, \mathbb{T})$ be as stated and let $T = \bigvee_\sigma T_\sigma$ be the Bellman operator. Fix $(v_n) \subset V$ with $v_n \uparrow \bar{v} \in V$. The sequence (Tv_n) is well-defined because $(V, \mathbb{T})$ is regular, and increasing because T is order preserving. Since V is σ-chain complete, $\bigvee_n Tv_n$ exists in V. We claim that $\bigvee_n Tv_n = T\bar{v}$.

On one hand, $T\bar{v}$ is an upper bound for (Tv_n). On the other hand, if $w \in V$ is another upper bound of (Tv_n) so that $Tv_n \preceq w$ for all n, then $T_\sigma v_n \preceq w$ for all n and σ. Fixing $\sigma \in \Sigma$, taking the supremum over n and using order continuity of T_σ gives $T_\sigma \bar{v} \preceq w$. Hence $T\bar{v} \preceq w$, which means that $T\bar{v}$ is a least upper bound of (Tv_n). This confirms that $\bigvee_n Tv_n = T\bar{v}$. Hence T is order continuous. □

Exercise 1.2.3. Let $(V, \mathbb{T})$ be an ADP and let V be σ-Dedekind complete. Prove that if $(V, \mathbb{T})$ is order continuous and order stable, then $(V, \mathbb{T})$ is strongly order stable.

1.2.1.6 Subsets of the Value Space

Let $(V, \mathbb{T})$ be an ADP. We often refer to the following three subsets of the value space V, the first of which was already introduced in §1.2.1.2:

- $V_G :=$ all $v \in V$ such that at least one v-greedy policy exists.
- $V_U :=$ all $v \in V$ with $v \preceq Tv$.
- $V_\Sigma :=$ all $v \in V$ such that $T_\sigma v = v$ for some $T_\sigma \in \mathbb{T}$.

In the last definition, we typically assume that $(V, \mathbb{T})$ is well-posed with policy set Σ, so V_Σ is the set of all σ-value functions (lifetime values) generated by $(V, \mathbb{T})$.

Example 1.2.2. For the optimal savings problem described in §1.1, the utility function u is in V_U. Indeed, for any $w \in \mathbb{R}_+$, by the definition of the Bellman operator,

$$u(w) \leqslant \max_{0 \leqslant c \leqslant w} \left\{ u(c) + \beta \int u(R(w-c)+y)\varphi(\mathrm{d}y) \right\} = (Tu)(w).$$

In the next lemma, $(V, \mathbb{T})$ is any ADP and V_G, V_U, and V_Σ are as defined above.

Lemma 1.2.3. *If $V_\Sigma \subset V_G$, then $V_\Sigma \subset V_U$.*

Proof. Fix $v_\sigma \in V_\Sigma$ and let T_σ be such that v_σ is a fixed point. By Lemma 1.2.1, we have $v_\sigma = T_\sigma v_\sigma \preceq T v_\sigma$ on V_G. □

EXERCISE 1.2.4. Show that V_U is nonempty whenever $(V, \mathbb{T})$ is regular and well-posed.

1.2.1.7 Optimality and the Bellman Equation

We say that a policy $\sigma \in \Sigma$ is **optimal** for $(V, \mathbb{T})$ if v_σ is a greatest element of V_Σ. In other words, σ is optimal if it attains the "highest possible" lifetime value.

Example 1.2.3. Consider the optimal savings ADP $(V, \mathbb{T})$ introduced in §1.2.1.4. According to the definition just given, a policy σ is optimal if and only if $v_\tau(w) \leqslant v_\sigma(w)$ for every $w \in \mathbb{R}_+$ and every feasible policy τ. (We are using the fact that the partial order on V is the pointwise order.) This is identical to the original definition of optimality we gave for the savings problem in §1.1.2.

Perhaps the most important part of aspect of the theory of dynamic programming is the link between optimality and the Bellman equation. To clarify this link we introduce some terminology. Let $(V, \mathbb{T})$ be a well-posed ADP and set

$$v_\top := \bigvee_\sigma v_\sigma := \bigvee V_\Sigma \quad \text{whenever the supremum exists.}$$

When $v_\top$ exists (i.e., when the supremum exists) we call $v_\top$ the **value function** of the ADP. The following statements are obvious from the definitions:

- Existence of an optimal policy σ implies that $v_\top$ exists and is equal to v_σ.
- If $v_\top = v_\sigma$ for some $\sigma \in \Sigma$, then σ is optimal.

At the same time, existence of $v_\top$ does not generally imply existence of a greatest element (and hence an optimal policy).

We say that **Bellman's principle of optimality holds** if

$$\{\sigma \in \Sigma : \sigma \text{ is optimal }\} = \{\sigma \in \Sigma : \sigma \text{ is } v_\top\text{-greedy }\}. \tag{1.18}$$

(When $v_\top$ does not exist both sets are understood as empty.)

Now let $(V, \mathbb{T})$ be both regular and well-posed. We say that the **fundamental optimality properties hold** for $(V, \mathbb{T})$ if

(B1) at least one optimal policy exists,

(B2) $v_\top$ is the unique solution to the Bellman equation in V, and

(B3) Bellman's principle of optimality holds.

Note that (B2) makes sense when (B1) holds because (B1) implies existence of the value function $v_\top$. Also, $(V, \mathbb{T})$ is regular, by assumption, so the Bellman equation is always well defined (in the sense that the supremum exists).

All three parts of (B1)–(B3) are important. (B1) tells us that the problem at hand has a solution. (B3) implies that a solution can be computed by taking a $v_\top$-greedy policy. To operationalize this idea we need to calculate $v_\top$. (B2) provides a restriction that can help us with this task.

Properties (B1)–(B3) are not independent. For example, (B2) implies (B3), as the next lemma shows.

Lemma 1.2.4. *Let $(V, \mathbb{T})$ be well-posed. If $v_\top$ exists in V and satisfies the Bellman equation, then Bellman's principle of optimality holds.*

Proof. Let $v_\top$ exist in V. We prove the equality in (1.18) when $Tv_\top = v_\top$. Suppose first that $\sigma \in \Sigma$ is optimal, so that $v_\sigma = v_\top$. Since $T_\sigma v_\sigma = v_\sigma$, this implies $T_\sigma v_\top = v_\top$. But $T v_\top = v_\top$, so $T_\sigma v_\top = Tv_\top$. Hence σ is $v_\top$-greedy (by Lemma 1.2.1). Suppose next that σ is $v_\top$-greedy, so that $T_\sigma v_\top = Tv_\top = v_\top$. But v_σ is the unique fixed point of T_σ in V, so $v_\sigma = v_\top$. Hence σ is an optimal policy. □

When regularity also holds we can say more:

Lemma 1.2.5. *When $(V, \mathbb{T})$ is well-posed and regular, the following statements are valid:*

(i) *$v_\top$ exists and satisfies the Bellman equation if and only if an optimal policy exists and Bellman's principle of optimality holds.*

(ii) *$v_\top$ exists and is the unique solution to the Bellman equation in V if and only if all of the fundamental optimality properties hold.*

Proof. Let $(V, \mathbb{T})$ be as stated. Regarding part (i, $\Rightarrow$), suppose that $v_\top$ exists and $Tv_\top = v_\top$. Using regularity, we take σ to be $v_\top$-greedy. Then $T_\sigma v_\top = Tv_\top = v_\top$. But v_σ is the unique fixed point of T_σ, so $v_\sigma = v_\top$. Hence σ is an optimal policy. Bellman's principle of optimality follows from Lemma 1.2.4. Regarding (ii, $\Leftarrow$), let $\sigma \in \Sigma$ be optimal, so that $v_\sigma = v_\top$. By Bellman's principle of optimality, the policy σ is $v_\top$-greedy. As a result, $Tv_\top = T_\sigma v_\top = T_\sigma v_\sigma = v_\sigma = v_\top$. In particular, $v_\top$ exists and satisfies the Bellman equation.

Regarding (ii, ⇒), if $v_\top$ exists and is the unique solution to the Bellman equation in V, then, by (i), an optimal policy exists and Bellman's principle of optimality holds. This confirms all of (B1)–(B3), so ⇒ of (ii) is valid. The claim (ii, ⇐) is trivial, so the proof of Lemma 1.2.5 is complete. □

The key message of Lemma 1.2.5 is that, at least for regular ADPs, our main goal is to construct conditions under which $v_\top$ exists and is the unique fixed point of T in V. We begin this task in §1.2.3.

The next exercise generalizes a well-known result from more traditional dynamic programming frameworks (see, e.g., Puterman (2005), Theorem 6.2.6).

Exercise 1.2.5. Let $(V, \mathbb{T})$ be regular and well-posed. Prove the following: If $v_\top$ exists and is the unique fixed point of T in V, then $\sigma \in \Sigma$ is optimal if and only if $T v_\sigma = v_\sigma$.

1.2.1.8 From Fixed Points to Value Functions

Now we state conditions under which properties (B1)–(B3) on page 17 hold. The theorem assumes existence of a fixed point for the Bellman operator and will be used as an intermediate result. In the statement, V_Σ and V_G are as defined in §1.2.1.6 and T is the Bellman operator.

Theorem 1.2.6. *If $(V, \mathbb{T})$ is downward stable and T has a fixed point in V_G, then*

(i) *the fixed point of T equals $v_\top$ and $v_\top$ is the greatest element of V_Σ,*

(ii) *T has no other fixed point in V_G, and*

(iii) *Bellman's principle of optimality holds.*

Proof. Let $(V, \mathbb{T})$ be as stated and let v be a fixed point of T in V_G. By the definition of V_G and the characterization of greedy policies in Lemma 1.2.1, we can choose a $\sigma \in \Sigma$ such that $v = T\,v = T_\sigma\,v$. By well-posedness, T_σ has a unique fixed point v_σ in V, so $v = v_\sigma$. Moreover, if τ is any policy, then $T_\tau\,v_\sigma \preceq T\,v_\sigma = v_\sigma$ and hence, by downward stability, $v_\tau \preceq v_\sigma$. These facts imply that v is a greatest element of V_Σ. Hence (i) holds. Moreover, if v' is another fixed point of T in V_G, then, by the same argument, v' is also a greatest element of V_Σ. Since greatest elements are unique, we have $v' = v$. This proves part (ii) of the proposition. Part (iii) follows from (i) and Lemma 1.2.4. □

The conclusions of Theorem 1.2.6 are similar to but weaker than the fundamental optimality properties on page 17. In particular, it is possible that T has other fixed points outside of V_G. Of course, if $V_G = V$, then this possibility is removed and the fundamental optimality properties hold. We state this as a corollary.

Corollary 1.2.7. *Let $(V, \mathbb{T})$ be regular. If $(V, \mathbb{T})$ is downward stable and T has a fixed point in V, then the fundamental optimality properties hold.*

Like Theorem 1.2.6, Corollary 1.2.7 is mainly useful as an input. We use it below to establish a range of optimality results based on various sufficient conditions. These later results are aimed at applications.

1.2.2 Algorithms

In this section we discuss algorithms for dynamic programming and present some preliminary results. The three algorithms we consider are value function iteration (VFI), Howard policy iteration (HPI) and optimistic policy iteration (OPI). These algorithms generalize the ones presented for the optimal savings case in §1.1.2.3.

1.2.2.1 Operators

As a first step, we introduce two operators. Throughout §1.2.2.1 we take $(V, \mathbb{T})$ to be a fixed ADP. As usual, when $(V, \mathbb{T})$ is well-posed, the unique fixed point of $T_\sigma \in \mathbb{T}$ in V is denoted by v_σ. We define the **Howard policy operator** corresponding to $(V, \mathbb{T})$ via

$$H \colon V_G \to V_\Sigma, \qquad Hv = v_\sigma \quad \text{where } \sigma \text{ is } v\text{-greedy},$$

as well as, for each $m \in \mathbb{N}$, the **optimistic policy operator**

$$W \colon V_G \to V, \qquad Wv = T_\sigma^m v \quad \text{where } \sigma \text{ is } v\text{-greedy}. \tag{1.19}$$

(Here and below, the dependence of W on m is often suppressed to simplify notation.)

Like the Bellman operator T, the map W is well-defined on all of V when $(V, \mathbb{T})$ is regular. The Howard policy operator H is well-defined on all of V when $(V, \mathbb{T})$ is well-posed and regular. Note that, for both of these maps, we always select the same v-greedy policy when applying them to some fixed v.[2]

Below we will associate VFI, OPI, and HPI with fixing a $v \in V$ and then iteratively applying the operators T, W, and H respectively. A small amount of thought will convince you that, for the optimal savings ADP described in §1.2.1.4, these iterative procedures coincide with our earlier description of VFI, HPI, and OPI from §1.1.2.3.

[2]In general, designating a v-greedy policy for all $v \in V$ requires the Axiom of Choice. In practice, many applications induce some structure on the policy set that can be used to produce simple selection mechanisms.

The next lemma collects useful results for the operators introduced above. In the statement, $(V, \mathbb{T})$ is an ADP with Howard policy operator H, optimistic policy operator W and Bellman operator T. We assume regularity, so that $V_G = V$.

Lemma 1.2.8. *If $(V, \mathbb{T})$ is regular and well-posed, then the following statements hold.*

(L1) *If $v \in V$ with $Hv = v$, then $Tv = v$.*

(L2) *The operators T, W and H all map V_U to itself.*

(L3) *If $v \in V_U$, then $Tv \preceq Wv \preceq T^m v$.*

Proof. Regarding (L1), fix $v \in V$ with $H\,v = v$ and let σ be a v-greedy policy such that $H\,v = v_\sigma$. Then $v_\sigma = v$. Since σ is v-greedy, $T_\sigma\,v = T\,v$. Since v_σ is fixed for T_σ, we also have $T_\sigma\,v = v$. Combining the last two equalities proves (L1).

Regarding (L2), fix $v \in V_U$. Since $v \preceq Tv$ and T is order preserving on V_U, we have $Tv \preceq TTv$. Hence $Tv \in V_U$. Regarding W, let σ be v-greedy with $W = T_\sigma^m v$. Since T and T_σ are order preserving and $v \preceq Tv$, we have $Wv = T_\sigma T_\sigma^{m-1} v \preceq TT_\sigma^{m-1} v \preceq TT_\sigma^{m-1} Tv = TT_\sigma^m v = TWv$. Hence $Wv \in V_U$. Finally, regarding H, we observe that $Hv \in V_\Sigma$ and, by Lemma 1.2.3, $V_\Sigma \subset V_U$.

To prove (L3) we fix $v \in V_U$. Letting σ be v-greedy, we have $v \preceq Tv = T_\sigma\,v$. Iterating on this inequality with T_σ proves that $(T_\sigma^k\,v)$ is increasing. In particular, $Tv = T_\sigma\,v \preceq Wv$. For the second inequality in (L3) we use the fact that $T_\sigma \preceq T$ on V and T and T_σ are both order preserving to obtain $Wv = T_\sigma^m v \preceq T^m v$. □

The next lemma adds upward stability and derives additional implications.

Lemma 1.2.9. *If $(V, \mathbb{T})$ is regular, well-posed, and upward stable, then*

$$v \in V_U \qquad \Longrightarrow \qquad T^n v \preceq W^n v \quad \textit{and} \quad T^n v \preceq H^n v \tag{1.20}$$

for all $n \in \mathbb{N}$. Moreover, the sequences $(T^n v)$, $(H^n v)$, and $(W^n v)$ are all increasing.

Proof. Our first claim is that

$$u, v \in V_U \text{ with } u \preceq v \implies Tu \preceq Wv \text{ and } Tu \preceq Hv. \tag{1.21}$$

To show this we fix such u, v and take σ to be v-greedy. Let v_σ be the σ-value function, so that $T_\sigma\,v_\sigma = v_\sigma$ and $v_\sigma = Hv$. Since $v \in V_U$ we have

$$v \preceq Tv = T_\sigma\,v \preceq T_\sigma^m\,v = Wv \preceq v_\sigma = Hv. \tag{1.22}$$

The second inequality is by iterating on $v \preceq T_\sigma v$, while the third is by upward stability. Since $Tu \preceq Tv$, we can use (1.22) to obtain (1.21). Iterating on (1.21) produces (1.20). The last claim in Lemma 1.2.9 follows from (1.22), which tells us that elements of V_U are mapped up by T, W, and H. □

1.2.2.2 Convergence

In this section, we assume that $(V, \mathbb{T})$ *is a regular ADP and the fundamental optimality properties on page 17 hold*. Let $v_\top$ denote the value function. In this setting, we say that

- **VFI converges** if $T^n v \uparrow v_\top$ for all $v \in V_U$,
- **OPI converges** if $W^n v \uparrow v_\top$ for all $v \in V_U$, and
- **HPI converges** if $H^n v \uparrow v_\top$ for all $v \in V_U$.

If, for all $v \in V_U$, there exists an $n \in \mathbb{N}$ with $H^n v = v_\top$, we say that HPI converges in finitely many steps. For OPI convergence, the meaning is that convergence holds for any choice of the OPI step size $m \in \mathbb{N}$.

EXERCISE 1.2.6. Prove that convergence of OPI implies convergence of VFI.

Remark 1.2.2. Despite the implication in Exercise 1.2.6, our statements of results often mention VFI explicitly – mainly for the benefit of readers who skim through books reading only the main theorems.

The following result is a direct corollary of Lemma 1.2.9.

Corollary 1.2.10. *If* $(V, \mathbb{T})$ *is regular and upward stable, then*

$$\text{convergence of VFI} \quad \Longrightarrow \quad \text{convergence of OPI and HPI}.$$

Proof. Assume the conditions of the corollary and fix $v \in V_U$. Since an optimal policy exists, $v_\top$ exists and is the greatest element of V_Σ. Lemma 1.2.9 yields $v \preceq T^n v \preceq W^n v \preceq v_\top$ for all n, where the last inequality follows from (1.22) and the fact that $v_\sigma \preceq v_\top$ for all σ. Hence convergence of VFI implies convergence of OPI. The proof for HPI is similar. □

The following lemma is also useful when studing convergence of the algorithms.

Lemma 1.2.11. *If* $(V, \mathbb{T})$ *is regular and upward stable, then* $v \preceq v_\top$ *for all* $v \in V_U$.

Proof. Let $(V, \mathbb{T})$ be regular and upward stable, and fix $v \in V_U$. By regularity, there exists a $\sigma \in \Sigma$ with $Tv = T_\sigma v$. Since $v \in V_U$, we have $v \preceq T_\sigma v$. Combining this inequality and upward stability gives $v \preceq v_\sigma$, where v_σ is the fixed point of T_σ. As a consequence, $v \preceq v_\top$. □

1.2.3 Optimality: Sufficient Conditions for Well-Posed ADPs

In this section we introduce conditions for optimality of ADPs that are well-posed and regular. These high-level results will later be used as inputs for lower level results that are more straightforward to check in applications.

1.2.3.1 Case I: Finite ADPs

In applications we often deal with dynamic programs that have finitely many states and actions. This finiteness implies that the set of feasible policies is finite. The next result deals with this case.

Theorem 1.2.12. *Let $(V, \mathbb{T})$ be regular and well-posed. If $(V, \mathbb{T})$ is order stable and $\mathbb{T}$ is finite, then*

(i) *the fundamental optimality properties hold and*

(ii) *HPI converges in finitely many steps.*

Proof. Let $(V, \mathbb{T})$ be as stated. Let v be any element of V_U (which is nonempty by Lemma 1.2.3). Since $(V, \mathbb{T})$ is well-posed and regular, the Howard policy operator H is well-defined on V. Let $v_n = H^n v$ for all $n \geqslant 0$. By Lemma 1.2.9, we have $v_n \preceq v_{n+1}$ for all n. Since (v_n) is contained in the finite set V_Σ, it must be that $v_{n+1} = v_n$ for some $n \in \mathbb{N}$. But then $Hv_n = v_n$, so, by Lemma 1.2.8, we have $Tv_n = v_n$. Since T has a fixed point in V, the fundamental optimality properties hold (Corollary 1.2.7). We have also shown that HPI converges in finitely many steps. □

1.2.3.2 Case II: Chain Complete Value Space

Now we replace finiteness with certain forms of order completeness (see §A.5.1). These completeness conditions will imply that V is order bounded, so that we can take $v_\perp, v_\top \in V$ with $v_\perp \preceq v \preceq v_\top$ for all $v \in V$. In many applications, such bounds can be constructed by working with the primitives. (Later, in §1.2.3.3, we will weaken this restriction.)

Theorem 1.2.13. *Let V be chain complete and let $(V, \mathbb{T})$ be regular and well-posed. In this setting, the fundamental optimality properties hold. If, in addition, $(V, \mathbb{T})$ is strongly order stable, then VFI, OPI and HPI all converge.*

Proof. Let $(V, \mathbb{T})$ be as stated and let T be the Bellman operator generated by $(V, \mathbb{T})$. By Corollary 1.2.7, it suffices to show that T has a fixed point in V and $(V, \mathbb{T})$ is downward stable. Since T is order preserving (Lemma 1.2.1) and V is chain complete, the first property follows from the Knaster–Tarski fixed point theorem. The second holds by Lemma A.5.3.

Regarding convergence of VFI, fix $v \in V_U$ and let $v_n := T^n v$. By Lemma 1.2.11, we have $v_n \preceq v_\top$ for all n. Also, let $\perp$ be the least element of V, let σ be an optimal policy and let $w_n := T_\sigma^n \perp$. We have $w_n \preceq T^n \perp \preceq T^n v = v_n$ and, by strong order stability, $w_n \uparrow v_\top$. Hence $v_n \uparrow v_\top$. Convergence of OPI and HPI now follow from Corollary 1.2.10. □

Theorem 1.2.13 is rather striking: It says that, for regular well-posed ADPs, chain completeness of the value space is enough to guarantee all of the optimality and convergence properties we seek. In other words, in nice environments (from an order-theoretic perspective), ADPs are essentially well-behaved.

At the same time, Theorem 1.2.13 has obvious limitations. The chain completeness assumption is relatively strong and the ADP is assumed to be well-posed, which is often challenging to establish. Below we examine what happens when we weaken these assumptions. (At the same time, we will always impose enough conditions to ensure well-posedness, without which our dynamic programming problems are not well defined.)

Next we add order continuity to the conditions of Theorem 1.2.13, which allows us to obtain both the fundamental optimality properties (as given on page 17) and convergence of algorithms (see page 22).

Theorem 1.2.14. *Let V be σ-chain complete and let $(V, \mathbb{T})$ be regular and well-posed. If $(V, \mathbb{T})$ is order continuous, then*

(i) *the fundamental optimality properties hold and*

(ii) *VFI, OPI and HPI all converge.*

Proof. Let $(V, \mathbb{T})$ be an ADP satisfying the conditions of Theorem 1.2.14. Under these conditions, T is order continuous (Lemma 1.2.2) and $(V, \mathbb{T})$ is order stable (Lemma A.5.8). In view of Corollary 1.2.7, to obtain the fundamental optimality results, we need only to show that T has a fixed point in V and $(V, \mathbb{T})$ is downward stable. The first property holds by the Tarski–Kantorovich theorem, while the second holds by order stability.

To show that VFI, OPI and HPI converge, we fix $v \in V_U$, so that $v \preceq Tv$. Since T is order continuous, the Tarski–Kantorovich theorem (page 205) implies that $\bar{v} := \bigvee_n T^n v$ is a fixed point of T with $T^n v \uparrow \bar{v}$. Since the fundamental optimality properties hold, $v_\top$ is the only fixed point of T in V (this is (B2) on page 17). Hence $\bar{v}$ is equal to $v_\top$ and VFI converges. Convergence of OPI and HPI now follow from Corollary 1.2.10. □

1.2.3.3 Case III: Dedekind Complete Value Space

The conditions in §1.2.3.2 require that V is order bounded. Here we loosen this restriction.

Theorem 1.2.15. *Let V be Dedekind complete and let $(V, \mathbb{T})$ be regular and well-posed. If $(V, \mathbb{T})$ is also bounded above and order stable, then the fundamental optimality properties hold.*

Proof. Let $(V, \mathbb{T})$ be as stated and let T be the corresponding Bellman operator. We claim that T has a fixed point in V. To see this, set $\bar{v} := \bigvee_\sigma v_\sigma$, which exists because $(V, \mathbb{T})$ is bounded above (which implies that V_Σ is bounded above, as shown in Exercise 1.2.2) and V is Dedekind complete. For any $\sigma \in \Sigma$, we have $v_\sigma = T_\sigma v_\sigma \preceq T_\sigma \bar{v} \preceq T\bar{v}$. Taking the supremum over σ gives $\bar{v} \preceq T\bar{v}$. Thus, $\bar{v} \in V_U$. Applying Lemma 1.2.9 and $HV_\Sigma \subset V_\Sigma$ again, we get $\bar{v} \preceq T\bar{v} \preceq H\bar{v} \preceq \bar{v}$. In particular, $\bar{v}$ is a fixed point of T. Since $(V, \mathbb{T})$ is order stable and regular, the claim in Theorem 1.2.15 now follows from Corollary 1.2.7. □

The next result replaces Dedekind completeness with σ-Dedekind completeness and order continuity. Strong order stability is not required.

Theorem 1.2.16. *Let V be σ-Dedekind complete and let $(V, \mathbb{T})$ be regular and well-posed. If $(V, \mathbb{T})$ is bounded above, order stable, and order continuous, then*

(i) *the fundamental optimality properties hold and*

(ii) *VFI, OPI and HPI all converge.*

Proof. In view of Corollary 1.2.7, the fundamental optimality properties will hold when T has a fixed point in V. To see that this is true, fix any $v \in V_U$ (which is nonempty by Lemma 1.2.3) and set $v_n := T^n v$. Since $(V, \mathbb{T})$ is bounded above and V is σ-Dedekind complete, there exists a $\bar{v} \in V$ with $v_n \uparrow \bar{v}$. We claim that $T\bar{v} = \bar{v}$. Indeed, $v_{n+1} = Tv_n \preceq T\bar{v}$ for all n, so, taking the supremum, $\bar{v} \preceq T\bar{v}$. For the reverse inequality we take σ to be $\bar{v}$-greedy and use order continuity of T_σ to obtain

$$T\bar{v} = T_\sigma \bar{v} = T_\sigma \bigvee_n v_n = \bigvee_n T_\sigma v_n \preceq \bigvee_n T v_n = \bigvee_n v_{n+1} = \bar{v}.$$

The fundamental optimality properties are now proved. In view of these properties, the only fixed point of T in V is $v_\top$. Hence $T^n v = v_n \uparrow \bar{v} = v_\top$. This proves convergence of VFI. Convergence of OPI and HPI follow from Corollary 1.2.10. □

1.2.4 Example: Back to Optimal Savings

In this section we investigate optimality properties of the simple optimal savings problem introduced in §1.1.

In §1.2.1.4, we converted the optimal savings problem from §1.1 into an ADP $(V, \mathbb{T}_{\mathrm{OS}})$, where $V = bm\mathbb{R}_+$ and each $T_\sigma \in \mathbb{T}_{\mathrm{OS}}$ takes the form

$$(T_\sigma v)(w) = u(\sigma(w)) + \beta \int v(R(w - \sigma(w)) + y)\varphi(\mathrm{d}y) \qquad (w \in \mathbb{R}_+). \tag{1.23}$$

We can now prove the following result.

Proposition 1.2.17. *The optimal savings ADP $(V, \mathbb{T}_{\mathrm{OS}})$ obeys the fundamental optimality properties and VFI, OPI and HPI converge.*

Proof. In §1.2.1.5, we showed that $(V, \mathbb{T}_{\mathrm{OS}})$ is regular, bounded above, order stable, and order continuous. Since V is σ-Dedekind complete (Corollary A.5.15), the claims in Proposition 1.2.17 follow from Theorem 1.2.16. □

Using these facts, we can recover the well-known optimality results for the optimal savings problem stated in §1.1.2.2. For example, we showed in §1.2.1.4 that $v \in V$ satisfies the ADP Bellman equation $\bigvee_\sigma T_\sigma v = v$ if and only if it satisfies

$$v(w) = \max_{0 \leqslant c \leqslant w} \left\{ u(c) + \beta \int v(R(w-c)+y)\varphi(\mathrm{d}y) \right\} \quad \text{for all } w \geqslant 0. \tag{1.24}$$

By this fact and the fundamental optimality properties and, the value function $v_\top$ exists and is the unique solution to (1.24) in V.

Under the ADP definition, a policy σ is v-greedy if and only if $T_\tau v \leqslant T_\sigma v$ for all $\tau \in \Sigma$. Since Σ contains all feasible Borel measurable maps, we see that σ has this property if and only if

$$\sigma(w) \in \operatorname*{argmax}_{0 \leqslant c \leqslant w} \left\{ u(c) + \beta \int v(R(w-c)+y)\varphi(\mathrm{d}y) \right\} \quad \text{for all } w \geqslant 0. \tag{1.25}$$

Applying Bellman's principle of optimality, a policy is optimal if and only if it satisfies (1.25) with v replaced by $v_\top$.

The optimal savings model considered in this section is relatively simplistic. Later, in §1.3.2, we will add additional features.

1.3 ADPs in Topological Space

In §1.2.3, we provided optimality and convergence results for well-posed ADPs. In practice, this condition can be difficult to establish. To handle this issue, we begin to add algebraic and topological structure to the value space. In this section, we focus on topological properties. Our aim is to provide optimality and convergence results build on earlier results while avoiding the assumption of well-posedness.

1.3.1 ADPs in Partially Ordered Space

We begin by adding topological structure to the value space. In particular, we assume that the value space is not only a poset but also a pospace (§A.5.3.1).

1.3.1.1 Globally Stable ADPs

Our first step is to introduce globally stable ADPs and then investigate their properties. Let $(V, \mathbb{T})$ be an ADP where $V = (V, \preceq)$ is a partially ordered space (pospace). We say that

> $(V, \mathbb{T})$ is **globally stable** if each $T_\sigma \in \mathbb{T}$ is globally stable on V.

Obviously, if $(V, \mathbb{T})$ is globally stable, then $(V, \mathbb{T})$ is well-posed. We also have the following useful preliminary result, which is an immediate consequence of Lemma A.5.17.

Lemma 1.3.1. *If $(V, \mathbb{T})$ is globally stable, then $(V, \mathbb{T})$ is strongly order stable.*

The next theorem handles globally stable ADPs that are not necessarily regular. In the statement, $(V, \mathbb{T})$ is an ADP and $V = (V, \preceq)$ is a partially ordered space. As usual V_G is all $v \in V$ for which at least one v-greedy policy exists. The set V_0 can be thought of as a "nice" subset of the value space where we can find greedy policies.

Theorem 1.3.2. *Let V be a pospace and let $(V, \mathbb{T})$ be globally stable. If there exists a $V_0 \subset V_G$ such that the Bellman operator T is globally stable on V_0, then*

(i) *at least one optimal policy exists,*

(ii) *the fixed point $v_\top$ of T is the greatest element of V_Σ,*

(iii) *σ is optimal if and only if σ is $v_\top$-greedy, and*

(iv) *$T^n v \to v_\top$ as $n \to \infty$ for all $v \in V_0$.*

Proof. Let the stated conditions hold. Part (i) holds whenever (ii) holds because existence of a greatest element of V_Σ implies existence of an optimal policy. Parts (ii)–(iii) follow from Theorem 1.2.6, since global stability of T on V_0 implies that T has a fixed point $v_\top$ in $V_0 \subset V_G$ and $(V, \mathbb{T})$ is downward stable and well-posed (Lemma 1.3.1).

Regarding (iv), fix $v \in V_0$. The sequence $(v_n) := (T^n v)$ is well-defined and takes values in V_0 because T is globally stable on V_0 and hence maps this set into itself (and T is always well-defined on V_G). In addition, global stability of T on V_0 implies $v_n \to v_\top$. □

Exercise 1.3.1. Show that, under the conditions of Theorem 1.3.2, VFI converges in order, in the sense that $T^n v \uparrow v_\top$ from any $v \in V_U \cap V_0$.

If, in Theorem 1.3.2, we can take $V_0 = V$, then $(V, \mathbb{T})$ is regular (because $V_0 \subset V_G$). In this setting, we obtain the fundamental optimality results (i.e., (B1)–(B3) on page 17). The next result contains this conclusion, while also weakening the stability condition on the Bellman operator T and proving convergence of all algorithms.

Theorem 1.3.3. *Let V be a pospace and let $(V, \mathbb{T})$ be regular and globally stable. If T has a fixed point in V, then*

(i) *the fundamental optimality properties hold and*

(ii) *VFI, OPI and HPI all converge.*

Proof. Let $(V, \mathbb{T})$ and T be as stated. Since $(V, \mathbb{T})$ is order stable (Lemma 1.3.1), part (i) follows from Corollary 1.2.7. Regarding convergence of VFI, fix $v \in V_U$, let σ be an optimal policy and let $v_\top$ be the value function. Since $v \preceq Tv$ and $T_\sigma v \preceq Tv \preceq v_\top$, (with the last inequality by Lemma 1.2.11) we have $T_\sigma^n v \preceq T^n v \preceq v_\top$ for all n. T_σ is globally stable with unique fixed point $v_\top$ and $T_\sigma^n v \preceq v_\top$ for all n, so $\bigvee_n T_\sigma^n v = v_\top$ (Lemma A.5.16). By this fact and $T_\sigma^n v \preceq T^n v \preceq v_\top$ for all n, we also have $T^n v \uparrow v_\top$. Hence VFI converges. Since $(V, \mathbb{T})$ is regular and upward stable (by Lemma 1.3.1), convergence of VFI implies convergence of OPI and HPI (Corollary 1.2.10). □

One disadvantage of Theorem 1.3.3 is that conditions are placed on the derived object T, rather than the primitives. Next we present an alternative result that uses global stability of the policy operators without placing assumptions directly on T. The conditions of the theorem are similar to those of Theorem 1.2.16, after replacing order stability and order continuity with global stability.

Theorem 1.3.4. *Let V be a pospace and let $(V, \mathbb{T})$ be regular, globally stable and bounded above. If V is σ-Dedekind complete, then*

(i) *the fundamental optimality properties hold and*

(ii) *VFI, OPI and HPI all converge.*

Proof. First we show that T has a fixed point in V. To see this, let $v_\top$ be such that $T_\sigma v_\top \preceq v_\top$ for all $\sigma \in \Sigma$ and let v be an element of V_Σ. By downward stability, $v \preceq v_\top$. Moreover, $v \preceq Tv$, since $V_\Sigma \subset V_U$, and $Tv_\top \preceq v_\top$. Hence T is a self-map on $[v, v_\top]$. Letting $v_n := T^n v$ and applying σ-Dedekind completeness, we have $v_n \uparrow \bar{v}$ for some $\bar{v} \in [v, v_\top]$. We claim that $T\bar{v} = \bar{v}$. To see this, first observe that $v_{n+1} = Tv_n \preceq T\bar{v}$ for all n, so $\bar{v} \preceq T\bar{v}$. Letting σ be $\bar{v}$-greedy, we have $\bar{v} \preceq T\bar{v} = T_\sigma \bar{v}$, so, by upward stability (which follows from Lemma A.5.17), $\bar{v} \preceq v_\sigma$. Also, letting $w_n := T_\sigma^n v$, we have $w_n \preceq v_n \preceq \bar{v} \preceq v_\sigma$ for all $n \in \mathbb{N}$. By global stability, $w_n \to v_\sigma$. Applying Lemma A.5.16, we also have $\vee_n w_n = v_\sigma$. From this fact and the previous chain of inequalities, it must be that $\bar{v} = v_\sigma$. Hence $T\bar{v} = T_\sigma \bar{v} = T_\sigma v_\sigma = v_\sigma = \bar{v}$. This completes the proof that $\bar{v}$ is a fixed point of T. Since $(V, \mathbb{T})$ is order stable (by global stability and Lemma 1.3.1) and regular, it follows that the fundamental optimality properties all hold (Corollary 1.2.7).

Regarding convergence of VFI, fix $v \in V_U$. Upward stability implies that $v \preceq v_{\top}$, where $v_{\top}$ is the value function (Lemma 1.2.11). Hence $v_n := T^n v$ is an increasing sequence dominated by $v_{\top}$. Moreover, letting σ be an optimal policy and setting $w_n := T_{\sigma}^n v$, we have $w_n \preceq v_n \preceq v_{\top}$ for all n and $w_n \to v_{\top}$. Applying Lemma 1.3.1 again, we have $\vee_n w_n = v_{\top}$ and hence $v_n \uparrow v_{\top}$. Since $(V, \mathbb{T})$ is regular and upward stable, convergence of VFI implies convergence of OPI and HPI (Corollary 1.2.10). □

1.3.1.2 Contracting ADPs

We conclude this section with a special case that handles many traditional dynamic programming problems. In the statement of the theorem, (V, d) is a metric space and $\preceq$ is a partial order on V. We understand $V = (V, d, \preceq)$ as a **partially ordered metric space**, by which we mean that d is a metric on V and $(V, \preceq)$ is a pospace under the topology generated by d. We use the notion of sup-nonexpansiveness introduced in §A.5.3.2.

Theorem 1.3.5. *Let $(V, \mathbb{T})$ be a regular ADP where $V = (V, d, \preceq)$ is a partially ordered metric space and d is complete and sup-nonexpansive. If each $T_{\sigma} \in \mathbb{T}$ is a contraction of modulus β on V, then*

(i) *the fundamental optimality properties hold and*

(ii) *VFI, OPI and HPI all converge.*

Proof. Let V and $\mathbb{T}$ have the stated properties. By completeness and Banach's fixed point theorem (page 173), the ADP $(V, \mathbb{T})$ is globally stable. Hence, by Theorem 1.3.3, we need only show that T has a fixed point in V. This follows from Banach's fixed point theorem and sup-nonexpansiveness of d (see Lemma A.5.19 on page 215). □

1.3.2 Example: Optimal Savings with Markov Income

Let's consider the optimal savings model from §1.2.4 again, but this time dropping the IID income assumption for income $(Y_t)_{t \geqslant 0}$ and assuming instead that that income is P-Markov on $\mathbb{R}_+$, where P is a stochastic kernel on $\mathbb{R}_+$.

Assumption 1.3.1. The utility function u is in $bc\mathbb{R}_+$ and the stochastic kernel P has continuous transition density p.

The last statement means that $P(y, B) = \int_B p(y, y') \, \mathrm{d}y'$ for all $y \in \mathbb{R}_+$ and all Borel sets B, where p is continuous on $\mathbb{R}^2_+$. The policy operator corresponding to $\sigma \in \Sigma$ becomes

$$(T_\sigma v)(w, y) = u(\sigma(w, y)) + \beta \int v(R(w - \sigma(w, y)) + y')p(y, y') \, \mathrm{d}y' \tag{1.26}$$

with $(w, y) \in \mathsf{X} := \mathbb{R}^2_+$. We let $\mathbb{T}_{\mathrm{OSM}}$ be the set of all such operators. We continue to assume that $\beta \in (0, 1)$. For the value space we take $V = bm\mathsf{X}$.

Exercise 1.3.2. By suitably modifying the proof of Lemma 1.1.2, show that

$$f(c, y) := u(c) + \beta \int v(R(w - c) + y')p(y, y') \, \mathrm{d}y' \qquad (0 \leqslant c \leqslant w)$$

is continuous for all $v \in V$ and $w, y \in \mathbb{R}_+$. Conclude that $(V, \mathbb{T}_{\mathrm{OSM}})$ is regular.

In the proof of the next proposition, we will use the fact that the supremum distance is sup-nonexpansive on V (see Example A.5.11 on page 215).

Proposition 1.3.6. *The Markov-income optimal savings ADP $(V, \mathbb{T}_{\mathrm{OSM}})$ obeys the fundamental optimality properties and VFI, OPI and HPI all converge.*

Proof. For fixed $\sigma \in \Sigma$ and $f, g \in V$ we have

$$|(T_\sigma f)(w, y) - (T_\sigma g)(w, y)| \leqslant \beta \int |f(h(w, y, y')) - g(h(w, y, y'))| p(y, y') \, \mathrm{d}y'$$

where $h(w, y, y') := R(w - \sigma(w, y)) + y'$. Hence

$$|(T_\sigma f)(w, y) - (T_\sigma g)(w, y)| \leqslant \beta \|f - g\|_\infty \quad \text{for all } (w, y) \in \mathsf{X}.$$

Taking the supremum over (w, y), we see that T_σ is a contraction of modulus β. Since $(V, \mathbb{T}_{\mathrm{OSM}})$ is regular, the claims in Proposition 1.2.17 follow from Theorem 1.3.5. □

1.3.3 ADPs in Banach Lattices

Our next step is to add algebraic structure to the problems that we study. We do this in the setting of Banach lattices, which are pospaces with algebraic structure. Banach lattices are particularly nice to work with, due to thier well-integrated order, algebraic and metric properties. Throughout this section, E is a Banach lattice with positive cone E_+.

1.3.3.1 Eventual Contractions on Banach Lattices

We call a self-map S on a subset V of E **eventually contracting** on V if there exists a positive linear operator $K: E \to E$ such that

(E1) $\rho(K) < 1$ and

(E2) $|S\,v - S\,w| \leqslant K|v - w|$ for all $v, w \in V$.

Eventually order contracting maps obey the following fixed point result.

Theorem 1.3.7. *If S is eventually contracting on V and V is closed in Banach lattice E, then S is globally stable on V.*

Proof. Let the stated conditions hold. We fix $n \in \mathbb{N}$ and use (E2) to obtain $|S^n v - S^n w| \leqslant K|S^{n-1}v - S^{n-1}w|$. Since K is order preserving, we can iterate on the last inequality to produce $|S^n v - S^n w| \leqslant K^n|v - w|$. Because E is a Banach lattice, its norm $\|\cdot\|$ is a lattice norm, which leads to

$$\|S^n v - S^n w\| \leqslant \|K^n|v - w|\| \leqslant \|K^n\|\|v - w\|.$$

Using Gelfand's formula $\lim_{n\to\infty} \|K^n\|^{1/n}$ for the spectral radius $\rho(K)$ and (E1), we obtain a $\lambda \in [0, 1)$ such that $\|K^n\| \leqslant \lambda^n$ for some $n \in \mathbb{N}$. Thus, S^n is a contraction on V. Since E is complete and V is a closed in E, a standard extension to the Banach fixed point theorem (see, e.g., Theorem A.2.10 on page 174) yields global stability of S on V. □

1.3.3.2 A Generalized Blackwell Condition

Next we study ADPs where the value space is a subset of a Banach lattice, as described in §A.5.3.3. The result in this section generalizes contraction-based conditions often used for dynamic programming in discounted environments (see, e.g., Puterman (2005), Bertsekas (2022) or Stokey and Lucas (1989)).

Let $(V, \mathbb{T})$ be an ADP where V is a subset of a Banach lattice E and $v + h \in V$ whenever $v \in V$ and $h \in E_+$. In this setting, We call $(V, \mathbb{T})$ **eventually Blackwell contracting** if V is a subset of E obeying $v + h \in V$ whenever $v \in V$ and $h \in E_+$, and, in addition, there exists a positive linear operator K on E such that

(C1) $\rho(K) < 1$ and

(C2) $T_\sigma(v + h) \leqslant T_\sigma v + Kh$ for all $T_\sigma \in \mathbb{T}$, $v \in V$ and $h \in E_+$.

This set of restricts generalizes a condition for contractivity often called "Blackwell's condition" that is popular in the literature on dynamic programming (see, e.g., Stokey and Lucas (1989), Theorem 3.3).

We can now state our main result for ADPs on Banach lattices.

Theorem 1.3.8. *If V is a closed subset of a Banach lattice and $(V, \mathbb{T})$ is regular and eventually Blackwell contracting, then*

(i) *the fundamental optimality properties hold, and*

(ii) *VFI, OPI and HPI all converge.*

Proof of Theorem 1.3.8. Let $(V, \mathbb{T})$ have the stated properties. Since V is a partially ordered space, it suffices to check the conditions of Theorem 1.3.3. Regarding global stability, fix $T_\sigma \in \mathbb{T}$. Given $v, w \in V$, we use (C2) to obtain

$$T_\sigma v = T_\sigma(w + v - w) \leqslant T_\sigma(w + |v - w|) \leqslant T_\sigma w + K|v - w|.$$

Rearranging gives $T_\sigma v - T_\sigma w \leqslant K|v - w|$. Reversing the roles of v and w yields $|T_\sigma v - T_\sigma w| \leqslant K|v - w|$. Hence T_σ is eventually contracting and therefore globally stable (by Theorem 1.3.7).

Regarding the Bellman operator, observe that

$$T_\sigma v = T_\sigma w + T_\sigma v - T_\sigma w \leqslant T_\sigma w + |T_\sigma v - T_\sigma w| \leqslant Tw + K|v - w|. \tag{1.27}$$

Taking the supremum over σ gives $Tv - Tw \leqslant K|v - w|$. Reversing the roles of v and w yields $|Tv - Tw| \leqslant K|v - w|$. Hence T is eventually contracting. In particular, T has a fixed point in V. This completes the proof of Theorem 1.3.8. □

1.3.3.3 Affine ADPs

In this section we examine affine ADPs, which include standard Markov decision processes and models with state-dependent discounting.

Theorem 1.3.9. *Let E be a Banach lattice and let $(E, \mathbb{T})$ be an affine ADP, where each $T_\sigma \in \mathbb{T}$ has the form*

$$T_\sigma \, v = r_\sigma + K_\sigma \, v \quad \text{for some } r_\sigma \in E \text{ and } K_\sigma \in \mathcal{B}_+(E),$$

Suppose that $(E, \mathbb{T})$ is regular. If either

(a) *there exists a $K \in \mathcal{B}(E)$ such that $K_\sigma \leqslant K$ for all $\sigma \in \Sigma$ and $\rho(K) < 1$, or*

(b) *E is σ-Dedekind complete, $(E, \mathbb{T})$ is bounded above and $\rho(K_\sigma) < 1$ for all $\sigma \in \Sigma$,*

then

(i) *the fundamental optimality properties hold, and*

(ii) *VFI, OPI and HPI all converge.*

Proof. Let $(E, \mathbb{T})$ be regular and affine, as in the statement of Theorem 1.3.9. Consider first case (a). We aim to check the conditions of Theorem 1.3.8. For this we need only show that (C1)–(C2) on page 32 hold. (C1) follows from $\rho(K) < 1$ and Gelfand's formula for the spectral radius (page 200). Regarding condition (C2), we fix $T_\sigma \in \mathbb{T}$ and observe that, for fixed $v \in E$ and $h \in E_+$, we have $T_\sigma(v+h) = r_\sigma + K_\sigma \, (v+h) \leqslant T_\sigma \, v + Kh$. Hence (C2) holds.

Now consider case (b). By assumption, $(E, \mathbb{T})$ is regular. Also, the conditions in (b) imply that $(E, \mathbb{T})$ is globally stable and bounded above. Claims (i) and (ii) now follow from Theorem 1.3.4. □

1.3.3.4 Concavity, Convexity, and Optimality

Some dynamic programs involve nonlinear policy operators that fail to be contractions. For example, models with recursive preferences or ambiguity often have these features. In this setting, we can deploy alternative point results related to concavity and convexity of operators. Here we apply such results to obtain optimality conditions for ADPs. Throughout, we take $E = (E, \leqslant)$ to be a Banach lattice and assume V has the form $V = [a, b]$ for some $a, b \in E$.

Theorem 1.3.10. *Let $(V, \mathbb{T})$ be an ADP where $V = [a, b]$ is containted in a σ-Dedekind complete Banach lattice and suppose that each $T_\sigma \in \mathbb{T}$ satisfies one of the following conditions:*

(a) *T_σ is concave and $T_\sigma a \geqslant a + \varepsilon(b - a)$ for some $\varepsilon \in (0, 1)$.*

(b) *T_σ is convex and $T_\sigma b \leqslant b - \varepsilon(b - a)$ for some $\varepsilon \in (0, 1)$.*

If, in addition, $(V, \mathbb{T})$ is regular, then

(i) *the fundamental optimality properties hold, and*

(ii) *VFI, OPI and HPI all converge.*

Proof. Let $(V, \mathbb{T})$ be as stated. Either of (a)–(b) in Theorem 1.3.10 is sufficient for global stability of T_σ on V by Du's Theorem (see, for example, Zhang (2012), Theorem 2.1.2). In particular, under (a)–(b), the ADP $(V, \mathbb{T})$ is globally stable. Since $(V, \mathbb{T})$ is also regular and V is bounded above, claims (i) and (ii) are implied by Theorem 1.3.4. □

When the positive cone of the Banach lattice has nonempty interior, there are sufficient conditions for (a)–(b) above that are sometimes convenient. To state them we write $x \ll y$ if $y - x$ is interior to E_+. For example, if $E = bc\mathsf{X}$ for some metric space X, then $0 \ll f$ if and only if $f(x) > 0$ for all $x \in \mathsf{X}$.

Corollary 1.3.11. *In Theorem 1.3.10, the alternatives* (a)–(b) *can be replaced by*

(a') *T_σ is concave and $a \ll T_\sigma a$ or*

(b') *T_σ is convex and $T_\sigma b \ll b$*

without changing the conclusions.

Proof. If (a') holds, then $T_\sigma a \geqslant a$ is interior to the positive cone E_+, so we can take a positive ε such that $T_\sigma a - a \geqslant \varepsilon(b - a)$. Hence (a) holds. The proof of (b) is similar. □

1.4 Additional Topics

Roadmap.

1.4.1 Nonstationary Policies

In all of the preceding discussion we focused on stationary policies. For example, in the context of the optimal savings problem from §1.1, we fixed a policy σ and computed its lifetime value v_σ by assuming that σ is applied at every t in $\{0, 1, \ldots\}$. In particular, in 1.1.1.2, we showed that, for v arbitrarily chosen from the value space,

$$v_\sigma = \lim_{n \to \infty} T_\sigma^n \, v. \tag{1.28}$$

This expression illustrates how lifetime value is obtained by repeatedly applying the same policy.

But is this focus on stationary policies justified? Could it be that higher lifetime value is available when we allow a change of policy in each period?

To address this question, suppose that we can select a **policy plan** $\bar\sigma := (\sigma_t)_{t \geqslant 0}$ in the infinite Cartesian product $\times_{t \geqslant 0} \Sigma$ and apply the t-th element σ_t at time t. Generalizing (1.28), the lifetime value of $\bar\sigma$ can be defined by

$$v_{\bar\sigma} = \lim_{n \to \infty} T_{\sigma_0} T_{\sigma_1} \cdots T_{\sigma_n} v. \tag{1.29}$$

Of course, for the definition in (1.29) to make sense we need to know that the limit exists. Ideally, it should also be independent of v. (In (1.29), we iterate backwards in time, applying T_{σ_j} first, because v is best thought of as a terminal condition, rather than an initial condition. See 1.1.1.2 for intuition.)

Since the expression (1.29) requires a topology, we consider an ADP $(V, \mathbb{T})$ where $V = (V, \preceq)$ is a partially ordered space. We also assume that the topology on V is generated by a metric d. As usual, $\mathbb{T} := \{T_\sigma : \sigma \in \Sigma\}$ is a family of order preserving self-maps on V. To ensure that (1.29) exists we also require the following:

Assumption 1.4.1. The metric d is complete and sup-nonexpansive (see §A.5.3.2). In addition, there exists a positive constant λ with $\lambda < 1$ and

$$d(T_\sigma \, v, T_\sigma \, w) \leqslant \lambda d(v, w) \quad \text{for all } v, w \in V \text{ and all } \sigma \in \Sigma.$$

In addition, for all $v \in V$ we have $\sup_{\sigma \in \Sigma} d(v, T_\sigma \, v) < \infty$.

We will make use of the following preliminary results.

Lemma 1.4.1. *If Assumption 1.4.1 holds, then*

(i) *for each $v \in V$ and policy plan $\hat\sigma := (\sigma_t)_{t \geqslant 0}$, the limit*

$$v_{\hat\sigma} := \lim_{n \to \infty} T_{\sigma_0} \cdots T_{\sigma_n} \, v$$

exists in V and is independent of v.

(ii) *Every $T_\sigma \in \mathbb{T}$ is continuous and globally stable on V, with unique fixed point v_σ satisfying*

$$v_\sigma = \lim_{j\to\infty} T_\sigma^j v \quad \textit{for all } v \in V. \tag{1.30}$$

(iii) *There exists a $v \in V$ such that $v \coloneqq \bigvee_{\sigma\in\Sigma} T_\sigma v$.*

Proof. Fix $v \in V$ and policy plan $\hat\sigma \coloneqq (\sigma_t)_{t\geqslant 0}$. Given the policy plan above and $m \leqslant n$, we adopt the following simplified notation:

$$T_{m,n} \coloneqq T_{\sigma_m} \circ T_{\sigma_{m+1}} \circ \cdots \circ T_{\sigma_n}.$$

Let $v_n = T_{0,n}v$. Our claim is that $\lim_n v_n$ exists in V and is independent of v. To see this, observe first that (v_n) is Cauchy, since, fixing $m, j \in \mathbb{N}$,

$$d(v_m, v_{m+j}) \leqslant \lambda^m d\left(v, T_{m+1,j}v\right),$$

and, by repeatedly applying the triangle inequality,

$$d(v, T_{m+1,j}v) \leqslant d(v, T_{m+1}v) + d(T_{m+1}v, T_{m+1}T_{m+2}v) \\ + \cdots + d(T_{m+1}\cdots T_{m+j-1}v, T_{m+1}\cdots T_{m+j-1}T_{m+j}v).$$

By Assumption 1.4.1, there exists a finite constant b satisfying $d(v, T_{\sigma_j}v) \leqslant b$ for all j. From this and the last bound we obtain

$$d(v, T_{m+1,j}v) \leqslant b + \lambda b + \cdots + \lambda^{j-m-1} b \leqslant \frac{b}{1-\lambda}.$$

As a result, $d(v_m, v_{m+j}) \leqslant \lambda^m b/(1-\lambda)$. This shows that (v_n) is Cauchy. Using completeness of V and letting $\bar v$ be the limit of this sequence, we argue that $\bar v$ is independent of v. Indeed, if $w_n \coloneqq T_{0,n}w$ for some $w \in V$, then $d(v_n, w_n) \leqslant \lambda^n d(v, w)$ for all n, so that (v_n) and (w_n) have the same limit. Hence $\lim_{n\to\infty} T_{0,n}v$ exists in V and is independent of the initial condition v. This proves claim (i).

The result in (ii) is immediate because, by Assumption 1.4.1, every $T_\sigma \in \mathbb{T}$ is a contraction mapping (and therefore continuous) on the complete metric space V. Finally, for (iii), applying Lemma A.5.19 on page 215, the Bellman operator T is also a contraction map and, therefore, has at least one fixed point in V. □

We can now prove the main result of this section, which shows that any policy plan is (weakly) dominated in value by a stationary policy.

Proposition 1.4.2. *If $(V, \mathbb{T})$ is regular and Assumption 1.4.1 holds, then the fundamental optimality properties hold. In addition, given any policy plan $\bar\sigma$, there exists a stationary policy plan σ such that $v_{\bar\sigma} \preceq v_\sigma$.*

Proof. Let the stated assumptions hold and let $\bar v$ be a fixed point of T in V, existence of which is guaranteed by Assumption 1.4.1. By Lemma 1.4.1, the ADP $(V, \mathbb{T})$ is globally stable and T has a fixed point in V. Since $(V, \mathbb{T})$ is also regular, the fundamental optimality properties hold (by Theorem 1.3.3).

Now let $\bar\sigma := (\sigma_t)_{t \geqslant 0}$ be an arbitrary policy plan and let σ be an optimal policy. Since v_σ is fixed for T, we have $\times_{t=0}^{j} T_{\sigma_t} v_\sigma \preceq T^j v_\sigma = v_\sigma$ for all j. Since Assumption 1.4.1 holds and the partial order is closed, taking the limit in j yields $v_{\bar\sigma} \preceq v_\sigma$. This proves that every policy plan is weakly dominated by a stationary continuation plan. □

1.4.2 Minimization Problems

In some dynamic programs, the objective is to minimize lifetime cost of a given policy, rather than maximizing rewards. While we focus primarily on maximization in this book, the present section discusses how to handle minimization problems. The content of this section can be summarized by the following statement: for a given ADP $(V, \mathbb{T})$, a minimization problem can be converted to a maximization problem by reversing the partial order on V. Further details are given below. Readers who prefer to focus on maximization results can safely skip ahead.

1.4.2.1 Definitions

Let $(V, \mathbb{T})$ be an ADP with policy set Σ. We define the **Bellman min-operator** $T_\perp$ corresponding to $(V, \mathbb{T})$ as by $T_\perp v = \bigwedge_\sigma T_\sigma v$ whenever the infimum exists.

Paralleling the maximization terminology, we say that

- $\sigma \in \Sigma$ is **v-min-greedy** if $T_\sigma v \preceq T_\tau v$ for all $\tau \in \Sigma$,
- $(V, \mathbb{T})$ is **min-regular** if, for each $v \in V$, at least one v-min-greedy policy exists, and
- v satisfies the **Bellman min-equation** if $T_\perp v = v$.

Now suppose $(V, \mathbb{T})$ is well-posed and let $V_\Sigma := \{v_\sigma\}$ be defined as before (i.e., the set of σ-value functions). In this setting we set $v_\perp = \bigwedge_\sigma v_\sigma$ and call it the **min-value function** whenever the infimum exists. Also, we say that

- $\sigma \in \Sigma$ is **min-optimal** for $(V, \mathbb{T})$ if $v_\sigma = v_\perp$, and
- $(V, \mathbb{T})$ obeys **Bellman's principle of min-optimality** if

$$\sigma \in \Sigma \text{ is min-optimal for } (V, \mathbb{T}) \quad \iff \quad \sigma \text{ is } v_\perp\text{-min-greedy.}$$

When $(V, \mathbb{T})$ is min-regular and well-posed, we define the **Howard policy min-operator** corresponding to $(V, \mathbb{T})$ via

$$H_\perp \colon V_G \to V_\Sigma, \qquad H_\perp v = v_\sigma \quad \text{where } \sigma \text{ is } v\text{-min-greedy,}$$

as well as, for each $m \in \mathbb{N}$, the **optimistic policy min-operator** via

$$W_\perp \colon V_G \to V, \qquad W_\perp v \coloneqq T_\sigma^m v \quad \text{where} \quad \sigma \text{ is } v\text{-min-greedy.} \tag{1.31}$$

We say that the **fundamental min-optimality properties hold** if

(B1') at least one min-optimal policy exists,

(B2') $v_\perp$ is the unique solution to the Bellman min-equation in V, and

(B3') Bellman's principle of min-optimality holds.

This definition parallels (B1)–(B3) from page 17.

Let V_D be all $v \in V$ with $T_\perp v \preceq v$. We say that

- **min-VFI converges** if $T_\perp^n v \downarrow v_\perp$ for all $v \in V_D$,
- **min-OPI converges** if $W_\perp^n v \downarrow v_\perp$ for all $v \in V_D$ and all $m \in \mathbb{N}$, and
- **min-HPI converges** if $H_\perp^n v \downarrow v_\perp$ for all $v \in V_D$.

To further increase clarity, when discussing maximization and minimization in the same section, we add a "max-" prefix to the previously introduced definitions that pertain to maximization. For example,

- "v-greedy policies" will be referred to as **v-max-greedy policies**,
- "optimal policies" will be referred to as **max-optimal policies**,
- "the Bellman equation" will be referred to as the **Bellman max-equation**,

and so on.

1.4.2.2 Dual ADPs

Let's now investigate how minimization problems can be converted to maximization problems in this abstract setting.

Recall from §A.1.2.5 that if $V := (V, \preceq)$ is a partially ordered set, then its order dual $V^\partial := (V, \preceq^\partial)$ is the partially ordered set obtained by setting $u \preceq^\partial v$ if and only if $v \preceq u$. If $(V, \mathbb{T})$ is any ADP, then we call $(V, \mathbb{T})^\partial := (V^\partial, \mathbb{T})$ the **dual** of $(V, \mathbb{T})$. In other words, the dual $(V, \mathbb{T})^\partial$ of $(V, \mathbb{T})$ is the ADP created by replacing the poset V with its order dual V^∂.

Exercise 1.4.1. Show that $(V, \mathbb{T})^\partial$ is an ADP.

Regarding notation for $(V, \mathbb{T})^\partial$,

- the Bellman max-operator will be denoted by T^∂,
- the Bellman min-operator will be denoted by $T^\partial_\perp$,
- the max-value function will be denoted by $v^\partial_\top$,
- etc.

Each ADP is self-dual, in the sense that $((V, \mathbb{T})^\partial)^\partial = (V, \mathbb{T})$. This follows from the fact that all partially ordered sets are order self-dual.

Exercise 1.4.2. Let $(V, \mathbb{T})$ be a well-posed ADP with dual $(V, \mathbb{T})^\partial$. Fix $v \in V$ and verify the following:

(i) σ is v-min-greedy for $(V, \mathbb{T})$ if and only if σ is v-max-greedy for $(V, \mathbb{T})^\partial$.

(ii) $(V, \mathbb{T})$ is min-regular if and only if $(V, \mathbb{T})^\partial$ is max-regular,

(iii) If $T^\partial v$ exists then so does $T_\perp v$, and, moreover, $T_\perp v = T^\partial v$.

(iv) If $W^\partial v$ exists then so does $W_\perp v$, and, moreover, $W_\perp v = W^\partial v$.

(v) If $H^\partial v$ exists then so does $H_\perp v$, and, moreover, $H_\perp v = H^\partial v$.

(vi) If the max-value function $v^\partial_\top$ exists for $(V, \mathbb{T})^\partial$, then the min-value function $v_\perp$ exists for $(V, \mathbb{T})$ and, moreover, $v_\perp = v^\partial_\top$.

(vii) $\sigma \in \Sigma$ is min-optimal for $(V, \mathbb{T})$ if and only if σ is max-optimal for $(V, \mathbb{T})^\partial$.

Self-duality implies corollaries to Exercise 1.4.2 that we treat as self-evident. For example, $\sigma \in \Sigma$ is min-optimal for $(V, \mathbb{T})^\partial$ if and only if σ is max-optimal for $(V, \mathbb{T})$.

Part (v) of Exercise 1.4.2 tells us that we can solve an ADP for a min-optimal policy by switching to the dual ADP and maximizing.

1.4.2.3 Optimality and Convergence

The next theorem links the fundamental max-optimality properties on on page 17 to analogous min-optimality properties for the dual. In the statement of the theorem, $(V, \mathbb{T})$ is a well-posed ADP with dual $(V, \mathbb{T})^\partial$.

Theorem 1.4.3. *The fundamental max-optimality properties hold for $(V, \mathbb{T})^\partial$ if and only if the fundamental min-optimality properties hold for $(V, \mathbb{T})$. Moreover,*

(i) *max-VFI converges for $(V, \mathbb{T})^\partial$ if and only if min-VFI converges for $(V, \mathbb{T})$,*

(ii) *max-OPI converges for $(V, \mathbb{T})^\partial$ if and only if min-OPI converges for $(V, \mathbb{T})$, and*

(iii) *max-HPI converges for $(V, \mathbb{T})^\partial$ if and only if min-HPI converges for $(V, \mathbb{T})$,*

The proof of Theorem 1.4.3 is left as an exercise.

1.5 Chapter Notes

The introduction to this chapter discusses applications of dynamic programming to atemporal problems, such as genome sequencing and the structure of production chains. For one discussion of the former see Gu et al. (2023); for the latter see, for example, Kikuchi et al. (2021). We mentioned also that many recent applications of dynamic programming are connected to machine learning and artificial intelligence. Introductions to the literature can be found in Bertsekas (2021) and Kochenderfer et al. (2022).

On the theory side, this chapter builds on the abstract dynamic programming framework developed by Denardo (1967), Bertsekas (1977), Verdu and Poor (1987), Szepesvari (1998), Bertsekas (2017), Li and Xie (2021), and, in particular, Bertsekas (2022). We have added an additional layer of abstraction over these earlier frameworks by shifting analysis to families of "policy operators" on partially ordered sets, rather that using the more specific operators on function spaces used in Bertsekas (2022). Motivation for this approach is described in Chapter 9 of Sargent and Stachurski (2025a) and the research paper Sargent and Stachurski (2025b). In essence, the benefit is that we can treat a wider class of problems and simplify some proofs. At the same time, we emphasize that Bertsekas (2022) inspired most of the research behind this book and contains many valuable results that are not treated here.

We make extensive use of order theory, which is a powerful and perhaps underutilized branch of mathematics. An introduction to the field can be found in Davey and Priestley (2002). High quality monographs on Riesz spaces, Banach lattices and positive operators include Aliprantis and Border (2006), Aliprantis and Burkinshaw (2006), Zaanen (2012), Meyer-Nieberg (2012), and Bátkai et al. (2017).

Chapter 2

Classes of Dynamic Programs

Roadmap.

2.1 Markov Decision Processes

Roadmap.

2.1.1 Finite State MDPs

Finite state Markov decision processes (MDPs) form the foundations of a vast range of quantitative modeling and reinforcement learning routines, as well as providing a benchmark setting for dynamic programming theory (see, e.g., Puterman (2005) or Chapter 5 of Sargent and Stachurski (2025a)). In this section we show that these MDPs are a special case of ADPs and discuss their optimality.

2.1.1.1 Model

We consider a **Markov decision process** where the objective is to maximize

$$\mathbb{E} \sum_{t \geqslant 0} \beta^t r(X_t, A_t) \quad \text{s.t.} \quad A_t \in \Gamma(X_t) \text{ for all } t \geqslant 0.$$

Here X_t takes values in finite set X (the state space), A_t takes values in finite set A (the action space), $\beta \in (0,1)$ is a discount factor, and r is a reward function. The action sequence (A_t) must also satisfy an information constraint: Each A_t is required to be measurable with respect to the σ-algebra generated by $(X_0, \ldots, X_t)$.

The feasible actions are defined by the nonempty correspondence Γ from X to A. Setting

$$\mathsf{G} := \{(x, a) \in \mathsf{X} \times \mathsf{A} : a \in \Gamma(x)\},$$

we assume $P \colon \mathsf{G} \times \mathsf{X} \to [0, 1]$ provides transition probabilities for the next period state given current state and action, so that $\sum_{x'} P(x, a, x') = 1$ for all $(x, a) \in \mathsf{G}$. The set of feasible policies is

$$\Sigma := \{\sigma \in \mathsf{A}^{\mathsf{X}} : \sigma(x) \in \Gamma(x) \text{ for all } x \in \mathsf{X}\}.$$

Given $\sigma \in \Sigma$, we set

$$P_\sigma(x, x') := P(x, \sigma(x), x') \quad \text{and} \quad r_\sigma(x) := r(x, \sigma(x)). \tag{2.1}$$

It follows from our assumptions on P that P_σ is a **stochastic matrix**, meaning that $P_\sigma \geqslant 0$ and $P_\sigma \mathbb{1} = \mathbb{1}$. Alternatively, P_σ is a Markov operator on $\mathbb{R}^{\mathsf{X}}$ (see §A.5.4). Thus, by choosing a policy σ, the controller determines a reward function r_σ on the state and Markov dynamics P_σ for the state process.

The lifetime value of σ given $X_0 = x$ is defined to be $v_\sigma(x) := \mathbb{E} \sum_{t \geqslant 0} \beta^t r(X_t, \sigma(X_t))$, where $(X_t)_{t \geqslant 0}$ is a Markov chain generated by P_σ with initial condition $X_0 = x \in \mathsf{X}$. Pointwise on X, we can express v_σ as

$$v_\sigma = \sum_{t \geqslant 0} (\beta P_\sigma)^t r_\sigma = (I - \beta P_\sigma)^{-1} r_\sigma, \tag{2.2}$$

where I is the identity map on $\mathbb{R}^{\mathsf{X}}$, the set of real-valued functions on X (see, e.g., Puterman (2005), Theorem 6.1.1, or Sargent and Stachurski (2025a), Ch. 5). The MDP optimality problem is to find a $\sigma \in \Sigma$ such that $v_\tau \leqslant v_\sigma$ for all $\tau \in \Sigma$.

Remark 2.1.1. When considering optimality we could also consider broader classes of policies, such as mappings that depend on the whole history of the processes, or policies that randomize across actions. However, none of these alternatives lead to strictly greater lifetime value. (See, e.g., Sargent and Stachurski (2025a), Section 9.2.1.6.) This is why we maintain focus on stationary Markov policies, where the same deterministic policy σ is applied at every point in time (leading to the valuation in (2.2)).

2.1.1.2 MDPs as ADPs

We can represent the MDP described in the previous section as an ADP. Fixing $\sigma \in \Sigma$, we define

$$(T_\sigma v)(x) = r(x, \sigma(x)) + \beta \sum_{x'} v(x')P(x, \sigma(x), x') \qquad (v \in \mathbb{R}^{\mathsf{X}},\ x \in \mathsf{X}). \tag{2.3}$$

In operator notation this is $T_\sigma v = r_\sigma + \beta P_\sigma v$. We set $\mathbb{T}_{\mathrm{MDP}} := \{T_\sigma : \sigma \in \Sigma\}$ and pair $\mathbb{R}^{\mathsf{X}}$ with the pointwise partial order $\leqslant$. Clearly each $T_\sigma \in \mathbb{T}_{\mathrm{MDP}}$ is an order preserving on $\mathbb{R}^{\mathsf{X}}$. Hence $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\mathrm{MDP}})$ is an ADP.

Also, since $\rho(P_\sigma) = 1$ (see Lemma A.5.29 on page 222) and hence $\rho(\beta P_\sigma) = \beta < 1$, each $T_\sigma \in \mathbb{T}_{\mathrm{MDP}}$ has a unique fixed point in $\mathbb{R}^{\mathsf{X}}$ given by $(I - \beta P)^{-1} r_\sigma$ (see Corollary A.4.11). Thus, $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\mathrm{MDP}})$ is well-posed. As expected, the fixed points of the policy operators represent lifetime values (see (2.2)).

EXERCISE 2.1.1. Show that $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\mathrm{MDP}})$ is order stable.

EXERCISE 2.1.2. Fix $v \in \mathbb{R}^{\mathsf{X}}$. Prove the following: If $\sigma \in \Sigma$ obeys

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} \left\{ r(x, a) + \beta \sum_{x'} v(x')P(x, a, x') \right\} \quad \text{for all } x \in \mathsf{X}, \tag{2.4}$$

then σ is v-greedy; that is, that $T_\sigma v \geqslant T_\tau v$ for all $\tau \in \Sigma$.

EXERCISE 2.1.3. Recall that the ADP Bellman operator is defined by the expression $Tv = \bigvee_\sigma T_\sigma v$. Show that, in the present setting, this can also be written as

$$(T\, v)(x) = \max_{a \in \Gamma(x)} \left\{ r(x, a) + \beta \sum_{x'} v(x')P(x, a, x') \right\} \qquad (x \in \mathsf{X}). \tag{2.5}$$

From (2.5) it follows that the Bellman equation for $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\mathrm{MDP}})$ is given by

$$v(x) = \max_{a \in \Gamma(x)} \left\{ r(x, a) + \beta \sum_{x'} v(x')P(x, a, x') \right\} \qquad (x \in \mathsf{X}). \tag{2.7}$$

This is the standard expression for the Bellman equation of the MDP we are considering (see, e.g., Puterman (2005) or Sargent and Stachurski (2025a)).

The next exercise shows that we can replace the value space $\mathbb{R}^{\mathsf{X}}$ with a smaller set if we wish.

EXERCISE 2.1.4. Letting

$$M := \max_{(x,a)\in \mathsf{G}} |r(x,a)| \quad \text{and} \quad \hat{V} := \left\{ v \in \mathbb{R}^{\mathsf{X}} \;:\; |v| \leqslant \frac{M}{1-\beta} \right\},$$

show that every T_σ is a self-map on $\hat{V}$. Show also that $v_\sigma \in \hat{V}$ for all $\sigma \in \Sigma$.

2.1.1.3 Properties

The ADP $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\mathrm{MDP}})$ is regular (i.e., greedy policies always exist). This fact is easily verified from finiteness of A, nonemptyness of Γ, and the characterization of greedy policies given in Exercise 2.1.2.

The ADP $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\mathrm{MDP}})$ is also globally stable, since each policy operator $T_\sigma \in \mathbb{T}_{\mathrm{MDP}}$ has the form $T_\sigma\, v = r_\sigma + \beta P_\sigma\, v$, and every such map is a contraction map on $\mathbb{R}^{\mathsf{X}}$ with respect to the supremum distance. Indeed, with $\|\cdot\|$ as the supremum norm and $v, v' \in \mathbb{R}^{\mathsf{X}}$, we have

$$\|T_\sigma\, v - T_\sigma\, v'\| = \beta\|P_\sigma\,(v - v')\| \leqslant \beta\|P_\sigma\|\,\|v - v'\| = \beta\|v - v'\|,$$

(In the last step we used $\|P_\sigma\| = 1$ from Lemma A.5.27 on page 221.) The set $\mathbb{R}^{\mathsf{X}}$ is closed in the complete space $(\mathbb{R}^{\mathsf{X}}, \|\cdot\|)$, so T_σ is globally stable on $\mathbb{R}^{\mathsf{X}}$.

Since each T_σ is a contraction map of modulus β on $\mathbb{R}^{\mathsf{X}}$ and the metric $d_\infty(v,w) = \|v - w\|$ is sup-nonexpansive, the discussion in §A.5.3.2 implies that the Bellman operator T is also a contraction of modulus β on $(\mathbb{R}^{\mathsf{X}}, d_\infty)$.

EXERCISE 2.1.5. Supply a more direct proof that T is a contraction of modulus β on $(\mathbb{R}^{\mathsf{X}}, d_\infty)$ using the expression for T in (2.5) and Lemma A.1.2.

2.1.1.4 MDP Optimality

Finite MDPs have strong optimality properties. Standard proofs can be found in Puterman (2005), Bäuerle and Rieder (2011), Hernández-Lerma and Lasserre (2012), Stachurski (2022), or Sargent and Stachurski (2025a). We can also derive the key results from ADP theory:

Proposition 2.1.1. *For the ADP* $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\mathrm{MDP}})$,

(i) *the fundamental optimality properties hold,*

(ii) *VFI, OPI and HPI all converge, and*

(iii) *and HPI converges in finitely many steps.*

Proof. We saw in §2.1.1.3 that $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\mathrm{MDP}})$ is regular, and that each $T_\sigma \in \mathbb{T}_{\mathrm{MDP}}$ is a contraction of modulus β on the complete and sup-nonexpansive space $(\mathbb{R}^{\mathsf{X}}, \leqslant, d_\infty)$. Hence Theorem 1.3.5 applies. This yields (i) and (ii). Part (iii) follows from Theorem 1.2.12 (since Σ is finite). □

The statement in Proposition 2.1.1 is easily translated into more standard MDP terminology. For example, it tells us that the value function $v_\top$ solves the Bellman equation, which we know is given by (2.7), and, by the characterization of greedy policies in (2.4) plus Bellman's principle of optimality, that a policy σ is optimal if and only if

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} \left\{ r(x, a) + \beta \sum_{x'} v_\top(x') P(x, a, x') \right\} \quad \text{for all } x \in \mathsf{X}.$$

Later we will show that ADP theory can also handle many extensions to the basic MDP framework.

As a class of dynamic programs, finite-state MDPs are very straightforward to work with. Because of this, almost all of the optimality results in Chapter 1 to can be used to prove (i) and (ii) in Proposition 2.1.1. The exercises ask you to verify this for some cases.

Exercise 2.1.6. Prove (i) and (ii) of Proposition 2.1.1 using Theorem 1.3.4.

Exercise 2.1.7. Prove (i) and (ii) of Proposition 2.1.1 using Theorem 1.3.9.

Exercise 2.1.8. Let $\hat{V}$ be as in Exercise 2.1.4. Using Theorem 1.2.13, show that the ADP $(\hat{V}, \mathbb{T}_{\mathrm{MDP}})$ obeys the fundamental optimality properties and VFI, OPI, and HPI all converge.

2.1.2 Finite MDPs in Continuous Time

In this section we modify the finite state MDP model from §2.1.1 to the continuous time setting and study optimality. For now we maintain the finite state setting.

2.1.2.1 Primitives and Values

As in the discrete time case, X and A are finite sets, while the controller is contstrainted by a feasible correspondence Γ from X to A. The definitions of G, Σ, and r are unchanged. Discounting is determined by a constant $\delta > 0$, referred to as the **discount rate**, while transitions are driven by an **intensity kernel** Q from G to X, which is a map Q from $\mathsf{G} \times \mathsf{X}$ to $\mathbb{R}$ satisfying

$$\sum_{x'} Q(x, a, x') = 0 \text{ for all } (x, a) \text{ in } \mathsf{G} \text{ and } Q(x, a, x') \geqslant 0 \text{ when } x \neq x'.$$

Informally, over the short interval from t to $t + h$, the controller receives instantaneous reward $r(x, a)h$ and the state transitions to state x' with probability $Q(x, a, x')h + o(h)$.

For any fixed $\sigma \in \Sigma$, we obtain an intensity operator (i.e., infinitesimal generator)

$$Q_\sigma(x, x') := Q(x, \sigma(x), x') \qquad (x, x' \in \mathsf{X})$$

that determines a continuous time Markov chain $(X_t)_{t \geqslant 0}$ with transition probabilities given by $P_t^\sigma := \mathrm{e}^{tQ_\sigma}$ for all $x \in \mathsf{X}$. In particular,

$$\mathbb{E}_x h(X_t) = (P_t^\sigma h)(x) \text{ for any } h \in \mathbb{R}^{\mathsf{X}}.$$

(For background see Chapter 10 of Sargent and Stachurski (2025a).) Continuing to define $r_\sigma(x) := r(x, \sigma(x))$, the lifetime value of following σ starting from state x is

$$v_\sigma(x) = \mathbb{E}_x \int_0^\infty \mathrm{e}^{-\delta t} r_\sigma(X_t) \, \mathrm{d}t = \int_0^\infty \mathrm{e}^{-\delta t} (P_t^\sigma r_\sigma)(x) \, \mathrm{d}t \tag{2.8}$$

(Passing the expectation through the integral can be justified by Fubini's theorem.) Using $\delta > 0$, we can rewrite v_σ as

$$v_\sigma = \int_0^\infty \mathrm{e}^{t(Q_\sigma - \delta I)} r_\sigma \, \mathrm{d}t = (\delta I - Q_\sigma)^{-1} r_\sigma. \tag{2.9}$$

The two representations for v_σ are the continuous time analogs of the discrete-time representations given in (2.2). A proof of the second equality is given in §10.2 of Sargent and Stachurski (2025a).

(Readers familiar with semigroup theory will recognize the two representations in (2.9) as alternative expressions for the resolvent of the semigroup (e^{tQ}) – see, for example, Engel and Nagel (2006), Theorem 1.10.)

2.1.2.2 Reformulation

We can treat this continuous-time problem from the ADP framework by reformulating (2.9) in a way that realizes v_σ as the fixed point of an order preserving policy operator. One way to do this in the finite state setting is by defining

$$P(x, a, x') := \mathbb{1}\{x = x'\} + \frac{Q(x, a, x')}{m} \quad \text{where} \quad m := \max_{x \in \mathsf{X},\, a \in \mathsf{A}} |Q(x, a, x)|.$$

In addition, we set

$$\beta := \frac{m}{m + \delta} \quad \text{and} \quad \hat{r}_\sigma := \frac{r_\sigma}{m + \delta}.$$

Analogous to the discrete time case, we also define, for each $\sigma \in \Sigma$, P_σ and $\hat{r}_\sigma$ via

$$P_\sigma(x, x') := P(x, \sigma(x), x') \quad \text{and} \quad \hat{r}_\sigma(x) = \hat{r}(x, \sigma(x)).$$

Exercise 2.1.9. Prove that, with these definitions,

(i) P_σ is a stochastic matrix and

(ii) the σ-value function v_σ obeys

$$v_\sigma = (I - \beta P_\sigma)^{-1} \hat{r}_\sigma. \tag{2.10}$$

From (2.10), we see that v_σ is the unique fixed point in $V := \mathbb{R}^{\mathsf{X}}$ of the policy operator

$$T_\sigma \, v = \hat{r}_\sigma + \beta P_\sigma \, v. \tag{2.11}$$

Letting $\mathbb{T}_{\text{CTMDP}} := \{T_\sigma : \sigma \in \Sigma\}$, we study the ADP $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\text{CTMDP}})$.

2.1.2.3 Optimality

Since (2.11) is the same as (2.3) after replacing r with $\hat{r}$ and P_σ with $\hat{P}_\sigma$, we can apply the discrete time MDP theory in §2.1.1. We see that $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\text{CTMDP}})$ is a well-posed ADP with Bellman equation given by

$$v(x) = \max_{a \in \Gamma(x)} \left\{ \hat{r}(x, a) + \beta \sum_{x'} v(x') P(x, a, x') \right\} \qquad (x \in \mathsf{X}). \tag{2.12}$$

Moreover, the fundamental optimality properties hold and VFI, OPI and HPI converge (Proposition 2.1.1). In particular, the unique solution $v_\top$ of the Bellman equation in

$\mathbb{R}^{\mathsf{X}}$ is the greatest element of V_Σ and a policy is optimal if and only if

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} \left\{ \hat{r}(x,a) + \beta \sum_{x'} v_\top(x') P(x,a,x') \right\} \quad \text{for all } x \in \mathsf{X},$$

Exercise 2.1.10. Show that $v \in \mathbb{R}^{\mathsf{X}}$ obeys (2.12) if and only if

$$\delta v(x) = \max_{a \in \Gamma(x)} \left\{ r(x,a) + \sum_{x'} v(x') Q(x,a,x') \right\} \qquad (x \in \mathsf{X}). \tag{2.13}$$

Equation (2.13) connects the exposition above to the traditional theory of continuous time MDPs (see, e.g., Guo and Hernández-Lerma (2009)). It is sometimes called the **Hamilton–Jacobi–Bellman** (**HJB**) equation, although that name is more commonly applied in the case where the state process is a diffusion.

Analogous to Exercise 2.1.10, it can be shown that $\sigma \in \Sigma$ is $v_\top$-greedy if and only if

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} \left\{ r(x,a) + \sum_{x'} v_\top(x') Q(x,a,x') \right\} \quad \text{for all } x \in \mathsf{X}.$$

2.1.3 Weak Feller MDPs

In §2.1.1 we introduced the MDP model with finite states and actions. Here we drop the finiteness assumption, taking the state space X and the action space A to be metric spaces. At the same time, we restrict attention to a plain vanilla version of the model in order to make the analysis relatively straightforward. Later we will consider additions and extensions.

2.1.3.1 Theory

Let X and A be separable metric spaces. We replace the finite-state policy operator (2.3) on page 45 with

$$(T_\sigma v)(x) = r(x, \sigma(x)) + \beta \int v(x') P(x, \sigma(x), \mathrm{d}x') \qquad (x \in \mathsf{X}). \tag{2.14}$$

Actions are restricted by a correspondence Γ from X to A. A feasible policy is a Borel measurable map $\sigma \colon \mathsf{X} \to \mathsf{A}$ such that $\sigma(x) \in \Gamma(x)$ for all $x \in \mathsf{X}$. Let $\mathsf{G} := \{(x,a) \in \mathsf{X} \times \mathsf{A} :$

$a \in \Gamma(x)\}$. We understand G as the set of feasible state-action pairs. We let Σ denote the set of all feasible policies. In addition,

- r is a reward function mapping G, the set of feasible state-action pairs, into $\mathbb{R}$, and
- P is a stochastic kernel (see §A.5.4.1) from G to X.

We impose the following boundedness and continuity conditions.

Assumption 2.1.1. The discount factor β takes values in $(0, 1)$ and Γ is nonempty, continuous and compact-valued. In addition, r is bounded and continuous, while P is weak Feller (see §A.5.4.5).

Example 2.1.1. Suppose that the state evolves according to

$$X_{t+1} = F(X_t, A_t, W_{t+1}) \quad \text{with} \quad (W_t)_{t \geqslant 0} \overset{\text{IID}}{\sim} \varphi \in \mathscr{D}(\mathsf{W}),$$

where W is a metric space and $F \colon \mathsf{G} \times \mathsf{W} \to \mathsf{X}$ is Borel measurable.

$$\int v(x')P(x, a, \mathrm{d}x') = \int v(F(x, a, w))\varphi(\mathrm{d}w)$$

As show in Example A.5.19, the kernel P is weak Feller whenever $(x, a) \mapsto F(x, a, w)$ is continuous for all $w \in \mathsf{W}$.

Assumption 2.1.1 is in force for the rest of §2.1.3.1. We construct the ADP for this model by selecting $bm\mathsf{X}$, the bounded Borel measurable functions on X, as the value space, and setting $\mathbb{T}_{\mathrm{WF}} := \{T_\sigma : \sigma \in \Sigma\}$, with each T_σ as defined in (2.14). Throughout this section,

- $\leqslant$ is the pointwise partial order on $bm\mathsf{X}$ and
- $\| \cdot \|$ is the supremum norm.

Exercise 2.1.11. Show that each policy operator T_σ maps $bm\mathsf{X}$ into itself.

Each T_σ is order preserving on $bm\mathsf{X}$ under $\leqslant$ so $(bm\mathsf{X}, \mathbb{T}_{\mathrm{WF}})$ is an ADP.

Lemma 2.1.2. *Every $T_\sigma \in \mathbb{T}_{\mathrm{WF}}$ is a contraction of modulus β on $bm\mathsf{X}$.*

Proof. Fixing $T_\sigma \in \mathbb{T}_{\mathrm{WF}}$, $x \in \mathsf{X}$, and a pair v, w in $bm\mathsf{X}$, we have

$$\begin{aligned} |(T_\sigma v)(x) - (T_\sigma w)(x)| &= \beta \left| \int v(x')P(x, \sigma(x), \mathrm{d}x') - \int w(x')P(x, \sigma(x), \mathrm{d}x') \right| \\ &\leqslant \int \beta \, |v(x') - w(x')| \, P(x, \sigma(x), \mathrm{d}x') \leqslant \beta \|v - w\|. \end{aligned}$$

Taking the supremum over all $x \in \mathsf{X}$ completes the proof. □

Lemma 2.1.3. *If $v \in bc\mathsf{X}$ then there exists a $\sigma \in \Sigma$ such that*

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} \left\{ r(x,a) + \beta \int v(x') P(x,a,\mathrm{d}x') \right\} \quad \textit{for all } x \in \mathsf{X}. \tag{2.15}$$

Proof. By Assumption 2.1.1, the map

$$\mathsf{G} \ni (x,a) \mapsto r(x,a) + \beta \int v(x') P(x,a,\mathrm{d}x') \in \mathbb{R}$$

is continuous. The claim now follows from Theorem A.3.3 on page 179. □

Exercise 2.1.12. Using Lemma 2.1.3, prove that, given $v \in bc\mathsf{X}$, a policy $\sigma \in \Sigma$ is v-greedy if and only if it satisfies (2.15).

Together, Lemma 2.1.3 and Exercise 2.1.12 imply that there exists a v-greedy policy for every $v \in bc\mathsf{X}$. At the same time, the ADP is *not* guaranteed to be regular without additional assumptions. The reason is that Lemma 2.1.3 uses continuity of v, which does not hold for every v in the value space $bm\mathsf{X}$.

Remark 2.1.2. Why not restrict the value space to $bc\mathsf{X}$, rather than $bm\mathsf{X}$, to obtain regularity? We could try to construct an argument along these lines but the path is not straightforward. The reason is that feasible policies are only required to be measurable, rather than continuous, so T_σ might not map $bc\mathsf{X}$ to itself.

Next we characterize the Bellman operator $T := \bigvee_\sigma T_\sigma$ generated by $(bm\mathsf{X}, \mathbb{T}_{\mathrm{WF}})$. Since $T_\sigma v = Tv$ whenever σ is v-greedy (Lemma 1.2.1), Lemma 2.1.3 and Exercise 2.1.12 imply that, for $v \in bc\mathsf{X}$, we have

$$(Tv)(x) = \max_{a \in \Gamma(x)} \left\{ r(x,a) + \beta \int v(x') P(x,a,\mathrm{d}x') \right\} \qquad (x \in \mathsf{X}). \tag{2.16}$$

(T is well-defined on $bc\mathsf{X}$ but not necessarily on $bm\mathsf{X}$. See Remark 2.1.2.)

Exercise 2.1.13. Prove that T is a self-map on $bc\mathsf{X}$ and a contraction of modulus β.

We can now prove the following

Proposition 2.1.4. *If Assumption 2.1.1 holds, then, for the ADP $(bm\mathsf{X}, \mathbb{T}_{\mathrm{WF}})$,*

(i) *the Bellman operator* T *has a unique fixed point* $v_\top$ *in* $bc\mathsf{X}$,

(ii) *a policy* σ *is optimal if and only if*

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} \left\{ r(x,a) + \beta \int v_\top(x') P(x,a,\mathrm{d}x') \right\} \quad \textit{for all } x \in \mathsf{X},$$

(iii) *at least one optimal policy exists, and*

(iv) *for each* $v \in bc\mathsf{X}$, *we have* $d(T^n v, v_\top) \to 0$ *as* $n \to \infty$.

Proof. The ADP $(bm\mathsf{X}, \mathbb{T}_{\mathrm{WF}})$ is globally stable (Lemma 2.1.2) and Exercise 2.1.13 implies that T has a fixed point in $bc\mathsf{X}$. Exercise 2.1.12 shows that every $v \in bc\mathsf{X}$ has at least one greedy policy. Hence Theorem 1.3.2 applies (with $V = bm\mathsf{X}$ and $V_0 = bc\mathsf{X}$). The results in Proposition 2.1.4 now follow from the conclusions of Theorem 1.3.2 (as well as Exercise 2.1.12, which characterizes greedy policies). □

2.1.3.2 Application: Optimal Savings with Persistent State Process

As an application, we study a version of the optimal savings model with a persistent state process and a stochastic rate of return on assets. Stochastic returns on assets appear to be important in generating sufficiently heavy right tails in wealth distributions when we take models to the data.[1]

In this model,

- the state is $x = (w, z)$, where $w \in \mathbb{R}_+$ is wealth and z is an exogenous state process on finite set Z with stochastic kernel (matrix) Q,
- the action a is current consumption c, taking values in $\mathbb{R}_+$,
- the feasible correspondence is $\Gamma(x) = \Gamma(w,z) = [0, w]$,
- the reward is $r(x,a) = r((w,z),c) = u(c)$, where u is bounded and continuous,
- the discount factor is $\beta \in (0,1)$, and
- the stochastic kernel takes the form

$$\int v(x') P(x,a,\mathrm{d}x') = \sum_{z'} \int v[R(z')(w-c) + y(z',s'), z'] \varphi(\mathrm{d}s') Q(z,z').$$

[1] See, for example, Benhabib et al. (2015) or Stachurski and Toda (2019). While the second reference shows that heavy-tailed wealth distributions can also be generated by time preference shocks, this channel is unrealistic (since it requires that all households in the economy simultaneously experience time preference shocks in the same direction)..

The kernel can be explained as follows: Labor income is affected by an IID shock s' drawn from distribution $\varphi \in \mathscr{D}(\mathsf{S})$, where S is a topological space. In addition, both the interest rate and labor income are impacted by a common persistent component z. The latter is driven by stochastic matrix Q. We give Z the discrete topology and $\mathsf{X} = \mathbb{R}_+ \times \mathsf{Z}$ the product topology.

Proposition 2.1.4 applies to this model. To show this we need to verify Assumption 2.1.1. The only nontrivial claims are that (a) the correspondence Γ is continuous and (b) that the weak Feller property holds. Claim (a) follows from Exercise A.3.2. Regarding claim (b), we need to show that, after fixing $v \in bc\mathsf{X}$, the mapping

$$m(w, z, c) := \sum_{z'} \int v[R(z')(w - c) + y(z', s'), z'] \varphi(\mathrm{d}s') Q(z, z')$$

is continuous on G. To verify this, we take (w_n, z_n, c_n) converging to (w, z, c). Since Z has the discrete topology, (z_n) is eventually constant at z. Hence it suffices to show that $m(w_n, z, c_n)$ converges to $m(w, z, c)$. This follows from continuity of u, continuity and boundedness of v, and the dominated convergence theorem.

The Bellman operator (2.16) becomes

$$(Tv)(w, z) = \max_{0 \leqslant c \leqslant w} \left\{ u(c) + \beta \sum_{z' \in \mathsf{Z}} \int v[R(z')(w - c) + y(z', s'), z'] \varphi(\mathrm{d}s') Q(z, z') \right\}.$$

Since Proposition 2.1.4 applies, we can fix a $v \in bc\mathsf{X}$ and approximate v_{T} via some iterate $T^k v$.

2.1.4 Strong Feller MDPs

In this section, we consider the general-state MDP model from §2.1.3 under stricter continuity conditions. Doing so will allow us to obtain some stronger conclusions.

2.1.4.1 Theory

To begin, let X and A be separable metric spaces. Let Γ be a nonempty correspondence from X to A. As before, G is the graph of Γ. Let μ be a σ-finite measure on X. In addition, let

- r be a reward function mapping G to $\mathbb{R}$,
- β be a discount factor, and

- p be a transition density on $\mathsf{G} \times \mathsf{X}$, so that that p is nonnegative, Borel measurable and
$$\int p(x, a, x') \mu(\mathrm{d}x') = 1 \quad \text{for all } (x, a) \in \mathsf{G}.$$

Assumption 2.1.2. The following conditions hold:

(i) The reward function r is bounded and continuous,

(ii) The discount factor β obeys $0 < \beta < 1$, and

(iii) $(x, a) \mapsto p(x, a, x')$ is continuous on G for all $x' \in \mathsf{X}$.

We define Σ to be all Borel measurable functions from X to A with $\sigma(x) \in \Gamma(x)$ for all $x \in \mathsf{X}$. We consider policy operators of the form

$$(T_\sigma v)(x) = r(x, \sigma(x')) + \beta \int v(x') p(x, \sigma(x), x') \mu(\mathrm{d}x') \qquad (x \in \mathsf{X}),$$

Let $\mathbb{T}_{\mathrm{SF}}$ be all such operators, with σ ranging over Σ. In the following discussion, Assumption 2.1.2 is in force.

EXERCISE 2.1.14. Show that $(bm\mathsf{X}, \mathbb{T}_{\mathrm{SF}})$ is an ADP.

EXERCISE 2.1.15. Prove that T_σ is a contraction of modulus β on $bm\mathsf{X}$ for every $\sigma \in \Sigma$.

We have the following result.

Proposition 2.1.5. *If Assumption 2.1.2 holds, then $(bm\mathsf{X}, \mathbb{T}_{\mathrm{SF}})$ satisfies the fundamental optimality properties and VFI, HPI, and OPI all converge.*

Proof. To prove Proposition 2.1.5 we first show that $(bm\mathsf{X}, \mathbb{T}_{\mathrm{SF}})$ is regular. To this end, fix $v \in bm\mathsf{X}$ and consider

$$f(x, a) := r(x, a) + \beta \int v(x') p(x, a, x') \mu(\mathrm{d}x') \qquad ((x, a) \in \mathsf{G}).$$

The function f is continuous on G by Assumption 2.1.2 and Lemma A.5.30 on page 223. By this fact and Theorem A.3.3, the ADP $(bm\mathsf{X}, \mathbb{T}_{\mathrm{SF}})$ is regular. The claims in Proposition 2.1.5 now follow from Exercise 2.1.15 and Theorem 1.3.5. □

EXERCISE 2.1.16. Show that the Bellman equation $v = \bigvee_\sigma T_\sigma v$ for the strong Feller model can be expressed as

$$v(x) = \max_{0 \leqslant c \leqslant w} \left\{ r(x, a) + \beta \int v(x') p(x, a, x') \mu(\mathrm{d}x') \right\}. \tag{2.17}$$

(You can use similar arguments to the ones we used to obtain (2.16).)

2.2 Recursive Decision Processes

In this section we consider a class of ADPs that can be thought of as a generalization of the MDP model. Elements of this class will be referred to as "recursive decision processes." Like the MDP model, the value space V of a recursive decision process is a set of real-valued functions on some state space X, and the partial order on V is the pointwise partial order. Unlike MDPs, recursive decision processes can accommodate many nonlinearities in the way that present values are calculated.

2.2.1 Definition and Examples

To a first approximation, recursive decision processes are dynamic programs with a Bellman equation of the form

$$v(x) = \max_{a \in \Gamma(x)} B(x, a, v) \qquad (x \in \mathsf{X}) \tag{2.18}$$

for some suitable choice of B. Here x is the state, a is an action, Γ is a feasible correspondence and B is an "aggregator," with interpretation

> $B(x, a, v)$ = total lifetime rewards, contingent on current action a, current state x and the use of v to evaluate future states.

In §2.2.1.1–2.2.1.2 we improve this definition and then provide examples. As usual, in a topological space setting, "measurable" means "Borel measurable" unless otherwise stated.

2.2.1.1 Definition

Let X and A be topological spaces, referred to henceforth as the **state** and **action spaces** respectively. Given these spaces, a **recursive decision process** (RDP) is a tuple (Γ, V, B) containing

(i) a nonempty correspondence Γ from X to A called the **feasible correspondence**, with an associated set of **feasible state-action pairs**

$$\mathsf{G} := \operatorname{graph}\Gamma = \{(x,a) \in \mathsf{X} \times \mathsf{A} : a \in \Gamma(x)\}$$

and an associated set of **feasible policies**

$$\Sigma := \{\text{all measurable } \sigma \colon \mathsf{X} \to \mathsf{A} \text{ satisfying } \sigma(x) \in \Gamma(x) \text{ for all } x \in \mathsf{X}\},$$

(ii) a subset V of $\mathbb{R}^{\mathsf{X}}$ called the **value space**,

(iii) a map $B \colon \mathsf{G} \times V \to \mathbb{R}$, referred to as an **aggregator**, satisfying the monotonicity condition

$$v \leqslant w \implies B(x,a,v) \leqslant B(x,a,w) \tag{2.19}$$

for every $v, w \in V$ and every $(x,a) \in \mathsf{G}$, and the consistency condition

$$\sigma \in \Sigma \text{ and } v \in V \quad\implies\quad m(x) := B(x, \sigma(x), v) \text{ is in } V. \tag{2.20}$$

The assumption that X and A are topological spaces is important in some applications and irrelevant in others. We maintain it for simplicity. When X and A are discrete, the topology is always the discrete topology. In representing the RDP by the tuple (Γ, V, B), we are treating X and A as understood from context.

The set $\Gamma(x)$ represents all actions available to a controller in state x. Figure 2.1 shows an illustration of one possible correspondence Γ when $\mathsf{A} = \mathsf{X} = \mathbb{R}_+$, along with G, the resulting set of feasible state-action pairs. Figure 2.2 shows a feasible policy σ in the same setting.

The value space V is a class of functions that assign values to states. The order on the left side of (2.19) is the usual pointwise partial order for functions. The monotonicity restriction is natural: relatively to v, if rewards are at least as high w in every future state, then the total rewards we can extract under w should be at least as high.

The final condition, in (2.20), is a consistency condition implying that V is large enough to capture the value of following a particular policy.

2.2.1.2 Example: Optimal Savings

Consider the optimal savings problem studied in §1.1. The state is $w \in \mathbb{R}_+$ and the action is $c \in \mathbb{R}_+$. The feasible correspondence is $\Gamma(w) = [0, w]$ and $V := bm\mathbb{R}_+$ is the

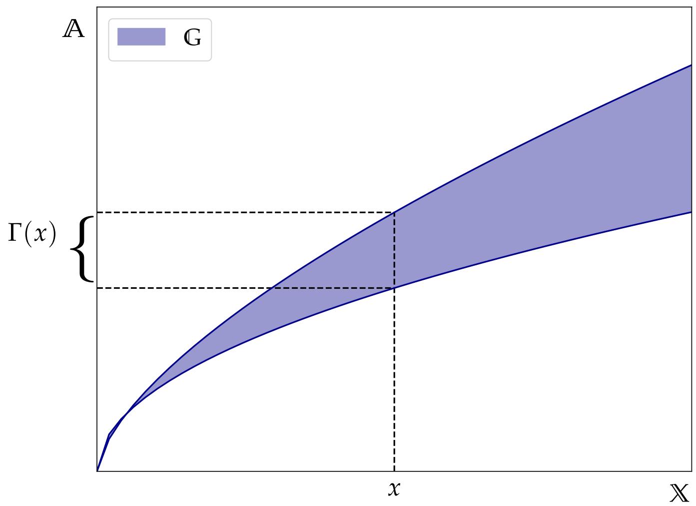


Figure 2.1: Feasible correspondence and feasible state-action pairs

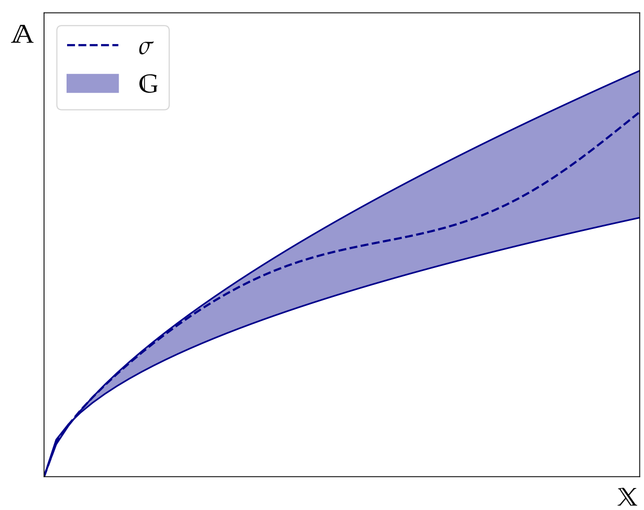


Figure 2.2: The action $\sigma(x)$ lies in $\Gamma(x)$ for all x

value space. We set

$$B(w, c, v) = u(c) + \beta \int v(R(w - c) + y)\varphi(\mathrm{d}y) \qquad (v \in V,\ 0 \leqslant c \leqslant w).$$

As in Assumption 1.1.1, we take u to be bounded and continuous. Under these restrictions, the tuple (Γ, V, B) is an RDP. The monotonicity condition (2.19) clearly holds. The consistency condition (2.20) holds because, by the definition of Γ, a policy is a Borel measurable map $\sigma \colon \mathbb{R}_+ \to \mathbb{R}_+$ with $0 \leqslant \sigma(w) \leqslant w$ for all w, and given any such policy and any $v \in bm\mathbb{R}_+$, the function

$$m(w) := u(\sigma(w)) + \beta \int v(R(w - \sigma(w)) + y)\varphi(\mathrm{d}y)$$

is measurable and bounded (since u is bounded and continuous).

For this model, the RDP Bellman equation $v(x) = \max_{a \in \Gamma(x)} B(x, a, v)$ from (2.18) agrees with the optimal savings Bellman equation in (1.12).

2.2.1.3 Example: MDPs

In §2.2.1.2 we shows that the basic optimal savings model can be framed as an RDP. This is true for all the MDP models we have studied, including the finite state MDPs in §2.1.1, the weak Feller MDPs in §2.1.3, and the strong Feller MDPs in §2.1.4. For example, in the strong Feller case, we take $V = bm\mathsf{X}$, Γ as given, and set

$$B(x, a, v) = r(x, a) + \beta \int v(x')p(x, a, x')\mu(\mathrm{d}x') \qquad ((x, a) \in \mathsf{G},\ v \in V).$$

Exercise 2.2.1. Confirm that, for the associated RDP (Γ, V, B), the monotonicity and consistency conditions (2.19) and (2.20) both hold.

For this model, the RDP Bellman equation $v(x) = \max_{a \in \Gamma(x)} B(x, a, v)$ from (2.18) agrees with the strong Feller Bellman equation in (2.17).

2.2.1.4 Example: MDPs with Modified Rewards

Some authors use an MDP framework where current rewards depend on the next period state, so that the Bellman equation has the form

$$v(x) = \max_{a \in \Gamma(x)} \sum_{x'} \left\{ r(x, a, x') + \beta \sum_{x'} v(x') \right\} P(x, a, x') \qquad (x \in \mathsf{X}). \tag{2.21}$$

Here r maps $\mathsf{G} \times \mathsf{X}$ to $\mathbb{R}$ and other primitives are as in §2.1.1. We take $V = \mathbb{R}^{\mathsf{X}}$, Γ as given, and set

$$B(x, a, v) = \sum_{x'} \left\{ r(x, a, x') + \beta \int v(x') \right\} P(x, a, x') \qquad ((x, a) \in \mathsf{G},\ v \in V).$$

Evidently, for the associated RDP (Γ, V, B), the monotonicity and consistency conditions (2.19) and (2.20) both hold. For this choice of B, the RDP Bellman equation $v(x) = \max_{a \in \Gamma(x)} B(x, a, v)$ agrees with the modified MDP Bellman equation in (2.21).

2.2.1.5 Example: Risk-Sensitive Preferences

Consider a variation on the finite-state MDP model from §2.1.1 where the Bellman equation has the form

$$v(x) = \max_{a \in \Gamma(x)} \left\{ r(x, a) + \frac{\beta}{\theta} \ln \left[\sum_{x'} \exp(\theta v(x')) P(x, a, x') \right] \right\}. \tag{2.22}$$

Here θ is a nonzero parameter in $\mathbb{R}$. In (2.22), the transform $f(v) = \exp(\theta v)$ is applied to v before the expectation is taken and then reversed via $f^{-1}(v) = (1/\theta) \ln(v)$. The agent can be either risk-averse or risk-loving with respect to future outcomes, depending on the value of θ.

To set this model up as an RDP, we take $V = \mathbb{R}^{\mathsf{X}}$, Γ as given, and set

$$B(x, a, v) := r(x, a) + \frac{\beta}{\theta} \ln \left[\sum_{x'} \exp(\theta v(x')) P(x, a, x') \right] \tag{2.23}$$

Exercise 2.2.2. Confirm that, for the associated RDP (Γ, V, B), the monotonicity and consistency conditions (2.19) and (2.20) both hold.

The RDP Bellman equation $v(x) = \max_{a \in \Gamma(x)} B(x, a, v)$ from (2.18) agrees with the Bellman equation in (2.22) when B is given by (2.23).

Add a computational example that helps to illustrate how θ affects choices.

2.2.1.6 Example: Optimal Harvests

A firm owns a timber plantation with biomass s_t at time t. It decides at each t whether to harvest or not. When making this decision, the firm observes current timber price p_t and, should it decide to harvest, sells at that same price. If not then it waits to the next period, where a new price is drawn, biomass updates, and the process repeats.

Biomass takes values in S, which is a closed and bounded interval in $\mathbb{R}_+$, and evolves according to $s_{t+1} = q(s_t)$, where q is a continuous self-map on S. If $q(0) > 0$, then the plantation regenerates after each harvest. If not, the plantation never regenerates and the problem below is an optimal stopping problem.

We assume that the price sequence (p_t) is IID with distribution φ on closed and bounded interval $\mathsf{E} \subset \mathbb{R}_+$. The cost of harvesting given biomass s is $m(s)$. The cost of maintaining the plantation is for one period rather than harvesting is $c(s)$. Both m and c are continuous real-valued functions on S. The firm is risk neutral and discounts the future using discount factor $\beta < 1$.

The state space for the model is $\mathsf{S} \times \mathsf{E}$. The Bellman equation can be expressed as

$$v(s, p) = \max \left\{ ps - m(s) + \beta \int v(q(0), p') \varphi(\mathrm{d}p'),\ -c(s) + \beta \int v(q(s), p') \varphi(\mathrm{d}p') \right\}. \tag{2.24}$$

We construct an RDP that produces this Bellman equation. To do so, we set $V = bm(\mathsf{S} \times \mathsf{E})$, $\Gamma(s, p) = \{0, 1\}$ for all (s, p), and, in addition,

$$B(s, p, a, v) = r(s, p, a) + \beta \int v[f(s, a), p'] \varphi(\mathrm{d}p'),$$

where

$$r(s, p, a) := a(ps - m(s)) - (1 - a)c \quad \text{and} \quad f(s, a) := q[(1 - a)s].$$

Here a is the action, with $a = 1$ indicates the decision to harvest, whereas $a = 0$ indicates the decision to continue without harvesting.

EXERCISE 2.2.3. Show that, for the associated RDP (Γ, V, B), the monotonicity and consistency conditions (2.19) and (2.20) both hold.

EXERCISE 2.2.4. Show that RDP Bellman equation $v(x) = \max_{a \in \Gamma(x)} B(x, a, v)$ from (2.18) is equivalent to (2.24).

The preceding discussion shows that the optimal harvest model is an RDP. In fact we can say more:

EXERCISE 2.2.5. Show that the harvest model can be framed as a weak Feller MDP.

2.2.2 RDPs vs ADPs

As we show below, every RDP generates an ADP. Hence we can use ADP optimality results to understand RDP dynamics. In fact we can use the additional structure provided by RDPs to sharpen our ADP sufficient conditions and bring them closer to applications. We pursue this agenda over the remainder of the chapter.

We also note that not all ADPs are RDPs: there are interesting dynamic programs that cannot be treated from the RDP framework. We clarify this point in §2.2.2.1.

2.2.2.1 RDPs are ADPs

Every RDP generates an ADP. To see this, let (Γ, V, B) be an RDP with state space X and action space A. The set V is paired with the pointwise partial order. With Σ as the set of feasible policies and given σ in Σ, we define T_σ by

$$(T_\sigma v)(x) = B(x, \sigma(x), v) \qquad (x \in \mathsf{X},\ v \in V).$$

The monotonity and consistency conditions (2.19)–(2.20) imply that T_σ is an order-preserving self-map on V. Hence, with $\mathbb{T}$ as the set of all policy operators, the pair $(V, \mathbb{T})$ is an ADP. We call $(V, \mathbb{T})$ the **ADP generated by** (Γ, V, B).

For ADPs generated by RDPs, we can provide intuitive representations of greedy policies and the Bellman equation. For example, we recall from our ADP discussion in §1.2.1.2 that a policy $\sigma \in \Sigma$ is v-greedy for ADP $(V, \mathbb{T})$ if $T_\tau v \leqslant T_\sigma v$ for all $\tau \in \Sigma$. If $(V, \mathbb{T})$ is generated by (Γ, V, B), then this is equivalent to the statement that

$$B(x, \tau(x), v) \leqslant B(x, \sigma(x), v) \quad \text{for all } \tau \in \Sigma \text{ and } x \in \mathsf{X}. \tag{2.25}$$

Also, we recall that the ADP Bellman operator is defined by $Tv = \bigvee_\sigma T_\sigma v$ whenever the supremum exists. When $(V, \mathbb{T})$ is generated by (Γ, V, B), this is equivalent to the statement

$$(Tv)(x) = \sup_{\sigma \in \Sigma} B(x, \sigma(x), v) \qquad (x \in \mathsf{X})$$

whenever the pointwise supremum exists (see Exercise A.1.11).

EXERCISE 2.2.6. Let A and X be finite. Given v in V, show that

(i) $\sigma \in \Sigma$ is v-greedy if and only if

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} B(x, a, v) \qquad \text{for all } x \in \mathsf{X}, \text{ and} \tag{2.26}$$

(ii) the Bellman operator simplifies to

$$(Tv)(x) = \max_{a \in \Gamma(x)} B(x, a, v) \qquad (x \in \mathsf{X}). \tag{2.27}$$

Note that, in the simple setting of Exercise 2.2.6, the Bellman equation takes the form of (2.18). Later we investigate more complex settings where this is still true.

2.2.2.2 Terminology

Let (Γ, V, B) be an RDP and let $(V, \mathbb{T})$ be the ADP generated by (Γ, V, B), as described in §2.2.2.1. To simplify the exposition, we say that

- (Γ, V, B) is **well-posed** (resp., regular, order stable, etc.) if $(V, \mathbb{T})$ is well-posed (resp., regular, order stable, etc.).
- T is **the Bellman operator** for (Γ, V, B) when T is the Bellman operator for $(V, \mathbb{T})$,
- v_σ is the **σ-value function** for (Γ, V, B) when v_σ is the σ-value function for $(V, \mathbb{T})$,
- σ is **optimal** for (Γ, V, B) when σ is optimal for $(V, \mathbb{T})$,
- **VFI converges** for (Γ, V, B) when VFI converges for $(V, \mathbb{T})$, etc.

For example, σ is optimal for (Γ, V, B) when $v_\tau \leqslant v_\sigma$ for all $\tau \in \Sigma$.

2.2.2.3 Not All ADPs are RDPs

Although the RDP framework is broad, there are significant dynamic programs that fall outside this framework.

Example 2.2.1. Later, in §4.1.4, we will consider firm exit problem with Bellman equation given by

$$v(x) = \pi(x) + \int \max\{q(x'), v_\sigma(x')\} K(x, \mathrm{d}x') \quad (x \in \mathsf{X}).$$

While the problem can be represented as an ADP, the ADP cannot be directly placed in the RDP framework for two reasons. One is that the $\max$ operator in the preceding display is inside the integral (as compared to (2.18), where the $\max$ is on the outside). The second is that, in the setting §4.1.4, the value space will be an L_p space, rather than a space of real-valued functions. (Recall that L_p spaces are sets of equivalence classes rather than functions.) This difference is not trivial when considering optimization.

Example 2.2.2. Kristensen et al. (2021) consider a Bellman equation of the form

$$v(z, d) = \int \int \max_{d' \in D} \{u(z', e, d') + \beta v(z', d')\} \, F(e \mid z') F(z' \mid z, d)$$

where d is a discrete choice variable, u is a utility function and β is a discount factor. (Further details can be found in Kristensen et al. (2021).) Similar to Example 2.2.1, this cannot be directly placed in the RDP framework because the $\max$ operator is inside the expectation.

In the two examples above, it is possible to rearrange the problem so that the $\max$ operator is shifted to the outside and, thereby, construct a version that fits the RDP framework. (Later, in Chapter 3, we discuss such rearrangements.) But there are good reasons *not* to do so, such as improved smoothness and lower dimensionality. Related discussion can be found in Kristensen et al. (2021) and Rust (1994).

Example 2.2.3. Later, in §2.3.3, we will investigate linear-quadratic (LQ) problems with Bellman equations such as

$$v(x) = \max_u \left\{ u^\top R u + x^\top Q x + v(Ax + Bu) \right\}. \tag{2.28}$$

Here R, Q, A, and B are matrices, while u and x are vectors. This model looks similar to an RDP if we set $\mathsf{X} = \mathbb{R}^k$, $\mathsf{A} = \mathbb{R}^m$, $\Gamma(x) = \mathsf{A}$ for all $x \in \mathsf{X}$ and, for the aggregator,

$$B(x, u, v) = u^\top R u + x^\top Q x + v(Ax + Bu).$$

However, in §2.3.3 we take Σ to a set of *stable* controls, so that $F \in \Sigma$ means that F is a matrix and $\rho(A + BF) < 1$. Thus, we restrict Σ beyond just feasibility of actions, in contrast to the specification of Σ within the definition of an RDP. (The LQ problem is difficult to handle without such additional restrictions on policies.)

2.2.3 Existence of Greedy Policies

Exercises 2.2.6 investigated a simple finite action RDP setting where greedy policies always exist. In general RDP settings, existence is less trivial. Here we state one useful result.

Lemma 2.2.1. *Let (Γ, V, B) be an RDP such that Γ is continuous and compact-valued on X. Fix $v \in V$. If $(x, a) \mapsto B(x, a, v)$ is continuous on G, then*

(i) *there exists a v-greedy policy σ such that*

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} B(x, a, v) \quad \textit{for all } x \in \mathsf{X} \tag{2.29}$$

and $x \mapsto B(x, \sigma(x), v)$ is continuous on X, and

(ii) *the Bellman operator is well-defined at v, the function Tv is continuous on X, and*

$$(Tv)(x) = \max_{a \in \Gamma(x)} B(x, a, v) \quad (v \in V; x \in \mathsf{X}).$$

If, in addition, $\sigma(x)$ is the unique maximizer of $B(x, a, v)$ over $\Gamma(x)$ for each $x \in \mathsf{X}$, then σ is continuous on X.

Proof of Lemma 2.2.1. Regarding part (i), the assumptions in the lemma and Theorem A.3.3 on page 179 imply the existence of a $\sigma \in \Sigma$ such that (2.29) holds. This policy clearly satisfies (2.25), so σ is v-greedy. Part (ii) follows directly from the assumptions and Theorem A.3.3, as does the final claim under the additional condition that the maximizer is unique at every x. □

Example 2.2.4. We saw in §2.1.3 that the conditions of Lemma 2.2.1 are valid when the RDP is generated by a weak Feller MDP and, in addition, $v \in bc\mathsf{X}$.

Example 2.2.5. We saw in §2.1.4 that the conditions of Lemma 2.2.1 are valid when the RDP is generated by a strong Feller MDP and v is any element of $bm\mathsf{X}$.

2.2.4 Weighted Contractions

RDPs have strong optimality properties when there exists a uniform contraction of values. The current section investigates this case. Throughout this section, A and X

are topological spaces and ℓ is a continuous weight function on X (see §A.5.3.4). Also, $\|\cdot\|$ denotes the ℓ-weighted supremum norm, so that

$$\|f\| := \sup_{x \in \mathsf{X}} \frac{|f(x)|}{\ell(x)}.$$

2.2.4.1 Theory

Let Γ be a continuous nonempty compact-valued correspondence from X to A, referred to henceforth as the feasible correspondence, and let G be the feasible state-action pairs (see §2.2.1.1). Let

- $V = b_\ell m\mathsf{X}$ and
- $B \colon \mathsf{G} \times V \to \mathbb{R}$ be such that $(x, a) \mapsto B(x, a, v)$ is measurable on G for all $v \in V$.

Assumption 2.2.1. The function B is such that

(U1) $(x, a) \mapsto B(x, a, v)$ is continuous on G whenever $v \in b_\ell c\mathsf{X}$,

(U2) $B(x, a, w) \leqslant B(x, a, v)$ for all $w \leqslant v$ in V and $(x, a) \in \mathsf{G}$.

Assumption 2.2.1 is used to obtain monotonicity and existence of greedy policies. The next two conditions are related to contractivity and ℓ-boundedness.

Assumption 2.2.2. The following statements are true:

(U3) There exists a $\lambda \in [0, 1)$ such that

$$|B(x, a, v) - B(x, a, w)| \leqslant \lambda \|v - w\| \, \ell(x) \quad \text{for all } (x, a) \in \mathsf{G} \text{ and } v, w \in V, \text{ and}$$

(U4) for any $v \in V$, there exist constants $M, N \in \mathbb{R}_+$ such that

$$|B(x, a, v)| \leqslant M + N\ell(x) \quad \text{for all } (x, a) \in \mathsf{G}.$$

In this setting, we have the following optimality result for RDPs.

Proposition 2.2.2. *If Assumptions 2.2.1 and 2.2.2 hold, then* (Γ, V, B) *is an RDP. Moreover, for this RDP,*

(i) *at least one optimal policy exists,*

(ii) T *has a unique fixed point* $v_\top$ *in* $b_\ell c\mathsf{X}$ *and* $v_\top$ *is the greatest element of* V_Σ*,*

(iii) σ *is optimal if and only if* σ *is* $v_\top$*-greedy, and*

(iv) $\|T^n v - v_\top\| \to 0$ *as* $n \to \infty$ *for all* $v \in b_\ell c\mathsf{X}$*.*

Proof. First we show that (Γ, V, B) is an RDP. The only nontrivial claim is that the consistency condition (2.20) holds. To see that this is so, let Σ be all measurable $\sigma \colon \mathsf{X} \to \mathsf{A}$ with $\sigma(x) \in \Gamma(x)$ for all x. Fix $v \in V$ and $\sigma \in \Sigma$, and consider the policy operator $T_\sigma v$ defined by $(T_\sigma v)(x) = B(x, \sigma(x), v)$. Since σ is, by definition, measurable, our restrictions on B imply that $T_\sigma v$ is measurable. Moreover, (U4) yields constants M, N such that $|T_\sigma v| \leqslant M + N\ell$, so $T_\sigma v$ is ℓ-bounded. Hence $T_\sigma v \in V$ and the consistency condition holds.

Next we prove (i)–(iv) by checking the conditions of Theorem 1.3.2 on page 28. To this end, we fix $v, w \in V$ and $\sigma \in \Sigma$. We use (U3) to obtain

$$\frac{|T_\sigma v - T_\sigma w|}{\ell} \leqslant \lambda \|v - w\|. \tag{2.30}$$

It follows that T_σ is a contraction map on V under the complete metric induced by the ℓ-weighted sup norm $\|\cdot\|$, and therefore globally stable on V. This proves that $(V, \mathbb{T})$, the ADP generated by (Γ, V, B), is globally stable.

Next we claim that the Bellman operator T is globally stable on $b_\ell c\mathsf{X}$. The first step is to prove that $T b_\ell c\mathsf{X} \subset b_\ell c\mathsf{X}$. To see this, note that, by Lemma 2.2.1, we have $(Tv)(x) = \max_{a \in \Gamma(x)} B(x, a, v)$ when $v \in b_\ell c\mathsf{X}$. This function is continuous by Theorem A.3.3 on page 179. The second step is to establish the contraction property. For this purpose we fix $v, w \in b_\ell c\mathsf{X}$ and use the sup inequality in Lemma A.1.2 and (2.30) to obtain

$$\frac{|Tv - Tw|}{\ell} \leqslant \sup_{\sigma \in \Sigma} \frac{|T_\sigma v - T_\sigma w|}{\ell} \leqslant \lambda \|v - v'\|.$$

Taking the supremum over the left hand side verifies contractivity. Since $b_\ell c\mathsf{X}$ is closed in the complete metric space V (by Theorem A.5.21), global stability holds.

The final step of the proof is to verify that each $v \in b_\ell c\mathsf{X}$ has at least one greedy policy (since, in this case, all the conditions in Theorem 1.3.2 hold). But this is true by

Lemma 2.2.1, given (U1) and our restrictions on Γ. The results in Proposition 2.2.2 now follow from the conclusions of Theorem 1.3.2. □

EXERCISE 2.2.7. Consider the setting of Assumption 2.2.1 but suppose that, instead of (U3), the following condition holds:

$$B(x, a, v + c\ell) \leqslant B(x, a, v) + \lambda c \ell \quad \text{for all } (x, a, v) \in \mathsf{G} \times V \text{ and all } c \in \mathbb{R}_+. \tag{2.31}$$

Prove that this condition is sufficient for (U3).

The next exercise asks you to check the conditions of Proposition 2.2.2 for a model where rewards are bounded (which is the relatively easy case).

EXERCISE 2.2.8. Consider an optimal savings model identical to the one described in §2.1.3.2 except that agents die with a time-dependent probability in each period (see, e.g., De Nardi et al. (2020)). To accommodate this feature, we modify the state to $x = (w, z, t) \in \mathsf{X} := \mathbb{R}_+ \times \mathsf{Z} \times \mathbb{Z}_+$, where t represents time, and set

$$B((w, z, t), c, v) = u(c) + \beta\, q(t) \sum_{z' \in \mathsf{Z}} \int v[R(z')(w - c) + y(z', s'), z', t + 1] \varphi(\mathrm{d}s') Q(z, z').$$

Here $q(t)$ is the survival probability for the agent at age t. Higher probability of dying increases the rate at which the future is discounted. Show that the conditions of Proposition 2.2.2 are satisfied when $\ell = \mathbb{1}$. (Impose the discrete topology on $\mathbb{Z}_+$.)

2.2.4.2 Application: Unbounded Weak Feller MDPs

We consider a weak Feller MDP where, unlike the discussion in §2.1.3, the reward function r is allowed to be unbounded. As before, the reward function is continuous on G, the discount factor β takes values in $(0, 1)$, and the feasible correspondence Γ is nonempty, continuous and compact-valued. A weight function ℓ is introduced below and we continue to use $\| \cdot \|$ to represent the ℓ-weighted supremum norm.

We also adopt the following conditions.

Assumption 2.2.3. There exists a continuous function $\ell \colon \mathsf{X} \to [1, \infty)$ such that, for some nonnegative constants γ, δ, η with $\eta < 1/\beta$,

$$r(x, a) \leqslant \gamma + \delta \ell(x) \quad \text{and} \quad \int \ell(x') P(x, a, \mathrm{d}x') \leqslant \eta \ell(x) \tag{2.32}$$

for all $(x, a) \in \mathsf{G}$. In addition, $(x, a) \mapsto \int h(x')P(x, a, \mathrm{d}x')$ is continuous on G whenever $h \in b_\ell c\mathsf{X}$.

We convert the problem into an RDB by setting $V = b_\ell m\mathsf{X}$, taking Γ as given, and defining

$$B(x, a, v) := r(x, a) + \beta \int v(x')P(x, a, \mathrm{d}x') \qquad ((x, a) \in \mathsf{G},\ v \in V).$$

We claim that (U1)–(U4) in Assumptions 2.2.1 and 2.2.2 all hold. Conditions (U1) and (U2) are immediate from Assumption 2.2.3 and the definition of B. Regarding (U3), we have, for $v, w \in V$,

$$\begin{aligned} |B(x, a, v) - B(x, a, w)| &\leqslant \beta \int \frac{|v(x') - w(x')|}{\ell(x')} \ell(x')P(x, a, \mathrm{d}x') \\ &\leqslant \beta \|v - w\| \int \ell(x')P(x, a, \mathrm{d}x'). \end{aligned}$$

By Assumption 2.2.3, the last term is dominated by $\beta\eta\|v - w\|\ell(x)$ and $\lambda := \beta\eta < 1$. Hence (U3) is satisfied.

Regarding (U4), we fix $v \in V$ and use Assumption 2.2.3 to obtain

$$|B(x, a, v)| \leqslant |r(x, a)| + \beta \int \frac{|v(x')|}{\ell(x')} \ell(x')P(x, a, \mathrm{d}x') \leqslant \gamma + \delta\ell(x) + \beta\|v\|\eta\ell(x)$$

for every $(x, a) \in \mathsf{G}$. Hence (U4) holds. This gives the following result for (V, Γ, B).

Proposition 2.2.3. *If Assumption 2.2.3 holds for the weak Feller RDP, then so do Assumptions 2.2.1 and 2.2.2. As a result, the conclusions of Proposition 2.2.2 are valid.*

The next section gives one example of how Proposition 2.2.3 can be applied.

2.2.4.3 Application: Optimal Savings with Unbounded Utility

To illustrate Proposition 2.2.3, we consider a savings problem with IID income where $\mathsf{X} = \mathsf{A} = \mathbb{R}_+$. We suppose that $\bar{y} := \int y\varphi(\mathrm{d}y) < \infty$, and that utility function u is increasing and concave as a map from $\mathbb{R}_+$ to itself, with $u(0) = 0$ and $\lim_{c\to\infty} u'(c) = 0$. We also assume that $1 < R < 1/\beta$.

EXERCISE 2.2.9. Prove that, under the stated assumptions, there exist nonnegative constants γ, δ such that $\delta \leqslant 1$ and $u(c) \leqslant \gamma + \delta c$ for all $c \in \mathbb{R}_+$.

In Exercise 2.2.9, we can obviously choose γ such that

$$\frac{\bar{y}}{R-1} \leqslant \gamma. \tag{2.33}$$

Then we set

$$\ell(w) = \gamma + w \quad \text{and} \quad V = b_\ell m \mathbb{R}_+.$$

We also set $\Gamma(w) = [0, w]$ and

$$B(w, c, v) = u(c) + \beta \int v(R(w-c) + y)\varphi(\mathrm{d}y).$$

We claim that, under these restrictions, Assumption 2.2.3 holds for the RDP (Γ, V, B). Indeed, for this choice of ℓ and feasible (w, c) pair, we have

$$u(c) \leqslant u(w) \leqslant \gamma + \delta w \leqslant \gamma + w = \ell(w),$$

so the first bound in (2.32) holds. Regarding the second,

$$\int \ell(w')P(w, c, \mathrm{d}w') = R(w-c) + \bar{y} + \gamma \leqslant Rw + \bar{y} + \gamma = R\left(w + \frac{\bar{y}+\gamma}{R}\right).$$

Rearranging (2.33) gives $(\bar{y}+\gamma)/R \leqslant \gamma$, and apply this to the last display yields

$$\int \ell(w')P(w, c, \mathrm{d}w') \leqslant R\,(w + \gamma) = R\ell(w).$$

Since $R < 1/\beta$, the condition in Assumption 2.2.3 holds with $\eta := R$.

2.2.5 Properties of Solutions

In this section, we seek sufficient conditions for the value and policy functions to have useful shape and continuity properties. We adopt the setting of §2.2.4 and study the properties of the RDP (Γ, V, B) discussed in Proposition 2.2.2. In the proofs below, we repeatedly use Lemma A.2.8 on page 173.

2.2.5.1 Monotone Values

First, we seek conditions under which the value function is increasing. We consider an RDP (Γ, V, B) where (a) the state space X is partially ordered by $\preceq$, and (b) the conditions of Proposition 2.2.2 are satisfied for some continuous weight function ℓ on X. Let

$$ib_\ell c\mathsf{X} := \text{ the set of increasing functions in } b_\ell c\mathsf{X}.$$

EXERCISE 2.2.10. Show that $ib_\ell c\mathsf{X}$ is a closed subset of $b_\ell c\mathsf{X}$.

Assumption 2.2.4. If $x \preceq x'$, then, for all $v \in ib_\ell c\mathsf{X}$ and all $a \in \Gamma(x')$,

$$\Gamma(x) \subset \Gamma(x') \quad \text{and} \quad B(x, a, v) \leqslant B(x', a, v).$$

Both conditions in Assumption 2.2.4 are monotonicity conditions. The first is equivalent to stating that Γ is order preserving when viewed as a map from $(\mathsf{X}, \preceq)$ to $(\wp(\mathsf{A}), \subset)$. Here $\wp(\mathsf{A})$ is the set of all subsets of A and $\subset$ is the partial order induced by set inclusion (Example A.1.2).

Proposition 2.2.4. *If the conditions of Proposition 2.2.2 and Assumption 2.2.4 both hold, then $v_\top$ is increasing on X.*

Proof. It suffices to show that T is invariant on $ib_\ell c\mathsf{X}$, since, by Proposition 2.2.2, T is globally stable on $b_\ell c\mathsf{X}$ and, in addition, $ib_\ell c\mathsf{X}$ is closed in $b_\ell c\mathsf{X}$. To see that this holds, pick any $v \in ib_\ell c\mathsf{X}$ and fix x and x' with $x \preceq x'$. Since T is invariant on $b_\ell c\mathsf{X}$, we need only show that Tv is increasing. But this must be so, since, by Assumption 2.2.4,

$$\sup_{a \in \Gamma(x)} B(x, a, v) \leqslant \sup_{a \in \Gamma(x')} B(x, a, v) \leqslant \sup_{a \in \Gamma(x')} B(x', a, v).$$

Hence $Tv(x) \leqslant Tv(x')$ and T is invariant on $ib_\ell c\mathsf{X}$. □

To illustrate the conditions, let us consider again the optimal savings problem from §2.2.4.3, where $V = b_\ell m\mathbb{R}_+$, $\Gamma(w) = [0, w]$, and

$$B(w, c, v) = u(c) + \beta \int v(R(w - c) + y)\varphi(\mathrm{d}y).$$

We saw in that section that the conditions of Proposition 2.2.2 hold, so, to show that $v_\top$ is increasing on $\mathbb{R}_+$, we only need to verify the conditions of Assumption 2.2.4.

EXERCISE 2.2.11. Prove that Assumption 2.2.4 holds under the conditions in §2.2.4.3.

2.2.5.2 Concavity

Next we seek sufficient conditions for the value function to be concave. In this section, we assume that both X and A are convex subsets of a vector space.

Assumption 2.2.5. The set of feasible state-action pairs G is convex and $(x, a) \mapsto B(x, a, v)$ is concave on G whenever v is concave on X.

The convexity requirement on G in Assumption 2.2.5 is equivalent to the statement that, for all x, x' in X, all $a \in \Gamma(x)$ all $a' \in \Gamma(x')$ and all $\lambda \in [0, 1]$, we have

$$\lambda a + (1 - \lambda)a' \in \Gamma(\lambda x + (1 - \lambda)x').$$

By taking $x = x'$, we see that each set $\Gamma(x)$ is convex in A.

Proposition 2.2.5. *Let the conditions of Proposition 2.2.2 hold. If, in addition, Assumption 2.2.5 holds, then $v_\top$ is concave on X.*

Proof. Let $cb_\ell c\mathsf{X}$ be the concave functions in $b_\ell c\mathsf{X}$. By a similar argument to the one used in the proof of Proposition 2.2.4, it suffices to show that T is invariant on $cb_\ell c\mathsf{X}$. To this end, fix v in $cb_\ell c\mathsf{X}$, λ in $[0, 1]$ and $x_0, x_1 \in \mathsf{X}$. Let a_i satisfy $Tv(x_i) = B(x_i, a_i, v)$ for each i. Let $x_\lambda = \lambda x_0 + (1 - \lambda)x_1$ and $a_\lambda = \lambda a_0 + (1 - \lambda)a_1$. By convexity of G, we know that a_λ lies in $\Gamma(x_\lambda)$, which gives

$$\lambda B(x_0, a_0, v) + (1 - \lambda)B(x_1, a_1, v) \leqslant B(x_\lambda, a_\lambda, v) \leqslant Tv(x_\lambda).$$

The left-hand side is $\lambda Tv(x_0) + (1 - \lambda)Tv(x_1)$, so we have proved concavity of Tv. Hence T is invariant on $cb_\ell c\mathsf{X}$, and the claim in Proposition 2.2.5 holds. □

EXERCISE 2.2.12. Consider the optimal savings problem studied at the end of §2.2.5.1. Prove that if u is also concave, then $v_\top$ is increasing and concave.

2.2.5.3 Uniqueness and Continuity

When the conditions of Proposition 2.2.2 are force, we know that at least one optimal policy exists in Σ. The question we ask now is, when is it unique? Not surprisingly, uniqueness can be obtained with a form of strict concavity.

Assumption 2.2.6. Assumption 2.2.5 is satisfied and, in addition, $a \mapsto B(x, a, v)$ is strictly concave on $\Gamma(x)$ for all x in X and all concave v in X.

Proposition 2.2.6. *Let the conditions of Proposition 2.2.2 hold. If Assumption 2.2.6 also holds, then the optimal policy is both unique and continuous.*

Proof. Let Assumption 2.2.6 hold. An optimal policy exists, since we are assuming the conditions of Proposition 2.2.2. Thus, only uniqueness needs to be shown. By the same theorem, a policy is optimal if and only if it is $v_\top$-greedy. So, to prove uniqueness, it suffices to show that there cannot be two such policies in Σ. To see this, observe that $v_\top$ is concave on X by Proposition 2.2.5. Hence, under Assumption 2.2.5, the map $a \mapsto B(x, a, v_\top)$ is strictly concave at each x. Strictly concave functions have unique maximizers, so the $v_\top$-greedy policy is unique in Σ. Continuity now follows from Lemma 2.2.1 on page 65. □

EXERCISE 2.2.13. Consider the optimal savings in the setting of Exercise 2.2.12. Suppose that, in addition, u is strictly concave. Show that, in this setting, the optimal policy is unique and continuous.

2.3 Beyond RDPs

In this section we consider models with general state spaces that do not fit into the RDP specification.

2.3.1 Q-Factors

Next we examine the Q-factor variation of the MDP model, which provides an alternative view on the Bellman equation. This variation is used extensively in Q-learning (Watkins, 1989; Tsitsiklis, 1994), which is a branch of reinforcement learning. We can study optimality of the Q-factor variation either by treating it directly, as an ADP in its own right, or by inferring optimality properties from the original MDP version of the problem in §2.1.1. The first approach is treated here and the second is treated in Chapter 3.

2.3.1.1 The Q-Factor Model

To begin, we take the MDP model from §2.1.1 and, given $v \in \mathbb{R}^{\mathsf{X}}$, set

$$q(x, a) := r(x, a) + \beta \sum_{x'} v(x')P(x, a, x') \qquad ((x, a) \in \mathsf{G}). \tag{2.34}$$

The function q is called the **Q-factor** corresponding to v. We will convert the original MDP Bellman equation from §2.1.1 (see (2.7)) into an equation in Q-factors. The first step is to observe that, given q in (2.34), the Bellman equation can be written as $v(x) = \max_{a \in \Gamma(x)} q(x, a)$. Taking the expectaions and discounting on both sides of this equation yields

$$\beta \sum_{x'} v(x') P(x, a, x') = \beta \sum_{x'} \max_{a' \in \Gamma(x')} q(x', a') P(x, a, x').$$

Adding $r(x, a)$ and using the definition of q again gives

$$q(x, a) = r(x, a) + \beta \sum_{x'} \max_{a' \in \Gamma(x')} q(x', a') P(x, a, x'). \tag{2.35}$$

This is the **Q-factor Bellman equation**. To study it, we introduce a family of policy operators $\mathbb{S} := \{S_\sigma : \sigma \in \Sigma\}$ via

$$(S_\sigma q)(x, a) = r(x, a) + \beta \sum_{x'} q(x', \sigma(x')) P(x, a, x') \qquad ((x, a) \in \mathsf{G}). \tag{2.36}$$

Here S_σ acts on function $q \in \mathbb{R}^{\mathsf{G}}$. The set $\mathbb{R}^{\mathsf{G}}$ is paired with the pointwise partial order.

Exercise 2.3.1. Prove that each S_σ is a contraction of modulus β on $\mathbb{R}^{\mathsf{G}}$ with respect to the supremum norm.

Exercise 2.3.2. Prove that $(\mathbb{R}^{\mathsf{G}}, \mathbb{S})$ is an ADP.

Exercise 2.3.3. Fixing $q \in \mathbb{R}^{\mathsf{G}}$, show that $\sigma \in \Sigma$ is q-greedy for the ADP $(\mathbb{R}^{\mathsf{G}}, \mathbb{S})$ whenever

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} q(x, a) \quad \text{for all } x \in \mathsf{X}.$$

Exercise 2.3.4. Fix $q \in \mathbb{R}^{\mathsf{G}}$. The ADP Bellman operator corresponding to $(\mathbb{R}^{\mathsf{G}}, \mathbb{S})$ obeys $Sq := \bigvee_\sigma S_\sigma q$. Recall that σ is q-greedy if and only if $S_\sigma q = Sq$ (Lemma 1.2.1). Using this fact and Exercise 2.3.3, show that Sq can also be written as

$$(Sq)(x, a) = r(x, a) + \beta \sum_{x'} \max_{a' \in \Gamma(x')} q(x', a') P(x, a, x'). \tag{2.37}$$

Evidently $q \in \mathbb{R}^{\mathsf{G}}$ is a fixed point of S if and only if it is a solution to the Q-factor Bellman equation (2.35).

EXERCISE 2.3.5. Prove that the fundamental optimality properties hold for $(\mathbb{R}^{\mathsf{G}}, \mathbb{S})$ and that VFI, OPI and HPI all converge.

2.3.1.2 Risk-Sensitive Q-Factors

Some reinforcement learning applications use risk-sensitive Q-factor methods (see, e.g., Fei et al. (2021)), where the Bellman equation has the form

$$f(x,a) = r(x,a) + \frac{\beta}{\theta} \ln \left\{ \sum_{x'} \exp \left[\theta \max_{a' \in \Gamma(x')} f(x', a') \right] P(x, a, x') \right\} \tag{2.38}$$

for some nonzero $\theta \in \mathbb{R}$. Here $(x, a) \in \mathsf{G}$ and $f \in \mathbb{R}^{\mathsf{G}}$. We take X, A, Γ, Σ and G as in our discussion of Q-factors in the previous section. The Bellman equation can be compared with the standard Q-factor Bellman equation in (2.35). The new parameter θ controls attitude to risk, similar to its role in the risk-sensitive MDP from §2.2.1.5.

Given $\sigma \in \Sigma$, the policy operator corresponding to the Bellman equation (2.38) takes the form

$$(T_\sigma f)(x,a) = r(x,a) + \frac{\beta}{\theta} \ln \left[\sum_{x'} \exp \left[\theta f(x', \sigma(x')) \right] P(x, a, x') \right]. \tag{2.39}$$

Here $f \in \mathbb{R}^{\mathsf{G}}$. Let $\mathbb{T}$ be the set of all such operators, where $\sigma \in \Sigma$. Continuing to endow $\mathbb{R}^{\mathsf{G}}$ with the pointwise order $\leqslant$, we can easily confirm the pair $(\mathbb{R}^{\mathsf{G}}, \mathbb{T})$ is an ADP. The next exercise uses our assumption that A is finite, and that $\Gamma(x)$ is nonempty for all x.

EXERCISE 2.3.6. Show that $(\mathbb{R}^{\mathsf{G}}, \mathbb{T})$ is regular.

EXERCISE 2.3.7. The ADP Bellman operator is defined at $f \in \mathbb{R}^{\mathsf{G}}$ by $Tf = \bigvee_\sigma T_\sigma f$. Show that $f \in \mathbb{R}^{\mathsf{G}}$ solves $Tf = f$ if and only if it solves (2.38).

2.3.2 Structural Estimation

Structural estimation is one of the most important fields of quantitative economics. Under this approach to estimation, researchers model economic agents as if they solve dynamic programs. The econometric challenge is to infer parameters that bring the model outputs as close as possible to the data. In this section, we set aside the estimation problem and focus instead on the types of dynamic programs typically adopted in this field.

2.3.2.1 Post-Action Value Functions

Rust (1987) and many subsequent authors study discrete choice problems with nonstandard Bellman equations that take the form

$$g(x, a) = \int \max_{a' \in \mathsf{A}} [r(x', a') + \beta g(x', a')] \, P(x, a, \mathrm{d}x') \quad ((x, a) \in \mathsf{G} := \mathsf{X} \times \mathsf{A}). \tag{2.40}$$

Here “discrete choice” means that A is a finite set. The state space X can be any metric space. The reward function $r \in \mathbb{R}^{\mathsf{G}}$ is assumed to be bounded and Borel measurable, while P is a stochastic kernel from G to X. The function g is interpreted as an “expected post-action value function.” The advantages of working with this version of the Bellman equation are discussed in Rust (1994), Kristensen et al. (2021) and other sources.

We can set this problem up as an ADP by taking Σ to be the set of Borel measurable maps from X to A and, for each $\sigma \in \Sigma$, introducing the policy operator

$$(T_\sigma g)(x, a) = \int [r(x', \sigma(x')) + \beta g(x', \sigma(x'))] P(x, a, \mathrm{d}x'). \tag{2.41}$$

Exercise 2.3.8. Show that T_σ is an order preserving self map on $(bm\mathsf{G}, \leqslant)$, the set of bounded Borel measurable functions in $\mathbb{R}^{\mathsf{G}}$ with the pointwise partial order.

We consider the ADP $(bm\mathsf{G}, \mathbb{T})$ where each $T_\sigma \in \mathbb{T}$ is given by (2.41).

Exercise 2.3.9. Fix $\sigma \in \Sigma$. Show that

$$\sigma(x) \in \operatorname*{argmax}_{a \in A} \{r(x, a) + \beta g(x, a)\} \text{ for all } x \in \mathsf{X} \quad \implies \quad \sigma \text{ is } g\text{-greedy}.$$

Since A is finite and nonempty, Exercise 2.3.9 implies that $(bm\mathsf{G}, \mathbb{T})$ is regular.

Exercise 2.3.10. Show that the Bellman equation $g = \bigvee_\sigma T_\sigma g$ is equivalent to (2.40).

Proposition 2.3.1. *For the structural estimation ADP $(bm\mathsf{G}, \mathbb{T})$, the fundamental optimality properties hold and VFI, HPI, and OPI all converge.*

Proof. It is straightforward to confirm that each T_σ is a contraction of modulus β on $bm\mathsf{G}$ under the supremum norm. Proposition 2.3.1 then follows from Theorem 1.3.5. □

2.3.2.2 State-Dependent Discounting

Next we consider an extension of the structural estimation model in Section 2.3.2.1 with state-depdendent discounting. The Bellman equation is

$$g(x,a) = \sum_{x'} \max_{a' \in \mathsf{A}} \left[r(x',a') + \beta(x') g(x',a') \right] P(x,a,x') \tag{2.42}$$

where $(x,a) \in \mathsf{G} := \mathsf{X} \times \mathsf{A}$ and A, X are finite and nonempty. As usual, r is a reward function on G and P is a transition kernel from G to X. The discount factor β is allowed to be a function of the state. We let $\| \cdot \|$ be the supremum norm and take Σ to be the set of all functions from X to A.

Given $\sigma \in \Sigma$, let T_σ be defined at $g \in \mathbb{R}^{\mathsf{G}}$ and $(x,a) \in \mathsf{G}$ by

$$(T_\sigma g)(x,a) = \sum_{x'} \left[r(x', \sigma(x')) + \beta(x') g(x', \sigma(x')) \right] P(x,a,x')$$

Let $\mathbb{T} = \{T_\sigma\}_{\sigma \in \Sigma}$. Each T_σ is an order-preserving self-map on $\mathbb{R}^{\mathsf{G}}$, so $(\mathbb{R}^{\mathsf{G}}, \mathbb{T})$ is an ADP. For each $g \in \mathbb{R}^{\mathsf{G}}$, we can construct a g-greedy policy by taking a $\sigma \in \Sigma$ such that

$$\sigma(x) \in \operatorname*{argmax}_{a \in \mathsf{A}} [r(x,a) + \beta(x) g(x,a)] \quad \text{for all } x \in \mathsf{X}. \tag{2.43}$$

Such a σ exists by finiteness of A, so $(\mathbb{R}^{\mathsf{G}}, \mathbb{T})$ is regular.

Let T be the Bellman operator, so that $Tg := \bigvee_\sigma T_\sigma g$. When evaluated at a g-greedy policy σ, we have $T_\sigma g = Tg$. Using this equality and (2.43) yields

$$(Tg)(x,a) = \sum_{x'} \max_{a' \in \mathsf{A}} \left[r(x',a') + \beta(x') g(x',a') \right] P(x,a,x').$$

Thus, solutions to $Tg = g$ solve the original Bellman equation (2.42).

For each $\sigma \in \Sigma$ we set

$$(K_\sigma g)(x,a) = \sum_{x'} \beta(x') g(x', \sigma(x')) P(x,a,x') \qquad (x,a) \in \mathsf{G}$$

We have the following result:

Proposition 2.3.2. *If $\rho(K_\sigma) < 1$ for all $\sigma \in \Sigma$, then the fundamental optimality properties hold for $(\mathbb{R}^{\mathsf{G}}, \mathbb{T})$ and HPI converges in finitely many steps.*

This discounting condition generalizes the traditional assumption that β is constant and strictly less than one, in which case $\rho(K_\sigma) < 1$ always holds.

Proof of Proposition 2.3.2. We have already shown that $(\mathbb{R}^{\mathsf{G}}, \mathbb{T})$ is regular. Hence, by Theorem 1.2.12, we need only show that $(\mathbb{R}^{\mathsf{G}}, \mathbb{T})$ is order stable. As every globally stable and order preserving self-map is order stable (Lemma A.5.17), the conclusions of Proposition 2.3.2 hold whenever each $T_\sigma \in \mathbb{T}$ is globally stable on $\mathbb{R}^{\mathsf{G}}$.

To see that this is so, fix $\sigma \in \Sigma$ and $f, g \in \mathbb{R}^{\mathsf{G}}$. We then have

$$|(T_\sigma f)(x,a) - (T_\sigma g)(x,a)| \leqslant \sum_{x'} \beta(x')|f(x', \sigma(x')) - g(x', \sigma(x'))|P(x,a,x')$$

That is, $|T_\sigma f - T_\sigma g| \leqslant K_\sigma |f - g|$. Applying Theorem 1.3.7, we see that T_σ is eventually contracting and therefore globally stable. □

2.3.2.3 Beyond Expected Utility

Some studies find incompatibilities between data and predictions of models that use additively separable preferences and mathematical expectation to evaluate uncertain outcomes (see, e.g., Lu et al. (2024)). To further this line of analysis, we revisit the basic structural estimation model in §2.3.2.1 while replacing mathematical expectation with a general certainty equivalent operator.

As in §2.3.2.1, spaces of bounded real-valued functions are paired with the pointwise order $\leqslant$, and the supremum norm, to be denoted by $\|\cdot\|$. The state space X is a metric space, A is a finite choice set, $\mathsf{G} := \mathsf{X} \times \mathsf{A}$, $r \colon \mathsf{G} \to \mathbb{R}$ is a measurable reward function, and $\beta \in (0,1)$ is a constant discount factor. However, we modify the Bellman equation (2.40) for the post-action value function to

$$g(x,a) = (\mathscr{E} H g)(x,a) \quad \text{where } (Hg)(x') := \max_{a' \in \mathsf{A}} \left[r(x',a') + \beta g(x',a') \right]. \tag{2.44}$$

Here $\mathscr{E}$ is a certainty equivalent operator mapping $bm\mathsf{X}$ into $bm\mathsf{G}$, in the sense that $\mathscr{E}$ is order-preserving and $\mathscr{E}\lambda = \lambda$ whenever λ is constant. We assume in addition that $\mathscr{E}$ is **constant subadditive**, meaning that $\mathscr{E}(f + \lambda) \leqslant \mathscr{E} f + \lambda$ for all $f \in bm\mathsf{X}$ and $\lambda \in \mathbb{R}_+$.

Let Σ be the set of Borel measurable maps from X to A. Given $\sigma \in \Sigma$, we set

$$(T_\sigma g)(x,a) = (\mathscr{E} H_\sigma g)(x,a) \quad \text{where } (H_\sigma g)(x') := r(x', \sigma(x')) + \beta g(x', \sigma(x')).$$

With $\mathbb{T} := \{T_\sigma\}_{\sigma \in \Sigma}$, the pair $(bm\mathsf{G}, \mathbb{T})$ is an ADP and $\sigma \in \Sigma$ is g-greedy whenever

$$\sigma(x) \in \operatorname*{argmax}_{a' \in \mathsf{A}} \left[r(x', a') + \beta g(x', a') \right] \quad \text{for all } x \in \mathsf{X}. \tag{2.45}$$

Since A is finite and nonempty, such a policy always exists. (A measurable selection theorem can be used to obtain Borel measurability of σ. See, for example, Aliprantis and Border (2006), Theorem 18.19.) Hence $(bm\mathsf{G}, \mathbb{T})$ is regular.

Given $g \in bm\mathsf{G}$, the ADP Bellman operator T satisfies $Tg = T_\sigma g$ whenever σ is g-greedy. Using this fact and (2.45), we obtain $Tg = \mathscr{E} H g$. Hence any fixed point of T solves the original Bellman equation (2.40).

Proposition 2.3.3. *If $\mathscr{E}$ is constant subadditive, then the fundamental optimality properties hold for $(bm\mathsf{G}, \mathbb{T})$, and VFI, OPI, and HPI all converge.*

Proof. Fix $f, g \in bm\mathsf{G}$. Since $\mathscr{E}$ and H_σ are order-preserving, we have

$$T_\sigma f = \mathscr{E} H_\sigma (g + f - g) \leqslant \mathscr{E} H_\sigma (g + \|f - g\|) \leqslant \mathscr{E}(H_\sigma g + \beta \|f - g\|).$$

Using constant subadditivity of $\mathscr{E}$ and rearranging gives $T_\sigma f - T_\sigma g = \beta \|f - g\|$. Reversing the roles of f and g yields $|T_\sigma f - T_\sigma g| \leqslant \beta \|f - g\|$. Taking the supremum, we see that each T_σ is a contraction of modulus β on $bm\mathsf{G}$. As regularity was confirmed above, this shows that the conditions of Theorem 1.3.5 are all verifed, which proves the proposition. □

As an illustration, suppose that $\mathscr{E}$ is the risk-sensitive certainty equivalent

$$(\mathscr{E} f)(x, a) := \frac{1}{\theta} \ln \left\{ \int \exp\left[\theta f(x')\right] P(x, a, \mathrm{d}x') \right\} \qquad ((x, a) \in \mathsf{G}),$$

where P is a stochastic kernel from G to X and θ is a nonzero constant. This choice of certainty equivalent is constant subadditive, so, among other things, Proposition 2.3.3 tells us that $\sigma \in \Sigma$ is optimal if and only if

$$\sigma(x) \in \operatorname*{argmax}_{a' \in \mathsf{A}} \left[r(x', a') + \beta g_{\top}(x', a') \right] \quad \text{for all } x \in \mathsf{X},$$

where $g_{\top}$ is the unique solution to the functional equation

$$g(x, a) = \frac{1}{\theta} \ln \left\{ \int \exp\left\{ \theta \max_{a' \in \mathsf{A}} \left[r(x', a') + \beta g(x', a') \right] \right\} P(x, a, \mathrm{d}x') \right\}$$

in the value space $bm\mathsf{G}$.

Another example of a nonlinear certainty equivalent operator is the quantile operator studied in de Castro and Galvao (2019) and de Castro et al. (2022), which allows for separation of intertemporal elasticity of substitution and risk aversion. This certainty equivalent is also constant subadditive, so Proposition 2.3.3 extends the results in de Castro and Galvao (2019).

2.3.3 LQ Control

LQ control is a sub-field of dynamic programming often applied to problems engineering, economics, operations research and elsewhere. In this section we describe a canonical LQ problem and show how it can be solved using ADP methods. The LQ problem was briefly described in Example 2.2.3 and, as discussed there, the Bellman equation takes the form

$$v(x) = \max_{u \in \mathbb{R}^m} \left\{ u^\top R u + x^\top Q x + v(Ax + Bu) \right\} \quad \text{for all } x \in \mathbb{R}^k. \tag{2.46}$$

Here

- A is $k \times k$ and B is $m \times k$,
- Q is $k \times k$ and negative semidefinite, and
- R is $m \times m$ and negative definite.

The problem is infinite horizon and deterministic, without discounting. The vector x is the state variable and u is the action. In the present setting, u is also referred to as the **control**.

Since the solution to this problem is well-known, we will not seek new results. Rather, our aim is to illustrate how we can embed the LQ problem into the ADP framework and recover existing results in a relatively straightforward way.

2.3.3.1 Preamble: Riccati Equations

For the purposes of what follows, it will be helpful to note some standard results on the Riccati mapping. Let $\mathcal{N}$ be the set of negative semidefinite $k \times k$ matrices. The Riccati mapping $N \mapsto \mathbf{R}N$ [2] is a self-map on $\mathcal{N}$ defined by

$$\mathbf{R}N = A^\top (N - NB(B^\top NB + R)^{-1} B^\top N) A + Q, \tag{2.47}$$

[2] In this section we use caligraphic upper case to represent matrix-valued maps.

The matrices A, B, R, and Q are as described after the Bellman equation (2.46). The fixed point equation $N = \mathbf{R}N$ is called the **Riccati equation**.

EXERCISE 2.3.11. Confirm that $\mathbf{R}$ maps $\mathcal{N}$ into itself.

Add solution.

We will also make use of a second matrix-valued mapping $N \mapsto \mathbf{F}N$, defined by

$$\mathbf{F}N = -(B^\top N B + R)^{-1} B^\top N A. \tag{2.48}$$

Lemma 2.3.4. *If $N \in \mathcal{N}$, $F = \mathbf{F}N$ and $v(x) = x^\top N x$, then*

$$Fx = \operatorname*{argmax}_{u \in \mathbb{R}^m} \left\{ u^\top R u + x^\top Q x + v(Ax + Bu) \right\} \quad \textit{for all } x \in \mathbb{R}^k. \tag{2.49}$$

Moreover, the maximizer is unique.

Proof. See Section (add section number) of Bertsekas (2012). □

In stating the next result we let C be such that $C^\top C = Q$. We refer to Bertsekas (2012) for the definitions of observability and controllability.

Lemma 2.3.5. *If (A, B) is controllable and (A, C) is observable, then*

(i) *The Riccati map $\mathbf{R}$ has a unique fixed point in $\mathcal{N}$.*

(ii) *If $N = \mathbf{R}N$ and $F = \mathbf{F}N$, then $\rho(A + BF) < 1$.*

Proof. See Proposition 3.1.1 of Bertsekas (2012). □

2.3.3.2 Description

We consider a deterministic undiscounted **linear-quadratic (LQ) control problem**, defined as a tuple (Q, R, A, B), where the primitives are as in §2.3.3.1. The objective of the LQ problem is to solve

$$\max_{(u_t)} \sum_{t \geqslant 0} \left[u^\top R u + x^\top Q x \right] \tag{2.50}$$

subject to

$$x_{t+1} = Ax_t + Bu_t \qquad \text{for all } t \geqslant 0.$$

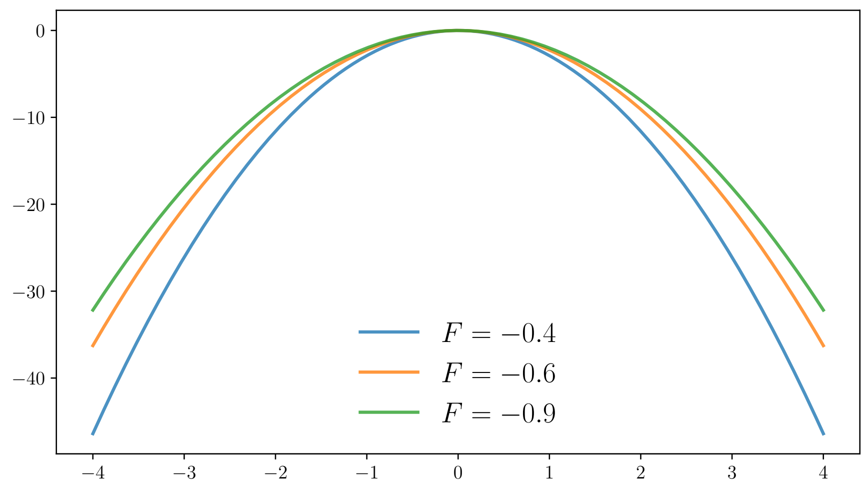


Figure 2.3: The function v_F for different choices of F

The vector $x_t \in \mathbb{R}^k$ is called the **state variable** and $u_t \in \mathbb{R}^m$ is called the **control**. Note that $u_t^\top R u_t + x_t^\top Q x_t \leqslant 0$ for all t, so the infinite sum (2.50) takes values in $[-\infty, 0]$.

In the LQ setting, a **control matrix** is any $F \in \mathbb{R}^{m \times k}$. Under a given control matrix F, the current control obeys $u_t = F x_t$ and the update rule for the state is $x_{t+1} = A x_t + B F x_t$. Hence the state evolves according to $x_t = (A + BF)^t x_0$. Following (2.50), the lifetime value of following F, starting at initial condition $x_0 \in \mathbb{R}^k$, is

$$v_F(x_0) := \sum_{t=0}^{\infty} x_t^\top \left(F^\top R F + Q\right) x_t \quad \text{with } x_t = (A + BF)^t x_0. \tag{2.51}$$

Example 2.3.1. Suppose that $m = k = 1$ and $Q = R = 1$, so that a control matrix is a scalar $F \in \mathbb{R}$. In view of (2.51), the lifetime value of F when starting at x_0 is

$$v_F(x_0) = c x_0^2 \quad \text{where} \quad c := (F^2 + 1) \sum_{t=0}^{\infty} (A + BF)^{2t}. \tag{2.52}$$

The sum is finite when $|A + BF| < 1$. Figure 2.3 plots v_F for different choices of F when $a = b = 1$ and the stability condition holds. Of the three alternatives, $F = -0.9$ attains the highest value from every state.

Returning to the general case, finite lifetime values require driving the state to zero fast enough for the sum (2.51) to converge. In this connection, extending the condition $|A+BF| < 1$ from the one-dimensional example, a control matrix F is called **stable** if the spectral radius condition $\rho(A + BF) < 1$ holds.

EXERCISE 2.3.12. Prove that, for any fixed $x_0 \in \mathbb{R}^k$, the sequence $x_t = (A + BF)^t x_0$ converges to zero as $t \to \infty$ when F is a stable control matrix.

2.3.3.3 From LQ to ADP

We wish to produce an ADP representation of the LQ problem. To this end, we endow $\mathcal{N}$ with the **Loewner partial order** $\preccurlyeq$, so that

$$A \preccurlyeq B \quad \iff \quad A - B \in \mathcal{N}.$$

Also, we set

- $\Sigma :=$ the set of stable control matrices, and
- $\mathbb{T} := \{\mathbf{T}_F : F \in \Sigma\}$ with each policy operator $N \mapsto \mathbf{T}_F N$ defined by

$$\mathbf{T}_F\, N = F^\top R F + Q + \mathbf{L}_F N \tag{2.53}$$

  where $\mathbf{L}_F\, N := (A + BF)^\top N(A + BF)$ is understood as a linear self-map on $\mathbb{R}^{k\times k}$.

In line with earlier ADP terminology, a stable control matrix is also referred to as a **policy**.

Lemma 2.3.6. *Every $\mathbf{T}_F \in \mathbb{T}$ is an order preserving self-map on $(\mathcal{N}, \preccurlyeq)$.*

Proof. Fix $F \in \Sigma$. $\mathbf{T}_F$ is a self-map on $\mathcal{N}$ because $\mathcal{N}$ is stable under addition and $\mathbf{L}_F \mathcal{N} \subset \mathcal{N}$. To see that $\mathbf{T}_F$ is order preserving, fix $N_1 \preccurlyeq N_2$. Since $N_1 - N_2$ is negative semidefinite, so is

$$\mathbf{T}_F\, N_1 - \mathbf{T}_F\, N_2 = (A + BF)^\top (N_1 - N_2)(A + BF). \tag{2.54}$$

Hence $\mathbf{T}_F\, N_1 \preccurlyeq \mathbf{T}_F\, N_2$. □

EXERCISE 2.3.13. Confirm that the expression on the right-hand side of (2.54) is negative semidefinite when $N_1 - N_2$ is negative semidefinite.

It follows directly from Lemma 2.3.6 that $(\mathcal{N}, \mathbb{T})$ is an ADP.

Lemma 2.3.7. *Every $\mathbf{T}_F \in \mathbb{T}$ is globally stable on $\mathcal{N}$ with unique fixed point*

$$N_F := \sum_{t=0}^{\infty} [(A + BF)^t]^\top \left(F^\top RF + Q\right) (A + BF)^t. \tag{2.55}$$

Proof. Fix $\mathbf{T}_F \in \mathbb{T}$. Since F is stable, $\rho(\mathbf{L}_F) < 1$ on $\mathbb{R}^{k\times k}$. Hence, by the Neumann series lemma (see, in particular, Corollary A.4.11 on page 201), the map $\mathbf{T}_F$ is globally stable on $\mathcal{N}$ with unique fixed point

$$N_F = \sum_{t=0}^{\infty} \mathbf{L}_F^t \left(F^\top RF + Q\right) = \sum_{t=0}^{\infty} [(A + BF)^t]^\top \left(F^\top RF + Q\right) (A + BF)^t.$$

This verifies the expression for N_F in (2.55). □

We have emphasized that fixed points of policy operators will be identified with lifetime values in applications. This is also true for the ADP $(\mathcal{N}, \mathbb{T})$. Indeed, comparing (2.55) with (2.51), we see that $x^\top N_F x = v_F(x)$ for all $x \in \mathbb{R}^k$. Hence N_F is the matrix representation of the lifetime value function.

Lemma 2.3.8. *The ADP $(\mathcal{N}, \mathbb{T})$ is well-posed and order stable.*

Proof. Fix any $\mathbf{T}_F \in \mathbb{T}$. By Lemma 2.3.7, $\mathbf{T}_F$ is globally stable. It follows that $(\mathcal{N}, \mathbb{T})$ is well-posed (since globally stable maps have unique fixed points). Moreover, by Lemma A.5.17 on page 214, globally stable order preserving self-maps are order stable. Hence $(\mathcal{N}, \mathbb{T})$ is order stable. □

2.3.3.4 Greedy Policies

Fixing $N \in \mathcal{N}$ and applying the ADP definition of greedy policies (page 12), we see that a policy $F \in \Sigma$ is N-greedy if and only if $T_G N \preccurlyeq T_F N$ for all $G \in \Sigma$. The next exercise is useful for characterizing greedy policies.

EXERCISE 2.3.14. Prove that, for $N \in \mathcal{N}$ and control matrix $F \in \mathbb{R}^{m\times k}$,

$$F = \mathbf{F}N \quad\iff\quad T_G\, N \preccurlyeq T_F\, N \text{ for every } G \in \mathbb{R}^{m\times k}.$$

The following lemma is immediate from Exercise 2.3.14.

Lemma 2.3.9. *Let $\mathcal{N}_S := \{N \in \mathcal{N} : \mathbf{F}N \in \Sigma\}$. If $N \in \mathcal{N}_S$, then $\mathbf{F}N$ is N-greedy.*

2.3.3.5 The Bellman Equation

Recalling the definitions from §1.2.1.3, the Bellman operator associated with the ADP $(\mathcal{N}, \mathbb{T})$ is defined by $\mathbf{T}N = \bigvee_{F \in \Sigma} \mathbf{T}_F N$ whenever the supremum exists. If $N \in \mathcal{N}_S$ and $F = \mathbf{F}N$, then, by Lemma 2.3.9, F is N-greedy. Hence $\mathbf{T}N = \mathbf{T}_F N$ (as follows from Lemma 1.2.1). In other words, for all $N \in \mathcal{N}_S$,

$$\mathbf{T}N = F^\top RF + Q + (A + BF)^\top N(A + BF) \quad \text{when} \quad F = \mathbf{F}N \tag{2.57}$$

The next lemma connects this to the Riccati equation and another version of the LQ Bellman equation.

Lemma 2.3.10. *For $N \in \mathcal{N}_S$, the following statements are equivalent:*

(i) *N solves the Riccati equation $\mathbf{R}N = N$.*

(ii) *N solves the ADP Bellman equation $\mathbf{T}N = N$.*

(iii) *The function $v(x) := x^\top N x$ obeys (2.46).*

Proof. The equivalence of (i) and (ii) can be proved by direct calculation (see, e.g., Proposition 3.1.1 in Bertsekas (2012)). Regarding (iii), suppose $N \in \mathcal{N}_S$ and $N = \mathbf{T}N$. Then, setting $v(x) := x^\top N x$ and using the expression for $\mathbf{T}N$ in (2.57) yields

$$v(x) = x^\top F^\top RFx + x^\top Qx + v(Ax + BFx).$$

Applying (2.49) yields (iii). Using similar logic we can show that (iii) implies (ii). □

2.3.3.6 Optimality

Suppose that (A, B) is controllable and (A, C) is observable. We saw above that $(\mathcal{N}, \mathbb{T})$ is well-posed and order stable. Moreover, the ADP Bellman operator $\mathbf{T}$ has at least one fixed point in $\mathcal{N}_S$. Indeed, Lemma 2.3.5 tells us that $\mathbf{R}$ has a fixed point $N_\top$ in $\mathcal{N}_S$, while Lemma 2.3.10 implies that $N_\top$ is also a fixed point of $\mathbf{T}$. Letting $\mathcal{N}_\Sigma$ be the set of lifetime values, so that

$$\mathcal{N}_\Sigma = \{N = N_F : F \in \Sigma\},$$

and applying Theorem 1.2.6, we obtain the following optimality results:

(a) the matrix $N_\top$ is the greatest element of $\mathcal{N}_\Sigma$,

(b) the matrix $N_\top$ obeys $\mathbf{T}N_\top = N_\top$, and

(c) a policy F is optimal if and only if F is $N_\top$-greedy.

We can translate (a)–(c) into more familiar optimality results for LQ problems. For example, (b) combined with Lemma 2.3.10 tells us that

$$v_\top(x) = \max_{u \in \mathbb{R}^m} \left\{ u^\top R u + x^\top Q x + v_\top(Ax + Bu) \right\} \quad (x \in \mathbb{R}^k), \tag{2.58}$$

where $v_\top$ is defined by $v_\top(x) = x^\top N_\top x$ for all x. Moreover, if we now set $F = \mathbf{F} N_\top$, then, by Lemma 2.3.9, the policy F is $N_\top$-greedy. Hence, by (c), F is optimal.

2.4 Chapter Notes

In §2.2.5 we extended some of the optimality results in Bäuerle and Jaśkiewicz (2018), giving conditions under which a unique and continuous optimal policy exists. Our framework is also similar to the contractive models in Bertsekas (2022), to which we add uniqueness, concavity and continuity results under additional assumptions.

RDPs were treated in Sargent and Stachurski (2025a). The main difference between the treatment in Sargent and Stachurski (2025a) and the theory in this chapter is that we drop the assumption of finite states and actions.

Chapter 3

Transformations

Roadmap. While the theory is very general, many examples in this chapter adopt state and action spaces are finite in order to help manage complexity and simplify the exposition. Make some connections to the transformations used in Ch. 5 of DP1.

3.1 Isomorphic Dynamic Programs

One core task of mathematics is drawing connections between distinct objects. For example, researchers study isomorphisms that connect distinct groups via their characteristics under group operations. Topological conjugacies help researchers identify similar dynamic systems. Topological isomorphisms play an analogous role for topological spaces. Order isomorphisms are used to connect posets.

In this section we introduce a concept of "isomorphic" dynamic programs. In particular, we describe an isomorphic relationship that leads to essentially equivalent optimality properties. This concept combines the idea of topological conjugacy for dynamic systems with the notion of order isomorphisms. We exploit these ideas to study additional optimization problems, including decision problems with ambiguity.

3.1.1 Isomorphisms and Anti-Isomorphisms

Roadmap.

3.1.1.1 Order Isomorphisms

A surjective map F from poset $(V, \preceq)$ to poset $(\hat{V}, \trianglelefteq)$ is called an

- **order isomorphism** if $v \preceq w \iff Fv \trianglelefteq Fw$, and an
- **order anti-isomorphism** if $v \preceq w \iff Fw \trianglelefteq Fv$.

(Surjective means that F maps V onto $\hat{V}$, so each $\hat{v} \in \hat{V}$ has a preimage.) When such an order isomorphism (resp., anti-isomorphism) exists, we say that V and $\hat{V}$ are isomorphic (resp., anti-isomorphic).

Example 3.1.1. Let $V = \hat{V} = \mathbb{R}^n_+$ with the usual pointwise order. Consider F mapping $v \in V$ to $v^2 \in \hat{V}$, where the operation $v \mapsto v^2$ acts pointwise. F is clearly onto and, for $v \geqslant 0$, we have $v \leqslant v'$ if and only if $v^2 \leqslant (v')^2$. Hence F is an order isomorphism.

In the next two exercises, X is any nonempty set and all spaces of real-valued functions have the pointwise partial order. Scalar actions on functions are applied pointwise; for example, given $h \in \mathbb{R}^{\mathsf{X}}$, the function $\exp h$ maps x to $\exp(h(x))$.

EXERCISE 3.1.1. Given $h \in \mathbb{R}^{\mathsf{X}}$, let $Fh = \exp(\theta h)$. Show that F is an order isomorphism (resp., anti-isomorphism) from $\mathbb{R}^{\mathsf{X}}$ to $(0, \infty)^{\mathsf{X}}$ whenever $\theta > 0$ (resp., $\theta < 0$).

The next exercise generalizes the previous one.

EXERCISE 3.1.2. Let

$$V = M^{\mathsf{X}} \quad \text{and} \quad \hat{V} = \hat{M}^{\mathsf{X}} \quad \text{where } M, \hat{M} \subset \mathbb{R}. \tag{3.1}$$

Let φ be a map from M onto $\hat{M}$ and let $Fv = \varphi \circ v$. Prove the following:

- If φ is an order isomorphism from M to $\hat{M}$, then F is an order isomorphism from V to $\hat{V}$.
- If φ is an order anti-isomorphism from M to $\hat{M}$, then F is an order anti-isomorphism from V to $\hat{V}$.

EXERCISE 3.1.3. Let V and $\hat{V}$ be posets. Show that every order isomorphism F from poset is a bijection. Show that every order anti-isomorphism is also a bijection.

EXERCISE 3.1.4. Let F be a bijection from $(V, \preceq)$ to $(\hat{V}, \trianglelefteq)$. Show that

(i) F is an order isomorphism if and only if F and F^{-1} are order preserving, and

(ii) F is an order anti-isomorphism if and only if F and F^{-1} are order reversing.

Lemma 3.1.1. *Let F be an order isomorphism from $(V, \preceq)$ to $(\hat{V}, \trianglelefteq)$. If the supremum of $\{v_\alpha\}_{\alpha\in\Lambda} \subset V$ exists in V, then*

$$\bigvee_\alpha Fv_\alpha \text{ exists in } \hat{V} \text{ and } \bigvee_\alpha Fv_\alpha = F \bigvee_\alpha v_\alpha.$$

Exercise 3.1.5. Prove Lemma 3.1.1.

The next exercise is related to Lemma 3.1.1.

Exercise 3.1.6. Let $V, \hat{V}$ be posets, let (v_n) be a sequence in V, and let F be a map from V to $\hat{V}$. Prove the following:

(i) If F is an order isomorphism, then $v_n \uparrow v$ in V if and only if $Fv_n \uparrow Fv$ in $\hat{V}$.

(ii) If F is an order anti-isomorphism, then $v_n \uparrow v$ in V if and only if $Fv_n \downarrow Fv$ in $\hat{V}$.

Exercise 3.1.7. Prove the following claims:

(i) If $V, \hat{V}$ are order isomorphic, then V is totally ordered if and only if $\hat{V}$ is totally ordered.

(ii) F is an order anti-isomorphism from V to $\hat{V}$ if and only if F is an order isomorphism from V to the dual $\hat{V}^\partial$.

3.1.1.2 Conjugate Dynamics

We recall that a (discrete time) **dynamical system** is a pair (V, S), where V is any set and S is a self-map on V. Two dynamical systems (V, S) and $(\hat{V}, \hat{S})$ are said to be **conjugate under F** (or just **conjugate**) if

$$F \text{ is a bijection from } V \text{ into } \hat{V} \text{ and } F \circ S = \hat{S} \circ F \text{ on } V.$$

We can also write the last equality as $S = F^{-1} \circ \hat{S} \circ F$. This helps us understand the conjugacy relationship: shifting a point $v \in V$ to Sv via S is equivalent to moving v

into $\hat{V}$ via $\hat{v} = Fv$, applying $\hat{S}$, and then moving the result back using F^{-1}:

$$
\begin{tikzcd}
v \arrow[r, "S"] \arrow[d, "F"] & Sv \\
\hat{v} \arrow[r, "\hat{S}"] & \hat{S}\hat{v} \arrow[u, "F^{-1}"']
\end{tikzcd}
$$

Example 3.1.2. Consider the dynamical systems $(\mathbb{R}, S)$ and $((0,\infty), \hat{S})$ where $Sx = ax + b$ and $\hat{S}x = \exp(b)x^a$. With $F = \exp$ we have

$$FSx = \exp(b)\exp(ax) = \exp(b)\exp(x)^a = \hat{S}Fx.$$

Hence $(\mathbb{R}, S)$ and $((0,\infty), \hat{S})$ are conjugate under F.

The next result lists some consequences of conjugacy.

Proposition 3.1.2. *If (V, S) and $(\hat{V}, \hat{S})$ are conjugate, then*

(i) $S^n = F^{-1}\hat{S}^n F$ *for all* $n \in \mathbb{N}$,

(ii) v *is a fixed point of* S *if and only if* Fv *is a fixed point of* $\hat{S}$,

(iii) $\hat{v}$ *is a fixed point of* $\hat{S}$ *if and only if* $F^{-1}\hat{v}$ *is a fixed point of* S, *and*

(iv) v *is the unique fixed point of* S *in* V *if and only if* Fv *is the unique fixed point of* $\hat{S}$ *in* $\hat{V}$.

The proofs of these claims are straightforward. For example, regarding (i), suppose that $S^n = F^{-1}\hat{S}^n F$ at some fixed n. Then, using conjugacy,

$$S^{n+1} = SS^n = SF^{-1}\hat{S}^n F = F^{-1}\hat{S}\hat{S}^n F = F^{-1}\hat{S}^{n+1} F.$$

EXERCISE 3.1.8. Prove part (v) of Proposition 3.1.2.

To define isomorphic ADPs and relationships between them, we will use a similarity notion for dynamical systems called "order conjugacy" that adds order to the concept of conjugate dynamics. To this end, we take V and $\hat{V}$ to be posets and consider two dynamical systems (V, S) and $(\hat{V}, \hat{S})$. We call these systems **order conjugate under F** if they are conjugate under F and, in addition, F is an order isomorphism. To indicate that such an F can be found, we simply say that (V, S) and $(\hat{V}, \hat{S})$ are **order conjugate**.

EXERCISE 3.1.9. Prove: Order conjugacy is an equivalence relation on the set of dynamical systems over partially ordered sets.

The next lemma shows one benefit of establishing order conjugacy.

Lemma 3.1.3. *If (V, S) and $(\hat{V}, \hat{S})$ are order conjugate under F, then S is is order stable on V if and only if $\hat{S}$ is order stable on $\hat{V}$.*

EXERCISE 3.1.10. Prove Lemma 3.1.3.

In the next section, we use order conjugacy to connect dynamic programs.

3.1.1.3 Isomorphic ADPs

Let $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ be ADPs with policy sets $\mathbb{T} := \{T_\sigma : \sigma \in \Sigma\}$ and $\hat{\mathbb{T}} := \{\hat{T}_\sigma : \sigma \in \Sigma\}$. We call these ADPs **isomorphic** under F if

(i) F is an order isomorphism from V to $\hat{V}$,

(ii) these two ADPs have the same policy set Σ, and

(iii) (V, T_σ) and $(\hat{V}, \hat{T}_\sigma)$ are order conjugate under F for all $\sigma \in \Sigma$.

Part (iii) obviously requires that

$$F \circ T_\sigma = \hat{T}_\sigma \circ F \quad \text{on } V \text{ for all } \sigma \in \Sigma. \tag{3.2}$$

Example 3.1.3. Fei et al. (2021) consider an "exponential" risk-sensitive Q-factor Bellman equation, which has policy operator $M_\sigma \colon (0, \infty)^{\mathsf{G}} \to (0, \infty)^{\mathsf{G}}$ defined by

$$(M_\sigma h)(x, a) = \exp\left\{\theta r(x, a) + \beta \ln\left[\sum_{x'} h(x', \sigma(x'))P(x, a, x')\right]\right\} \tag{3.3}$$

for all $(x, a) \in \mathsf{G}$. The parameter θ is a nonzero constant and all other primitives are the same as for the finite MDP model (see §2.1.1). The operator M_σ is related to the risk-sensitive Q-factor policy operator T_σ from (2.39), which we repeat here for convenience:

$$(T_\sigma f)(x, a) = r(x, a) + \frac{\beta}{\theta} \ln\left[\sum_{x'} \exp\left[\theta f(x', \sigma(x'))\right] P(x, a, x')\right].$$

Let $\mathbb{T} := \{T_\sigma : \sigma \in \Sigma\}$ and $\mathbb{M} := \{M_\sigma : \sigma \in \Sigma\}$. We saw in §2.3.1.2 that $(\mathbb{R}^\mathsf{G}, \mathbb{T})$ is a regular ADP. Suppose for now that $\theta > 0$ and let F be the order isomorphism from Exercise 3.1.1, which maps $h \in \mathbb{R}^\mathsf{G}$ into $Fh \in (0, \infty)^\mathsf{G}$ via $Fh = \exp(\theta h)$. Then, for $h \in \mathbb{R}^\mathsf{G}$ and $(x, a) \in \mathsf{G}$, we have

$$
\begin{aligned}
(F\,T_\sigma\,h)(x,a) &= \exp\left\{\theta r(x,a) + \beta \ln\left[\sum_{x'} \exp\left[\theta h(x', \sigma(x'))\right] P(x,a,x')\right]\right\} \\
&= \exp\left\{\theta r(x,a) + \beta \ln\left[\sum_{x'} (Fh)(x', \sigma(x')) P(x,a,x')\right]\right\}.
\end{aligned}
$$

Comparing with (3.3), we see that the last expression equals $(M_\sigma\,Fh)(x,a)$. This proves that $F \circ T_\sigma = M_\sigma \circ F$ on $\mathbb{R}^\mathsf{G}$. Hence $((0,\infty)^\mathsf{G}, \mathbb{M})$ and $(\mathbb{R}^\mathsf{G}, \mathbb{T})$ are isomorphic.

Example 3.1.4. Let (Γ, V, B) and $(\Gamma, \hat{V}, \hat{B})$ be two RDPs with identical state space X, action space A and feasible correspondence Γ. Let V and $\hat{V}$ be as in (3.1). If there exists an order isomorphism φ from M onto $\hat{M}$ such that

$$
B(x,a,v) = \varphi^{-1}[\hat{B}(x,a,\varphi \circ v)] \quad \text{for all } v \in V \text{ and } (x,a) \in \mathsf{G}, \tag{3.4}
$$

then $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are isomorphic. To see this, recall that $Fv = \varphi \circ v$ is an order isomorphism from V to $\hat{V}$ (Exercise 3.1.2). Moreover, the respective policy operators are linked by $T_\sigma = F^{-1} \circ \hat{T}_\sigma \circ F$ on V. This confirms our claim.

Lemma 3.1.4. *Isomorphism between ADPs is an equivalence relation on the set of ADPs.*

In other words, if $\mathbf{A}$ is the set of all ADPs and, for $(V, \mathbb{T}), (\hat{V}, \hat{\mathbb{T}}) \in \mathbf{A}$, the symbol $(V, \mathbb{T}) \sim (\hat{V}, \hat{\mathbb{T}})$ means $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are isomorphic, then $\sim$ is reflexive, symmetric and transitive.

Exercise 3.1.11. Prove Lemma 3.1.4. (Hint: Use Exercise 3.1.9.)

3.1.1.4 Isomorphisms and Optimality

We seek relationships between optimality properties of isomorphic ADPs. For all of this section, we take $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ to be two ADPs with $\mathbb{T} = \{T_\sigma : \sigma \in \Sigma\}$ and $\hat{\mathbb{T}} = \{\hat{T}_\sigma : \sigma \in \Sigma\}$. In particular, the two ADPs share the same policy set Σ. When they exist, we let

- v_σ (resp., $\hat{v}_\sigma$) be the unique fixed point of T_σ (resp., $\hat{T}_\sigma$),
- T (resp., $\hat{T}$) be the Bellman operator of $(V, \mathbb{T})$ (resp., $(\hat{V}, \hat{\mathbb{T}})$), and

- $v_\top$ (resp., $\hat{v}_\top$) be the value function of $(V, \mathbb{T})$ (resp., $(\hat{V}, \hat{\mathbb{T}})$).

The next theorem shows that isomorphic ADPs share the same regularity and optimality properties:

Theorem 3.1.5. *If $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are isomorphic under F, then*

(i) *σ is v-greedy for $(V, \mathbb{T})$ if and only if σ is Fv-greedy for $(\hat{V}, \hat{\mathbb{T}})$,*

(ii) *$(V, \mathbb{T})$ is regular if and only if $(\hat{V}, \hat{\mathbb{T}})$ is regular,*

(iii) *$(V, \mathbb{T})$ is well-posed if and only if $(\hat{V}, \hat{\mathbb{T}})$ is well-posed,*

(iv) *$(V, \mathbb{T})$ is order stable if and only if $(\hat{V}, \hat{\mathbb{T}})$ is order stable, and*

(v) *σ is optimal for $(V, \mathbb{T})$ if and only if σ is optimal for $(\hat{V}, \hat{\mathbb{T}})$.*

Proof. Let $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ be isomorphic under F. Regarding (i), fix $v \in V$ and suppose that σ is v-greedy for $(V, \mathbb{T})$. Then $T_\tau\, v \preceq T_\sigma\, v$ and hence $FT_\tau\, v \preceq FT_\sigma\, v$ for all $\tau \in \Sigma$. Conjugacy now implies that $\hat{T}_\tau\, Fv \preceq \hat{T}_\sigma\, Fv$ for all $\tau \in \Sigma$, so σ is Fv-greedy for $(\hat{V}, \hat{\mathbb{T}})$. The converse implication is symmetric.

Claim (ii) is immediate from claim (i). Claims (iii) and (iv) follow directly from order conjugacy of the policy operators (as in (3.2)) and Lemma 3.1.3. Regarding (v), we use order conjugacy of the policy operators to obtain $Fv_\sigma = \hat{v}_\sigma$ for all $\sigma \in \Sigma$, from which it follows that

$$v_\sigma = \bigvee_\tau v_\tau \quad \iff \quad Fv_\sigma = F \bigvee_\tau v_\tau = \bigvee_\tau Fv_\tau \quad \iff \quad \hat{v}_\sigma = \bigvee_\tau \hat{v}_\tau$$

In other words, σ is optimal for $(V, \mathbb{T})$ if and only if σ is optimal for $(\hat{V}, \hat{\mathbb{T}})$. □

The next theorem studies the case when $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are regular and well-posed.

Theorem 3.1.6. *Let $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ be regular, well-posed ADPs. If $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are isomorphic under F, then*

(i) *the Bellman operators obey*

$$F \circ T = \hat{T} \circ F \text{ on } V, \tag{3.5}$$

(ii) *when they exist, the value functions for $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are related by $\hat{v}_\top = Fv_\top$,*

(iii) *the fundamental optimality properties hold for $(V, \mathbb{T})$ if and only if they hold for $(\hat{V}, \hat{\mathbb{T}})$.*

Proof. Let $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ be as stated. Regarding (i), we fix $v \in V$ and apply (3.2) to obtain

$$FTv = F \bigvee_\sigma T_\sigma v = \bigvee_\sigma FT_\sigma v = \bigvee_\sigma \hat{T}_\sigma Fv = \hat{T} F v.$$

(The second equality follows from regularity and Lemma 3.1.1.) This confirms (3.5), so T and $\hat{T}$ are order conjugate under F.

Regarding (ii), suppose that $v_\top = \bigvee_\sigma v_\sigma$ exists. Then $\hat{v}_\top = \bigvee_\sigma Fv_\sigma = F \bigvee_\sigma v_\sigma = Fv_\top$.

Regarding (iii), suppose that the fundamental optimality properties hold for $(V, \mathbb{T})$. We need only show that they likewise hold for $(\hat{V}, \hat{\mathbb{T}})$, since the reverse implication then holds by symmetry. First, an optimal policy exists for $(\hat{V}, \hat{\mathbb{T}})$ by existence for $(V, \mathbb{T})$ and part (v) of Theorem 3.1.5. Second, $v_\top$ is the unique fixed point of T and, in addition, (V, T) and $(\hat{V}, \hat{T})$ are order conjugate, so $Fv_\top$ is the unique fixed point of $\hat{T}$. In view of (ii), this means that $\hat{v}_\top$ is the unique fixed point of $\hat{T}$. Bellman's principle of optimality also holds for $(\hat{V}, \hat{\mathbb{T}})$ by Lemma 1.2.4 on page 18 (or by (i) and (v) of Theorem 3.1.5). □

The next theorem considers convergence of algorithms.

Theorem 3.1.7. *Let $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ be regular, well-posed ADPs. If $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are isomorphic under F, then*

(i) *the respective optimistic policy operators W and $\hat{W}$ obey*

$$F \circ W = \hat{W} \circ F \text{ on } V, \text{ and} \tag{3.6}$$

(ii) *the respective Howard policy operators H and $\hat{H}$ obey*

$$F \circ H = \hat{H} \circ F \text{ on } V. \tag{3.7}$$

Moreover, if the fundamental optimality properties hold for one and hence both of these ADPs, then the following statements are true.

(iii) *VFI converges for $(V, \mathbb{T})$ if and only if VFI converges for $(\hat{V}, \hat{\mathbb{T}})$,*

(iv) *OPI converges for $(V, \mathbb{T})$ if and only if OPI converges for $(\hat{V}, \hat{\mathbb{T}})$, and*

(v) *HPI converges for $(V, \mathbb{T})$ if and only if HPI converges for $(\hat{V}, \hat{\mathbb{T}})$.*

Proof. Fix $m \in \mathbb{N}$. Let $W := W_m$ and H be the optimistic and Howard policy operators for $(V, \mathbb{T})$. Let $\hat{W} := \hat{W}_m$ and $\hat{H}$ be the optimistic and Howard policy operators for $(\hat{V}, \hat{\mathbb{T}})$. Fix $v \in V$ and let σ be v-greedy for $(V, \mathbb{T})$, so that $Wv = T_\sigma^m v$. By Theorem 3.1.5,

σ is Fv-greedy for $(\hat{V}, \hat{\mathbb{T}})$. Hence $\hat{W}Fv = \hat{T}_\sigma^m Fv = FT_\sigma^m v = FWv$. This proves that (3.6) holds. Similarly, continuing to assume that σ is v-greedy for $(V, \mathbb{T})$, we have $Hv = v_\sigma$ and, because σ is Fv-greedy for $(\hat{V}, \hat{\mathbb{T}})$, we also have $\hat{H}Fv = \hat{v}_\sigma$. As a result, $F^{-1}\hat{H}Fv = F^{-1}\hat{v}_\sigma = F^{-1}Fv_\sigma = v_\sigma = Hv$. Hence (3.7) also holds.

Regarding (iii)–(v), we prove only (v), since the remaining arguments are similar. Suppose that HPI converges for $(V, \mathbb{T})$ and fix $\hat{v} \in \hat{V}_U$. Then $v := F^{-1}\hat{v}$ is in V_U, since $v = F^{-1}\hat{v} \preceq F^{-1}\hat{T}\hat{v} = TF\hat{v} = Tv$. As a result, we have $H^n v \uparrow v_\top$. But then $FH^n v \uparrow Fv_\top = \hat{v_\top}$, where $\uparrow$ is by Exercise 3.1.6 and the equality is by (3.5). Since H and $\hat{H}$ are conjugate under F, we also have $FH^n v = \hat{H}^n Fv$. Combining the last two equalities gives us $\hat{H}^n \hat{v} = \hat{H}^n Fv = FH^n v \uparrow \hat{v_\top}$. As $\hat{v}$ was chosen arbitrarily from $\hat{V}_U$, we see that HPI converges for $(\hat{V}, \hat{\mathbb{T}})$. □

3.1.1.5 The Anti-Isomorphic Case

In this section, we switch to studying anti-isomorphic ADPs. In doing so, we will consider minimization as well as maximization. We follow the notational conventions and terminology introduced in §1.4.2. Most readers will find it helpful to review that section before reading this one.

Let $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ be ADPs with the same policy set. In line with the notation in §3.1.1.4, we let

- $T_\perp$ (resp., $\hat{T}_\perp$) be the Bellman min-operator of $(V, \mathbb{T})$ (resp., $(\hat{V}, \hat{\mathbb{T}})$),
- $v_\perp$ (resp., $\hat{v}_\perp$) be the min-value function of $(V, \mathbb{T})$ (resp., $(\hat{V}, \hat{\mathbb{T}})$),
- $H_\perp$ (resp., $\hat{H}_\perp$) be the Howard policy min-operator of $(V, \mathbb{T})$ (resp., $(\hat{V}, \hat{\mathbb{T}})$), and
- $W_\perp$ (resp., $\hat{W}_\perp$) be the optimistic policy min-operator of $(V, \mathbb{T})$ (resp., $(\hat{V}, \hat{\mathbb{T}})$).

As was the case in §1.4.2, we enhance clarity by adding a "max-" prefix to previously introduced definitions that pertain to maximization. For example,

- "optimal policies" will be referred to as "max-optimal policies",
- the "Bellman equation" will be referred to as the "Bellman max-equation",
- the "Bellman operator" will be referred to as the "Bellman max-operator",

and so on.

We call $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ **anti-isomorphic** under F if these two ADPs have the same policy set Σ and, in addition, F is an anti-isomorphism from V to $\hat{V}$ such that (3.2) holds.

We can also expressed this relationship in terms of isomorphisms and duality of ADPs, as defined in §1.4.2.2:

EXERCISE 3.1.12. Show that $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are anti-isomorphic under F if and only if $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})^\partial$ are isomorphic under F. (Hint: See Exercise 3.1.7).

Here is an optimality result for anti-isomorphic ADPs that parallels Theorem 3.1.5.

Theorem 3.1.8. *If $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are anti-isomorphic under F, then*

(i) *σ is v-max-greedy for $(V, \mathbb{T})$ if and only if σ is Fv-min-greedy for $(\hat{V}, \hat{\mathbb{T}})$,*

(ii) *$(V, \mathbb{T})$ is max-regular if and only if $(\hat{V}, \hat{\mathbb{T}})$ is min-regular,*

(iii) *$(V, \mathbb{T})$ is well-posed if and only if $(\hat{V}, \hat{\mathbb{T}})$ is well-posed,*

(iv) *$(V, \mathbb{T})$ is order stable if and only if $(\hat{V}, \hat{\mathbb{T}})$ is order stable, and*

(v) *σ is max-optimal for $(V, \mathbb{T})$ if and only if σ is min-optimal for $(\hat{V}, \hat{\mathbb{T}})$.*

Proof. Let $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ be anti-isomorphic, so that $(V, \mathbb{T})$ is isomorphic to $(\hat{V}, \hat{\mathbb{T}})^\partial$ (Exercise 3.1.12). Regarding (i), suppose that σ is v-max-greedy for $(V, \mathbb{T})$. Then, by Theorem 3.1.5, σ is Fv-max-greedy for $(\hat{V}, \hat{\mathbb{T}})^\partial$. Applying Exercise 1.4.2 on page 40, we see that σ is Fv-min-greedy for $(\hat{V}, \hat{\mathbb{T}})$. The proof of the reverse implication is analogous. This proves (i) and the proofs for the remaining claims are similar. Details are left to the reader. □

Theorem 3.1.9. *Let $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ be well-posed ADPs, where $(V, \mathbb{T})$ is max-regular and $(\hat{V}, \hat{\mathbb{T}})$ is min-regular. If $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are anti-isomorphic under F, then*

(i) *the Bellman operators T and $\hat{T}_\perp$ are connected via*

$$F \circ T = \hat{T}_\perp \circ F \text{ on } V, \tag{3.8}$$

(ii) *when they exist, the value functions for $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are related by $\hat{v}_\perp = Fv_\top$,*

(iii) *the fundamental max-optimality properties hold for $(V, \mathbb{T})$ if and only if the fundamental min-optimality properties hold for $(\hat{V}, \hat{\mathbb{T}})$.*

Proof. Let $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ be anti-isomorphic under F, so that $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})^\partial$ are isomorphic under F. Theorem 3.1.5 implies that $F \circ T = \hat{T}^\partial \circ F$, where $\hat{T}^\partial$ is the Bellman max-operator of $(\hat{V}, \hat{\mathbb{T}})^\partial$. Exercise 1.4.2 on page 40 gives $\hat{T}^\partial = \hat{T}_\perp$. Combining these results gives (3.8).

Regarding (ii), since $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})^\partial$ are isomorphic under F, Theorem 3.1.5 yields $(\hat{v}_\top)^\partial = Fv_\top$. Applying Exercise 1.4.2 again, we have $(\hat{v}_\top)^\partial = \hat{v}_\perp$. Hence $\hat{v}_\perp = Fv_\top$.

Regarding (iii), Theorem 3.1.6 implies that the fundamental max-optimality properties hold for $(V, \mathbb{T})$ if and only if these same max-optimality properties hold for $(\hat{V}, \hat{\mathbb{T}})^\partial$. Theorem 1.4.3 on page 41 tells us that the fundamental max-optimality properties hold for $(\hat{V}, \hat{\mathbb{T}})^\partial$ if and only if the fundamental min-optimality properties hold for $(\hat{V}, \hat{\mathbb{T}})$. □

Now we consider convergence of algorithms in the anti-isomorphic case.

Theorem 3.1.10. *Let $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ be regular, well-posed ADPs. If $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are anti-isomorphic under F, then*

(i) *the optimistic policy operators W and $\hat{W}$ obey*

$$F \circ W = \hat{W}_\perp \circ F \text{ on } V, \text{ and} \tag{3.9}$$

(ii) *the Howard policy operators H and $\hat{H}$ obey*

$$F \circ H = \hat{H}_\perp \circ F \text{ on } V. \tag{3.10}$$

Moreover, if the fundamental optimality properties hold for one and hence both of these ADPs, then the following statements are true.

(iii) *max-VFI converges for $(V, \mathbb{T})$ if and only if min-VFI converges for $(\hat{V}, \hat{\mathbb{T}})$,*

(iv) *max-OPI converges for $(V, \mathbb{T})$ if and only if min-OPI converges for $(\hat{V}, \hat{\mathbb{T}})$, and*

(v) *max-HPI converges for $(V, \mathbb{T})$ if and only if min-HPI converges for $(\hat{V}, \hat{\mathbb{T}})$.*

Proof. Let $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ be anti-isomorphic under F, so that $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})^\partial$ are isomorphic under F. Theorem 3.1.5 implies that $F \circ W = \hat{W}^\partial \circ F$, where $\hat{W}^\partial$ is the optimistic policy max-operator of $(\hat{V}, \hat{\mathbb{T}})^\partial$. Exercise 1.4.2 on page 40 gives $\hat{W}^\partial = \hat{W}_\perp$. Combining these results gives (3.9). The proof of (3.10) is similar.

Regarding (iii), if max-VFI converges for $(V, \mathbb{T})$, then by Theorem 3.1.10, max-VFI converges for $(\hat{V}, \hat{\mathbb{T}})^\partial$. But then min-VFI converges for $(\hat{V}, \hat{\mathbb{T}})$, by Theorem 1.4.3 on page 41. The proof of the converse implication is symmetric, and the proofs of (iv) and (v) are similar. □

3.1.2 Applications

Roadmap.

3.1.2.1 Application: Epstein–Zin Utility

Epstein–Zin utility is perhaps the best known and most popular specification of recursive preferences in economics and finance. (See Chapter 7 of Sargent and Stachurski (2025a) for more on recursive preferences.) In this section we study a relatively simple version of the Epstein–Zin model with finite states and actions. While the simple version can be generalized, we maintain it in order to keep focus on isomorphic and anti-isomorphic relationships.

The problem that we study concerns a variation of the finite state MDP from §2.1.1, where the Bellman equation is modified to

$$v(x) = \max_{a \in \Gamma(x)} \left\{ r(x,a)^\alpha + \beta \left[(Lv)(x,a) \right]^\alpha \right\}^{1/\alpha}, \tag{3.11}$$

with

$$(Lv)(x,a) := \left(\sum_{x'} v(x')^\gamma P(x,a,x') \right)^{1/\gamma}.$$

Here X, A, r, Γ, β, and P are all as in §2.1.1, while G is the feasible state-action pairs. If $\alpha = \gamma = 1$, then this Epstein–Zin model reduces to an ordinary finite state MDP. In this section, however, we allow α and γ to take any nonzero values. The parameter α determines elasticity of substitution between current and future payoffs, while γ parameterizes risk-aversion when facing uncertainty over intertemporal outcomes.

Using the notation

$$r_\sigma(x) := r(x, \sigma(x)) \quad \text{and} \quad (L_\sigma v)(x) := (Lv)(x, \sigma(x)),$$

we can write a corresponding policy operator T_σ as

$$T_\sigma \, v = \left\{ r_\sigma^\alpha + \beta \, (L_\sigma \, v)^\alpha \right\}^{1/\alpha}. \tag{3.12}$$

Let $\mathbb{T}$ be the set of all such T_σ.

We assume that r is strictly positive, so that T_σ maps $(0,\infty)^{\mathsf{X}}$ into itself. Sargent and Stachurski (2025a) establish optimality properties for this specification when P_σ is irreducible for all $\sigma \in \Sigma$. Here we drop this assumption and establish the same optimality properties.

To this end, let $\theta = \gamma/\alpha$. Fix $\varepsilon > 0$ with $\min r^\alpha - \varepsilon > 0$. Consider the constant functions $v_1 = m_1 \wedge m_2$ and $v_2 = m_1 \vee m_2$, where

$$m_1 := \left(\frac{\min r^\alpha - \varepsilon}{1-\beta} \right)^\theta \quad \text{and} \quad m_2 := \left(\frac{\max r^\alpha + \varepsilon}{1-\beta} \right)^\theta.$$

Let $\hat{V} = [v_1, v_2]$. Let F be defined by

$$F\,v = v^\gamma \qquad \text{with } v \in (0,\infty)^{\mathsf{X}},$$

where the exponent γ is applied pointwise to v, and set

$$V := F^{-1}\hat{V} = \{v \in (0,\infty)^{\mathsf{X}} : v_1 \leqslant v^\gamma \leqslant v_2\}. \tag{3.13}$$

We are interested in optimality properties of $(V, \mathbb{T})$. While we can try to tackle this ADP directly, the arguments become significantly easier after a transformation. To pursue this path, we introduce the auxillary ADP $(\hat{V}, \hat{\mathbb{T}})$ with $\hat{V}$ as defined above and

$$\hat{T}_\sigma\, v = \left\{ r_\sigma^\alpha + \beta\,(P_\sigma\, v)^{1/\theta} \right\}^\theta. \tag{3.14}$$

In the next exercise, $f \ll g$ means $f(x) < g(x)$ for all x, as in §1.3.3.4.

Exercise 3.1.13. Verify the following inequalities: For v_1 and v_2 as defined above, we have $v_1 \ll \hat{T}_\sigma\, v_1$ and $\hat{T}_\sigma\, v_2 \ll v_2$.

Exercise 3.1.14. Confirm the following statements.

(i) If $0 < \theta \leqslant 1$, then $\hat{T}_\sigma$ is convex on $\hat{V}$.

(ii) If $\theta < 0$ or $1 \leqslant \theta$, then $\hat{T}_\sigma$ is concave on $\hat{V}$.

The results from Exercises 3.1.13 and 3.1.14 allow us to establish optimality properties for the auxillary ADP $(\hat{V}, \hat{\mathbb{T}})$.

Lemma 3.1.11. *The following statements are both true.*

(i) *The fundamental max-optimality results hold for* $(\hat{V}, \hat{\mathbb{T}})$ *and max-VFI, max-OPI, and max-HPI all converge.*

(ii) *The fundamental min-optimality results hold for* $(\hat{V}, \hat{\mathbb{T}})$ *and min-VFI, min-OPI, and min-HPI all converge.*

Proof. Fix $\sigma \in \Sigma$. From Exercise 3.1.13 we have $v_1 \ll \hat{T}_\sigma\, v \ll v_2$ for all $v \in \hat{V}$. Also, by Exercise 3.1.14, T_σ is either concave or convex. Theorem 1.3.10 now implies that the fundamental max-optimality properties hold for $(\hat{V}, \hat{\mathbb{T}})$ and max-VFI, max-OPI, and max-HPI all converge. (See, in particular, Corollary 1.3.11 on page 35.)

Regarding the fundamental min-optimality results, it is simple to confirm that $(\hat{V}, \hat{\mathbb{T}})$ is min-regular, so $(\hat{V}, \hat{\mathbb{T}})^\partial$ is max-regular (Theorem 3.1.8). Also, the proof of Theorem 1.3.10 shows that $(\hat{V}, \hat{\mathbb{T}})$ is globally stable (as a result of Du's Theorem), so $(\hat{V}, \hat{\mathbb{T}})^\partial$ is also globally stable (as global stability is not affected by taking the dual). Since V is an order interval in $\mathbb{R}^\mathsf{X}$ and X is finite, Theorem 1.3.4 applies. Hence $(\hat{V}, \hat{\mathbb{T}})^\partial$ obeys the fundamental max-optimality properties and max-VFI, max-OPI, and max-HPI all converge. Applying Theorem 1.4.3, we see that the second statement in Lemma 3.1.11 is true. □

Exercise 3.1.15. Show that, for all $\sigma \in \Sigma$, we have $F \circ T_\sigma = \hat{T}_\sigma \circ F$ on V.

The next result follows easily from the conclusion of Exercise 3.1.15.

Lemma 3.1.12. *The following statements are true:*

(i) *If $\gamma > 0$, then $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are isomorphic.*

(ii) *If $\gamma < 0$, then $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are anti-isomorphic.*

Proof. If $\gamma < 0$, then F is an order anti-isomorphism from V to $\hat{V}$. (Obviously F is order-reversing. Also, F is clearly one-to-one and, by construction, F maps V onto $\hat{V}$.) From this fact and Exercise 3.1.15, the ADPs $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are anti-isomorphic. If $\gamma > 0$, then F is an order isomorphism from V to $\hat{V}$, so, applying the result of Exercise 3.1.15 gain, $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are isomorphic. □

We are now ready to state and prove the main result of this section.

Proposition 3.1.13. *The fundamental max-optimality properties hold for $(V, \mathbb{T})$. In addition, max-VFI, max-OPI, and max-HPI all converge.*

Proof of Proposition 3.1.13. Suppose first that $\gamma > 0$. Then, by Lemma 3.1.12, $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are isomorphic. Moreover, by Lemma 3.1.11, the fundamental max-optimality results hold for $(\hat{V}, \hat{\mathbb{T}})$ and max-VFI, max-OPI, and max-HPI all converge. Theorems 3.1.6 and 3.1.7 now imply that the same results hold for $(V, \mathbb{T})$.

Next, suppose that $\gamma < 0$. Then, by Lemma 3.1.12, $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ are anti-isomorphic. Moreover, by Lemma 3.1.11, the fundamental min-optimality results hold for $(\hat{V}, \hat{\mathbb{T}})$ and min-VFI, min-OPI, and min-HPI all converge. Applying Theorems 3.1.9 and 3.1.10, we see that the fundamental max-optimality results hold for $(V, \mathbb{T})$ and max-VFI, max-OPI, and max-HPI all converge. □

The relationship between $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ allows us to use either one to solve for an optimal policy. For example, if $\gamma < 0$, then, by Theorem 3.1.8, any min-optimal policy for $(\hat{V}, \hat{\mathbb{T}})$ will be max-optimal for $(V, \mathbb{T})$. Hence we can solve for the Epstein–Zin max-optimal policy either by directly solving $(V, \mathbb{T})$ or by solving $(\hat{V}, \hat{\mathbb{T}})$ for a min-optimal policy. The best choice depends on computational simplicity and numerical stability.

3.2 Factored Dynamic Programs

Next we introduce an asymmetric relationship between ADPs referred to as as subordination. In essence, a subordinate ADP is an ADP that is derived from another ADP via some rearrangement of the Bellman equation. In the applications we consider, the associated transformations are not bijective, which differentiates subordination from the isomorphic relationships considered in §3.1. Nonetheless, we show that subordination provides connections between ADPs in terms of optimality.

Often, the lack of bijections referred to in the previous paragraph occurs because one dynamic program evolves in a higher dimensional space than the other. Connecting optimality properties of such dynamic programs is valuable because the lower dimensional program can usually be solved more efficiently than the higher dimensional one.

This section is related to the dynamic programs associated with the modified Bellman equations we introduced in Chapter 5 of Sargent and Stachurski (2025a). Relative to that theory, the exposition below is more concise and more general. As a result, we can easily cover additional variations on the MDP Bellman equation, of which there are many, as well as studying relationships between ADPs beyond the traditional MDP setting.

3.2.1 Definition and Properties

We begin by introducing a similarity notion somewhat related to conjugacy. Then we define factored dynamic programs and investigate basic properties.

3.2.1.1 Semiconjugacy

Let (V, S) and $(\hat{V}, \hat{S})$ be dynamical systems, where V and $\hat{V}$ are posets. In this setting, we call (V, S) and $(\hat{V}, \hat{S})$ **mutually semiconjugate under F, G** when there exist order-preserving maps $F\colon V \to \hat{V}$ and $G\colon \hat{V} \to V$ such that

$$S = G \circ F \text{ on } V \qquad \text{and} \qquad \hat{S} = F \circ G \text{ on } \hat{V}. \tag{3.15}$$

The "semiconjugate" terminology comes from the fact that, when (3.15) holds,

$$F \circ S = \hat{S} \circ F \quad \text{and} \quad G \circ \hat{S} = S \circ G. \tag{3.16}$$

In particular, if either F or G is an order isomorphism, then S and $\hat{S}$ are order conjugate.

Exercise 3.2.1. Confirm that (3.15) implies (3.16).

Like order conjugacy, mutual semiconjugacy can be used to derive useful relationships between dynamical systems. The next lemma lists relationships that will be helpful when we turn to dynamic programming.

Lemma 3.2.1. *Let (V, S) and $(\hat{V}, \hat{S})$ be mutually semiconjugate under F, G. In this setting,*

(i) *if v is a fixed point of S in V, then Fv is a fixed point of $\hat{S}$ in $\hat{V}$.*

(ii) *if $\hat{v}$ is a fixed point of $\hat{S}$ in $\hat{V}$, then $G\hat{v}$ is a fixed point of S in V.*

(iii) *S has a unique fixed point in V if and only if $\hat{S}$ has a unique fixed point in $\hat{V}$.*

(iv) *S is order stable on V if and only if $\hat{S}$ is order stable on $\hat{V}$.*

Proof. Let (V, S) and $(\hat{V}, \hat{S})$ be as stated. If v is a fixed point of S in V, then $\hat{S}Fv = FSv = Fv$, so Fv is a fixed point of $\hat{S}$ in $\hat{V}$. Similarly, if $\hat{v}$ is a fixed point of $\hat{S}$ in $\hat{V}$, then $SG\hat{v} = G\hat{S}\hat{v} = G\hat{v}$, so $G\hat{v}$ is a fixed point of S in V. This proves (i)–(ii).

Regarding (iii), suppose that v is the only fixed point of S in v. We know that Fv is a fixed point of $\hat{S}$ in $\hat{V}$. Suppose in addition that $\hat{v}$ is fixed for $\hat{S}$. Then $FG\hat{v} = \hat{v}$ and hence $GFG\hat{v} = G\hat{v}$, or $SG\hat{v} = G\hat{v}$. Since v is the only fixed point of S in V, we have $G\hat{v} = v$. Applying F gives $\hat{S}\hat{v} = Fv$. But $\hat{v}$ is fixed for $\hat{S}$, so $\hat{v} = Fv$. This shows that $\hat{S}$ has exactly one fixed point in $\hat{V}$. The reverse implication holds by symmetry.

Regarding (iv), suppose that S is order stable on V, with unique fixed point $v \in V$. Then, by the preceding argument, Fv is the unique fixed point of $\hat{S}$ in $\hat{V}$. The map $\hat{S}$ is upward stable because if $\hat{v} \in \hat{V}$ and $\hat{v} \preceq \hat{S}\hat{v}$, then $G\hat{v} \preceq G\hat{S}\hat{v} = SG\hat{v}$ and so, by upward stability of S, $G\hat{v} \preceq v$. Applying F gives $\hat{S}\hat{v} \preceq Fv$, so $\hat{S}$ is upward stable. The proof of downward stability is similar. Hence $\hat{S}$ is order stable on $\hat{V}$. The reverse implication holds by symmetry. □

3.2.1.2 FDPs

A **factored dynamic program** (FDP) is a tuple $(V, \hat{V}, \mathbb{G}, F)$ where

(i) V and $\hat{V}$ are nonempty posets,

(ii) F is an order-preserving map from V to $\hat{V}$, and

(iii) $\mathbb{G} := \{G_\sigma\}_{\sigma \in \Sigma}$ is a family of order-preserving maps from $\hat{V}$ to V, and

(iv) the set $\{G_\sigma\, \hat{v}\}_{\sigma \in \Sigma}$ has a greatest element for every $\hat{v} \in \hat{V}$.

Given $(V, \hat{V}, \mathbb{G}, F)$ we set

$$G\, \hat{v} := \bigvee_\sigma G_\sigma\, \hat{v} \qquad (\hat{v} \in \hat{V}), \tag{3.17}$$

which is well-defined by (iv). In addition, we define $(V, \mathbb{T})$, where the policy operators in $\mathbb{T}$ take the form

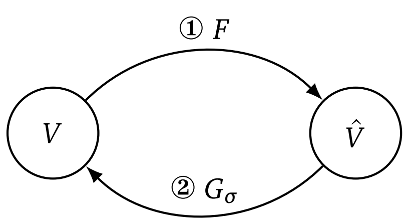
$$T_\sigma = G_\sigma \circ F \text{ for all } \sigma \in \Sigma.$$

We call $(V, \mathbb{T})$ the **primary ADP** generated by $(V, \hat{V}, \mathbb{G}, F)$.

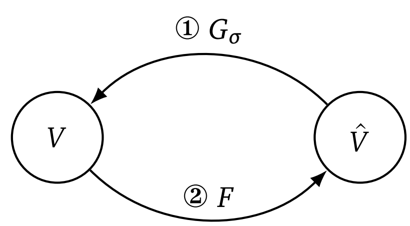
The factored dynamic program $(V, \hat{V}, \mathbb{G}, F)$ also produces an second ADP $(\hat{V}, \hat{\mathbb{T}})$, where the policy operators in $\hat{\mathbb{T}}$ take the form

$$\hat{T}_\sigma = F \circ G_\sigma \qquad \text{for all } \sigma \in \Sigma,$$

We call $(\hat{V}, \hat{\mathbb{T}})$ the **subordinate ADP** generated by $(V, \hat{V}, \mathbb{G}, F)$. The figure below illustrates, with the numbers indicating the order in which mappings are applied.



primary: $T_\sigma = G_\sigma \circ F$



subordinate: $\hat{T}_\sigma = F \circ G_\sigma$

3.2.1.3 Example: From MDPs to Q-Factors

Consider the finite-state MDP environment described in §2.1.1. Using these primitives, we form a factored dynamic program $(V, \hat{V}, \mathbb{G}, F)$ by setting

$$(Fv)(x,a) = r(x,a) + \beta \sum_{x'} v(x') P(x,a,x') \qquad \left(v \in \mathbb{R}^{\mathsf{X}}\right) \tag{3.18}$$

and

$$(G_\sigma f)(x) = f(x, \sigma(x)) \qquad \left(\sigma \in \Sigma,\ f \in \mathbb{R}^{\mathsf{G}}\right),$$

with $V := \mathbb{R}^{\mathsf{X}}$ and $\hat{V} := \mathbb{R}^{\mathsf{G}}$. As required, F maps V to $\hat{V}$, and G_σ maps $\hat{V}$ to V, and both are are order preserving. Also, fixing $f \in \hat{V}$ and choosing σ such that $\sigma(x) \in \operatorname{argmax}_{a \in \Gamma(x)} f(x, a)$, we verify the existence of a policy σ with $G_\tau f \leqslant G_\sigma f$ for all $\tau \in \Sigma$. Hence $(V, \hat{V}, \mathbb{G}, F)$ is a factored dynamic program, as claimed.

The primary ADP $(V, \mathbb{T})$ generated by $(V, \hat{V}, \mathbb{G}, F)$ is produced by setting $T_\sigma = G_\sigma \circ F$ for each σ, which gives

$$(T_\sigma\, v)(x) = (G_\sigma\, F\, v)(x) = r(x, \sigma(x)) + \beta \sum_{x'} v(x') P(x, \sigma(x), x').$$

Thus, $(V, \mathbb{T})$ is nothing but the standard ADP generated from an MDP (see §2.1.1.2).

The subordinate ADP $(\hat{V}, \hat{\mathbb{T}})$ generated by $(V, \hat{V}, \mathbb{G}, F)$ is produced by setting $\hat{T}_\sigma = F \circ G_\sigma$ for each σ, which gives

$$(\hat{T}_\sigma\, f)(x, a) = (F\, G_\sigma\, f)(x, a), = r(x, a) + \beta \sum_{x'} f(x', \sigma(x')) P(x, a, x').$$

This map $\hat{T}_\sigma$ is identical to the Q-factor policy operator S_σ we constructed in (2.36) (page 74). Thus, $(\hat{V}, \hat{\mathbb{T}})$ is the same as the Q-factor ADP we examined in §2.3.1.1 (where the ADP was written as $(\mathbb{R}^{\mathsf{G}}, \mathbb{S})$).

3.2.1.4 Expected Value Formulations

We consider again the finite MDP setting discussed in §3.2.1.3, but now with

$$(Fv)(x, a) = \sum_{x'} v(x') P(x, a, x') \tag{3.19}$$

and

$$(G_\sigma\, g)(x) = r(x, \sigma(x)) + \beta g(x, \sigma(x)). \tag{3.20}$$

As before, $V := \mathbb{R}^{\mathsf{X}}$ and $\hat{V} := \mathbb{R}^{\mathsf{G}}$. Clearly F is an order-preserving map from V to $\hat{V}$, while each G_σ is an order preserving map from $\hat{V}$ to V.

EXERCISE 3.2.2. Prove the existence of a policy σ with $G_\tau\, g \leqslant G_\sigma g$ for all $\tau \in \Sigma$.

While the factored dynamic program $(V, \hat{V}, \mathbb{G}, F)$ has been modified, the primary ADP $(V, \mathbb{T})$ is, once again, the standard ADP generated from an MDP. This holds because

$T_\sigma = G_\sigma \circ F$, which translates to

$$(T_\sigma v)(x) = (G_\sigma F v)(x) = r(x, \sigma(x)) + \beta \sum_{x'} v(x') P(x, \sigma(x), x').$$

For the subordinate ADP $(\hat{V}, \hat{\mathbb{T}})$, the policy operators are given by

$$(\hat{T}_\sigma g)(x, a) = (F G_\sigma g)(x, a) = \sum_{x'} \{r(x', \sigma(x')) + \beta g(x', \sigma(x'))\} P(x, a, x').$$

The Bellman equation for $(\hat{V}, \hat{\mathbb{T}})$ is

$$g(x, a) = \sum_{x'} \max_{a' \in \Gamma(x')} \{r(x', a') + \beta g(x', a')\} P(x, a, x'),$$

which is a version of the structural estimation Bellman equation discussed in §2.3.2.

3.2.1.5 Risk-Sensitive Q-Factors

Consider again the the risk-sensitive MDP environment from §2.2.1.5. Using these primitives, we set

$$(Fv)(x, a) = r(x, a) + \frac{\beta}{\theta} \ln \left[\sum_{x'} \exp(\theta v(x')) P(x, a, x') \right]$$

and

$$(G_\sigma f)(x) = f(x, \sigma(x)) \qquad \left(\sigma \in \Sigma, \; f \in \mathbb{R}^{\mathsf{G}}\right),$$

with $V := \mathbb{R}^{\mathsf{X}}$ and $\hat{V} := \mathbb{R}^{\mathsf{G}}$. As required, F maps V to $\hat{V}$, G_σ maps $\hat{V}$ to V, and both are are order preserving. An argument identical to the one given in §3.2.1.3 proves that, for each $f \in \hat{V}$, there exists a policy σ with $G_\tau f \leqslant G_\sigma f$ for all $\tau \in \Sigma$. Hence $(V, \hat{V}, \mathbb{G}, F)$ is a factored dynamic program.

The policy operators for the primary ADP $(V, \mathbb{T})$ generated by $(V, \hat{V}, \mathbb{G}, F)$ take the form

$$(T_\sigma v)(x) = (G_\sigma F v)(x) = r(x, \sigma(x)) + \frac{\beta}{\theta} \ln \left[\sum_{x'} \exp(\theta v(x')) P(x, \sigma(x), x') \right].$$

This is the same as the policy operators for the risk-sensitive RDP setting of §2.2.1.5.

For the subordinate ADP $(\hat{V}, \hat{\mathbb{T}})$, the policy operators are given by

$$(\hat{T}_\sigma\, v)(x, a) = (FG_\sigma\, v)(x, a) = r(x, a) + \frac{\beta}{\theta} \ln \left[\sum_{x'} \exp\left[\theta f(x', \sigma(x'))\right] P(x, a, x') \right].$$

Thus, $(\hat{V}, \hat{\mathbb{T}})$ is identical to the risk-sensitive Q-factor ADP from §2.3.1.2.

3.2.2 Optimality

In this section our main aim is to study the extent to which optimality properties are transferred under subordination. We begin with some preliminary results.

3.2.2.1 Preliminaries

Let $(V, \hat{V}, \mathbb{G}, F)$ be a factored dynamic program. We now state some preliminary results concerning the ADPs generated by $(V, \hat{V}, \mathbb{G}, F)$.

Lemma 3.2.2. *If $(V, \mathbb{T})$ is the primary ADP generated by $(V, \hat{V}, \mathbb{G}, F)$, then*

(i) *$(V, \mathbb{T})$ is regular,*

(ii) *the Bellman operator T obeys $T = G \circ F$ on V, and*

(iii) *σ is v-greedy for $(V, \mathbb{T})$ if and only if $G_\sigma F v = GFv$.*

Proof. Regarding regularity of $(V, \mathbb{T})$, fix $v \in V$. By definition, there exists a $\sigma \in \Sigma$ such that $G_\tau\, Fv \preceq G_\sigma\, Fv$ for all $\tau \in \Sigma$. But then $T_\tau\, v \preceq T_\sigma\, v$ for all $\tau \in \Sigma$, so σ is v-greedy. Regarding (ii), for any $v \in V$ we have $Tv = \bigvee_\sigma T_\sigma\, v = \bigvee_\sigma G_\sigma\, Fv = GFv$. Claim (iii) follows from (ii) by Lemma 1.2.1 on page 13. □

Lemma 3.2.3. *If $(\hat{V}, \hat{\mathbb{T}})$ is the subordinate ADP generated by $(V, \hat{V}, \mathbb{G}, F)$, then*

(i) *$(\hat{V}, \hat{\mathbb{T}})$ is regular,*

(ii) *the Bellman operator $\hat{T}$ obeys $\hat{T} = F \circ G$ on $\hat{V}$, and*

(iii) *if $G_\sigma\, \hat{v} = G\hat{v}$, then σ is $\hat{v}$-greedy for $(\hat{V}, \hat{\mathbb{T}})$.*

Proof. Regarding (i), fix $\hat{v} \in \hat{V}$ and choose $\sigma \in \Sigma$ such that $G_\tau\, \hat{v} \preceq G_\sigma\, \hat{v}$ for all $\tau \in \Sigma$. Since F is order-preserving, we then have $\hat{T}_\tau\, \hat{v} \preceq \hat{T}_\sigma\, \hat{v}$ for all $\tau \in \Sigma$. Hence σ is $\hat{v}$-greedy and $(\hat{V}, \hat{\mathbb{T}})$ is regular. Claim (ii) also holds because

$$\hat{T}\hat{v} = \bigvee_\sigma FG_\sigma\, \hat{v} = F \bigvee_\sigma G_\sigma\, \hat{v} = FG\hat{v}.$$

(The second equality in the last display is valid because $\{G_\sigma \hat{v}\}_{\sigma \in \Sigma}$ has a greatest element.) Regarding (iii), if $G_\sigma \hat{v} = G\hat{v}$, then applying F to both sides gives $\hat{T}_\sigma \hat{v} = \hat{T}\hat{v}$. Hence σ is $\hat{v}$-greedy for $(\hat{V}, \hat{\mathbb{T}})$. □

Let $(V, \hat{V}, \mathbb{G}, F)$ be a factored dynamic program with generated ADP $(V, \mathbb{T})$ and subordinate ADP $(\hat{V}, \hat{\mathbb{T}})$. The following lemma restates results we have already proved in order to emphasize them.

Lemma 3.2.4. *The policy operators of $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$ obey*

$$T_\sigma = G_\sigma \circ F \quad \text{and} \quad \hat{T}_\sigma = F \circ G_\sigma \tag{3.21}$$

for all $\sigma \in \Sigma$, while the Bellman operators are related by

$$T = G \circ F \qquad \text{and} \qquad \hat{T} = F \circ G. \tag{3.22}$$

As a result,

(i) *each pair of policy systems (V, T_σ) and $(\hat{V}, \hat{T}_\sigma)$ is mutually semiconjugate under F, G_σ and*

(ii) *the Bellman operator systems (V, T) and $(\hat{V}, \hat{T})$ are mutually semiconjugate under F, G.*

Proof. These results follow either from the definitions or from the Bellman operator results proved in Lemmas 3.2.2 and 3.2.3. □

The semiconjugacy results in Lemma 3.2.4 imply useful similarity properties for the two ADPs. The next lemma helps to illustrate.

Lemma 3.2.5. *The following relationships hold:*

(i) *$(\hat{V}, \hat{\mathbb{T}})$ is well-posed if and only if $(V, \mathbb{T})$ is well-posed, and*

(ii) *$(\hat{V}, \hat{\mathbb{T}})$ is order stable if and only if $(V, \mathbb{T})$ is order stable.*

In either case, the σ-value functions are linked by

$$\hat{v}_\sigma = F v_\sigma \quad \text{and} \quad v_\sigma = G_\sigma \hat{v}_\sigma \quad \text{for all} \quad \sigma \in \Sigma. \tag{3.23}$$

Proof. All claims follow from Lemma 3.2.1 and the observation that (V, T_σ) and $(\hat{V}, \hat{T}_\sigma)$ are mutually semiconjugate under F, G_σ at every $\sigma \in \Sigma$. □

Note that (3.22) implies that (V, T) and $(\hat{V}, \hat{T})$ are mutually semiconjugate under F, G. This fact allows us to connect optima for the two ADPs and we use it repeatedly below.

3.2.2.2 Optimality

Now we are ready to study the extent to which optimality properties are transferred under subordination. As before, the context is that $(V, \hat{V}, \mathbb{G}, F)$ is a factored dynamic program with primary ADP $(V, \mathbb{T})$ and subordinate ADP $(\hat{V}, \hat{\mathbb{T}})$. The symbols T and $\hat{T}$ denote their respective Bellman operators. When they exist,

$$v_\top = \bigvee_\sigma v_\sigma \quad \text{and} \quad \hat{v}_\top = \bigvee_\sigma \hat{v}_\sigma$$

will represent their respective value functions.

We begin with our main optimality result for factored dynamic programs and the two ADPs they generate.

Theorem 3.2.6. *The following statements are equivalent:*

(a) *$v_\top$ exists and is the unique fixed point of T in V.*

(b) *$\hat{v}_\top$ exists and is the unique fixed point of $\hat{T}$ in $\hat{V}$.*

If either and hence both of these statements are true, then

(i) *the value functions obey*

$$v_\top = G\,\hat{v}_\top \quad \textit{and} \quad \hat{v}_\top = F\,v_\top, \tag{3.24}$$

(ii) *the fundamental optimality properties hold for $(V, \mathbb{T})$,*

(iii) *the fundamental optimality properties hold for $(\hat{V}, \hat{\mathbb{T}})$,*

(iv) *if $G_\sigma\,\hat{v}_\top = G\,\hat{v}_\top$, then σ is optimal for $(V, \mathbb{T})$, and*

(v) *if σ is optimal for $(V, \mathbb{T})$, then σ is optimal for $(\hat{V}, \hat{\mathbb{T}})$.*

In the proof of Theorem 3.2.6, we repeatedly use the mutual semiconjugacy results in Lemma 3.2.1 to transfer fixed points from one value space to another.

Proof of Theorem 3.2.6. Suppose that (a) holds, so that $v_\top$ exists and is the unique fixed point of T. Lemma 3.2.4 states that (V, T) and $(\hat{V}, \hat{T})$ are mutually semiconjugate under F, G, so, by Lemma 3.2.1, $Fv_\top$ is the unique fixed point of $\hat{T}$ in $\hat{V}$. We claim that $\hat{v}_\top = Fv_\top$. To see this, observe first that, given $\sigma \in \Sigma$, we have $v_\sigma \preceq v_\top$, so, applying the fixed point translation in (3.23), we get $\hat{v}_\sigma = Fv_\sigma \preceq Fv_\top$. The last inequality becomes an equality if σ is optimal for $(V, \mathbb{T})$. This proves that the supremum $\hat{v}_\top = \bigvee_\sigma \hat{v}_\sigma$ is equal to $Fv_\top$. In particular, $\hat{v}_\top$ exists and is the unique fixed point of $\hat{T}$ in $\hat{V}$.

Now suppose that (b) holds, so that $\hat{v}_\top$ exists and is the unique fixed point of $\hat{T}$ in $\hat{V}$. Since the dynamical systems (V, T) and $(\hat{V}, \hat{T})$ are mutually semiconjugate under F, G, the element $G\hat{v}_\top$ is the unique fixed point of T in V. Thus, to prove (a) we need only show that $v_\top$ exists and $v_\top = G\hat{v}_\top$. First take any $\sigma \in \Sigma$. We have $\hat{v}_\sigma \preceq \hat{v}_\top$ and, by mutual semiconjugacy of (V, T_σ) and $(\hat{V}, \hat{T}_\sigma)$ under F, G_σ, the equality $v_\sigma = G_\sigma \hat{v}_\sigma$. Using these facts together gives $v_\sigma = G_\sigma \hat{v}_\sigma \preceq G_\sigma \hat{v}_\top \preceq G\hat{v}_\top$. This proves that $G\hat{v}_\top$ is an upper bound of V_Σ in V. Thus, we can complete the proof that $v_\top$ exists and $v_\top = G\hat{v}_\top$ by producing a $\sigma \in \Sigma$ with $v_\sigma = G\hat{v}_\top$.

To this end, we choose σ such that $G_\sigma \hat{v}_\top = G\hat{v}_\top$, which is possible by the definition of factored dynamic programs. Bellman's principle of optimality holds for $(\hat{V}, \hat{\mathbb{T}})$ (by (b) and Lemma 1.2.5 on page 18) and, applying F to the previous equality gives $\hat{T}_\sigma \hat{v}_\top = \hat{T}\hat{v}_\top$. Hence σ is optimal for $(\hat{V}, \hat{\mathbb{T}})$, so $\hat{v}_\sigma = \hat{v}_\top$. Combining this equality with $G_\sigma \hat{v}_\top = G\hat{v}_\top$ yields $v_\sigma = G_\sigma \hat{v}_\sigma = G_\sigma \hat{v}_\top = G\hat{v}_\top$. Thus, we have at least one $\sigma \in \Sigma$ with $v_\sigma = G\hat{v}_\top$. We have now shown that $v_\top$ exists and $v_\top = G\hat{v}_\top$, so the proof of (a) is done.

Now suppose that (a) and (b) hold. The arguments above also showed that (3.24) is valid. Also, Lemma 1.2.5 implies that the fundamental optimality properties hold for both $(V, \mathbb{T})$ and $(\hat{V}, \hat{\mathbb{T}})$. Regarding (iv), let $\sigma \in \Sigma$ be such that $G_\sigma \hat{v}_\top = G\,\hat{v}_\top$. Applying (3.24) yields $G_\sigma F v_\top = GFv_\top$, or $T_\sigma v_\top = T\, v_\top$. By Bellman's principle of optimality, σ is optimal for $(V, \mathbb{T})$.

Regarding part (v), let σ be optimal for $(V, \mathbb{T})$. Since $(V, \mathbb{T})$ obeys the fundamental optimality properties, σ is $v_\top$-greedy (i.e., $T_\sigma v_\top = Tv_\top$). Also, by (3.24), we have $\hat{v}_\top = Fv_\top$. Therefore,

$$\hat{T}_\sigma \hat{v}_\top = \hat{T}_\sigma F v_\top = FG_\sigma F v_\top = F\, T_\sigma v_\top = F\, T v_\top = Fv_\top = \hat{v}_\top = \hat{T}\, \hat{v}_\top.$$

Thus, σ is $\hat{v}_\top$-greedy for $(\hat{V}, \hat{\mathbb{T}})$. But Bellman's principle of optimality also holds for $(\hat{V}, \hat{\mathbb{T}})$, so σ is optimal for $(\hat{V}, \hat{\mathbb{T}})$. □

In the next section we consider further applications of Theorem 3.2.6.

3.2.3 Applications

Roadmap.

3.2.3.1 Expected Value ADPs

As a simple application of Theorem 3.2.6, consider the factored dynamic program $(V, \hat{V}, \mathbb{G}, F)$ from §3.2.1.4. We saw in §3.2.1.4 that the primary ADP $(V, \mathbb{T})$ is just the ordinary ADP generated by a finite state MDP. To find optimal policies for $(V, \mathbb{T})$, we can study instead the subordinate ADP $(\hat{V}, \hat{\mathbb{T}})$ also defined in §3.2.1.4. In view of Theorem 3.2.6, we can do this by computing the unique fixed point $g_\top$ of the corresponding Bellman operator $\hat{T}$ in $\hat{V}$ and then finding a policy σ obeying $G_\sigma g_\top = G g_\top$, where $G = \bigvee_\sigma G_\sigma$. By the definition of G_σ in (3.20), this means that we solve for σ satisfying

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} \{r(x,a) + \beta g_\top(x,a)\} \qquad (x \in \mathsf{X}).$$

To see why this approach might be beneficial, let's take the case of the optimal harvest model from §2.2.1.6. We recall that the Bellman operator has the form

$$(Tv)(s,p) = \max_a \left\{ r(s,p,a) + \beta \int v[f(s,a),p']\varphi(\mathrm{d}p') \right\},$$

where $a \in \{0,1\}$,

$$r(s,p,a) := a(ps - m(s)) - (1-a)c \quad \text{and} \quad f(s,a) := q[(1-a)s].$$

3.2.3.2 Subordination in an Epstein–Zin Setting

In this section we consider a special case of the Epstein–Zin ADP $(V, \mathbb{T})$ analyzed in § 3.1.2.1. The special case concerns optimal savings in the presence of an IID endowment process. We will produce a subordinate ADP via a transformation reminiscent of the expected value transformation of an ordinary MDP in §3.2.1.4. We will see that this subordinate ADP is easier to analyze and solve.

We begin with a finite set W of possible wealth values and a finite set E of possible values for the endowment process. (Finiteness helps simplify the exposition and can be replaced by continuity and compactness conditions.) The Bellman equation takes the form

$$v(w,e) = \max_{w' \in \Gamma(w,e)} \left\{ r(w,w',e)^\alpha + \beta \left(\sum_{e'} v(w',e')^\gamma \varphi(e') \right)^{\alpha/\gamma} \right\}^{1/\alpha}.$$

Here $\Gamma(w,e) \subset \mathsf{W}$ is the set of all feasible choices for next period wealth w' given current wealth w and current endowment e. The new endowment e' is drawn independently from distribution φ, which maps E into $[0,1]$.

This model is a special case of the Epstein–Zin ADP from §3.1.2.1. To see this we set $\mathsf{X} := \mathsf{W} \times \mathsf{E}$, with typical element $x = (w,e)$. Let V be the order interval $[v_1^{1/\gamma}, v_2^{1/\gamma}] \subset$

$(0, \infty)^{\mathsf{X}}$ defined in §3.1.2.1 (see (3.13)). With $a = w'$ and $\mathsf{A} = \mathsf{W}$, the Bellman equation from the previous paragraph is a special case of (3.11).

To improve analysis we produce a factored dynamic program where the primary ADP is the model just discussed and the subordinate ADP operates in a lower-dimensional state space. To do so we set V as above,

$$(Fv)(w) = \left\{ \sum_e v(w,e)^\gamma \varphi(e) \right\}^{1/\gamma} \qquad (w \in \mathsf{W}),$$

which maps V to $\hat{V} := F(V)$, and

$$(G_\sigma h)(w,e) = \{r(w, \sigma(w), e)^\alpha + \beta h(\sigma(w))^\alpha\}^{1/\alpha} \qquad ((w,e) \in \mathsf{X}), \tag{3.25}$$

Both F and G_σ are order-preserving. Assuming that $\Gamma(w,e)$ is nonempty at each $(w,e) \in \mathsf{X}$, one easily verifies the existence, for each $h \in \hat{V}$, of a policy σ such that

$$\sigma(w,e) \in \operatorname*{argmax}_{w' \in \Gamma(w,e)} \{r(w, w', e)^\alpha + \beta h(w')^\alpha\}^{1/\alpha} \quad \text{for all } (w,e) \in \mathsf{X}. \tag{3.26}$$

For this σ we have $G_\tau h \leqslant G_\sigma h$ for all $\tau \in \Sigma$. Hence, with $\mathbb{G} = \{G_\sigma\}_{\sigma \in \Sigma}$, the tuple $(V, \hat{V}, F, \mathbb{G})$ is a factored dynamic program.

Exercise 3.2.3. Show that the primary ADP for this factored dynamic program has the form $(V, \mathbb{T})$, where each $T_\sigma \in \mathbb{T}$ is given by

$$(T_\sigma v)(w,e) = \left\{ r(w, \sigma(w), e)^\alpha + \beta \left(\sum_{e'} v(\sigma(w), e')^\gamma \varphi(e') \right)^{\alpha/\gamma} \right\}^{1/\alpha}. \tag{3.27}$$

Since $(V, \mathbb{T})$ is a special case of the ADP discussed in §3.1.2.1, Proposition 3.1.13 implies that the fundamental optimality properties hold for $(V, \mathbb{T})$.

Exercise 3.2.4. Show that the subordinate ADP for the factored dynamic program $(V, \hat{V}, F, \mathbb{G})$ is $(\hat{V}, \hat{\mathbb{T}})$, where each $\hat{T}_\sigma \in \hat{\mathbb{T}}$ has the form

$$(\hat{T}_\sigma h)(w) = \left\{ \sum_e \{r(w, \sigma(w), e)^\alpha + \beta h(\sigma(w))^\alpha\}^{\gamma/\alpha} \varphi(e) \right\}^{1/\gamma}. \tag{3.28}$$

The benefit of working with $(\hat{V}, \hat{\mathbb{T}})$ is that $\hat{T}_\sigma$ acts on functions that depend only on w

rather than on both w and e (as is the case for T_σ). These lower dimensional operations are significantly more efficient, even when E is relatively small.

Since the fundamental optimality properties hold for $(V, \mathbb{T})$, Theorem 3.2.6 implies that they also hold for $(\hat{V}, \hat{\mathbb{T}})$. It also tells us that we can obtain an optimal policy for $(V, \mathbb{T})$ by finding the value function $\hat{v}_\top$ for $(\hat{V}, \hat{\mathbb{T}})$ and then calculating a policy σ obeying $G_\sigma \hat{v}_\top = G \hat{v}_\top$. By the definition of G_σ in (3.25), this means that we solve for σ satisfying (3.26) after setting $h = \hat{v}_\top$.

To compute $\hat{v}_\top$, we can use Theorem 1.2.12, which tells us that Howard policy iteration applied to $(\hat{V}, \hat{\mathbb{T}})$ converges to $\hat{v}_\top$ in finitely many steps. Summarizing this analysis, an optimal policy for $(V, \mathbb{T})$ can be computed via Algorithm 3.1.

Algorithm 3.1: Solving $(V, \mathbb{T})$ via $(\hat{V}, \hat{\mathbb{T}})$

input $\sigma_0 \in \Sigma$, set $k \leftarrow 0$ and $\varepsilon \leftarrow 1$

while $\varepsilon > 0$ **do**

  $h_k \leftarrow$ the fixed point of $\hat{T}_{\sigma_k}$

  $\sigma_{k+1} \leftarrow$ an h_k-greedy policy, satisfying

$$\sigma_{k+1}(w) \in \operatorname*{argmax}_{0 \leqslant s \leqslant w} \left\{ \sum_e \{r(w, s, e)^\alpha + \beta h(s)^\alpha\}^{\gamma/\alpha} \varphi(e) \right\}^{1/\gamma}$$

  $\varepsilon \leftarrow \mathbb{1}\{\sigma_k \neq \sigma_{k+1}\}$

  $k \leftarrow k + 1$

end

return the σ in (3.26) with $h = h_k$

Figure 3.1 shows $w \mapsto \sigma^*(w, e)$ for two values of e (smallest and largest) when σ^* is the optimal policy, calculated using Algorithm 3.1. In the computation we set $\Gamma(w, e) = [0, w]$ and $r(w, s, e) = w - s + e$. We chose α and γ to match the values used in Schorfheide et al. (2018).

In Figure 3.2 we display the relative speed gain from using the lower-dimensional model $(\hat{V}, \hat{\mathbb{T}})$ instead of $(V, \mathbb{T})$ across multiple choices of $|\mathsf{W}|$ and $|\mathsf{E}|$. The speed gain is the time required to solve an optimal policy for $(V, \mathbb{T})$ using HPI applied to $(V, \mathbb{T})$ (as in Theorem 1.2.12), divided by the time required to solve for the same optimal policy via Algorithm 3.1. The speed gain increases linearly in the size of E.

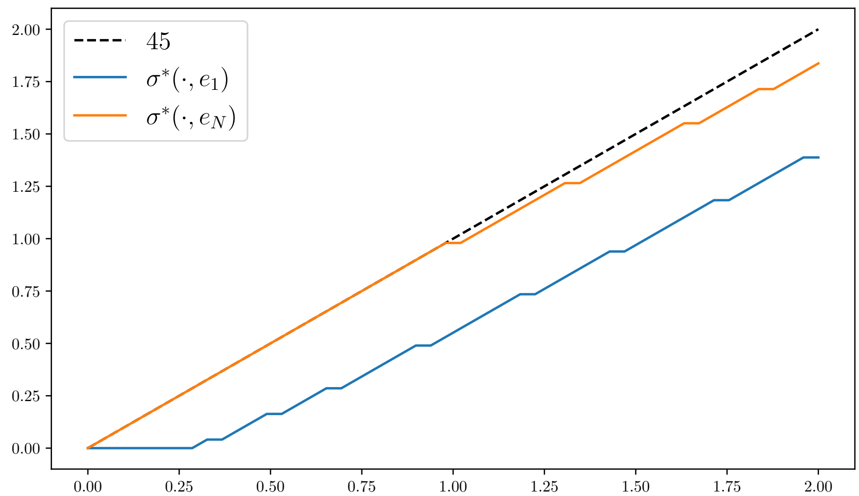


Figure 3.1: Optimal savings policy with Epstein–Zin preference

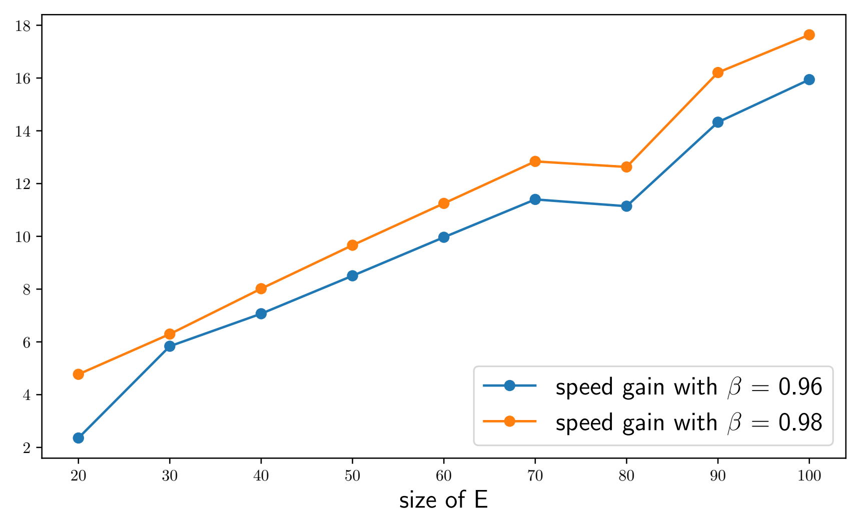


Figure 3.2: Speed gain from replacing $(V, \mathbb{T})$ with subordinate model $(\hat{V}, \hat{\mathbb{T}})$

Chapter 4

Discrete Choice

In this chapter we apply the theory from Chapter 1. We begin with optimal stopping problems and then examine some extensions. All models considered in this chapter have a finite action space. This is helpful because regularity (existence of greedy policies) is usually trivial to verify.

This chapter was written before the theory in Chapter 3 was established. The discussion of transformations needs to be rewritten to utilize that theory.

4.1 Optimal Stopping

In this section, we examine several optimal stopping problems, where a decision-maker receives a flow of rewards and chooses when to exit.

4.1.1 Job Search

We begin with the job search problem of McCall (1970), a finite state version of which was discussed at length in Chapter 1 of Sargent and Stachurski (2025a). Here we consider a general state version.

4.1.1.1 Description

In each period, an unemployed worker receives a wage offer W_t, drawn from some fixed distribution φ. The worker can accept or wait until the following period and consider a new offer. For now we require the wage sequence to be IID.

Assumption 4.1.1. The offer sequence $(W_t)_{t \geqslant 0}$ is IID and takes values in a nonempty Borel measurable set $\mathsf{W} \subset \mathbb{R}_+$. The distribution φ has finite mean, so that $\int w \varphi(\mathrm{d}w) < \infty$. The worker discounts future payoffs via constant discount factor $\beta \in (0, 1)$.

Let $L_1(\varphi) := L_1(\mathsf{W}, \mathscr{B}, \varphi)$ be all Borel measurable $f : \mathsf{W} \to \mathbb{R}$ with $\int |f| \, \mathrm{d}\varphi < \infty$. As usual, functions equal φ-almost everywhere are identified and $f \leqslant g$ means that $\{f > g\}$ has measure zero under φ. (See, e.g., Example A.5.13 on page 216.) Let Σ be all Borel measurable $\sigma : \mathsf{W} \to \{0, 1\}$. Each such σ can be understood as a policy, mapping states to actions: If $\sigma(w) = 1$, then the unemployed worker stops and accepts current offer w, while if $\sigma(w) = 0$, then she continues.

Consider first a two period problem. In period zero, the worker can either accept observed wage offer $w_0 \sim \varphi$ or continue to the next period, receiving unemployment compensation c and random payoff $v(W_1)$. The offer W_1 is drawn from φ and v is a given "terminal reward" function. Under policy σ, which maps the wage offer w_0 into an accept/reject decision, the expected present value of her payoff is

$$v_\sigma(w_0) := \sigma(w_0) w_0 + (1 - \sigma(w_0)) \left[c + \beta \int v(w') \varphi(\mathrm{d}w') \right]. \tag{4.1}$$

If $\sigma(w_0) = 1$ the worker accepts and and receives reward w_0. If $\sigma(w_0) = 0$, then she rejects and receives expected continuation reward $c + \beta \int v(w') \varphi(\mathrm{d}w')$.

Now we switch to an infinite horizon. Jobs are assumed to be permanent, so the present value of stopping with wage offer w in hand is

$$\frac{w}{1 - \beta} = w + \beta w + \beta^2 w \cdots \tag{4.2}$$

Fixing $\sigma \in \Sigma$, let $v_\sigma(w)$ be the lifetime value of following policy σ given initial wage offer w. By analogy with (4.1), we expect v_σ to obey the recursion

$$v_\sigma(w) = \sigma(w) \frac{w}{1 - \beta} + (1 - \sigma(w)) \left[c + \beta \int v_\sigma(w') \varphi(\mathrm{d}w') \right] \quad \text{for all } w \in \mathsf{W}. \tag{4.3}$$

Compared to (4.1), we have taken the value of stopping from (4.2) and also replaced the terminal value function v on the right-hand side of (4.1) with v_σ. This is because we now work with an infinite horizon, so that (4.3) becomes a recursion in v_σ.

Continuing to hold σ fixed, we introduce the policy operator $v \mapsto T_\sigma v$ via

$$(T_\sigma v)(w) = \sigma(w) \frac{w}{1 - \beta} + (1 - \sigma(w)) \left[c + \beta \int v(w') \varphi(\mathrm{d}w') \right]. \tag{4.4}$$

Since $L_1(\varphi)$ is closed under linear operations, policies are Borel measurable, and Assumption 4.1.1 is in force, we have $T_\sigma v \in L_1(\varphi)$ whenever $v \in L_1(\varphi)$. Clearly T_σ is order preserving on $(L_1(\varphi), \leqslant)$. Hence, with $\mathbb{T} := \{T_\sigma : \sigma \in \Sigma\}$, the pair $(L_1(\varphi), \mathbb{T})$ is an ADP. By construction, any fixed point of T_σ solves (4.3), so each such fixed point v_σ has the interpretation of assigning lifetime values to states under σ.

Exercise 4.1.1. Prove that $(L_1(\varphi), \mathbb{T})$ is well-posed.

Fix $v \in L_1(\varphi)$ and consider the policy σ given by

$$\sigma(w) = \mathbb{1}\left\{\frac{w}{1-\beta} \geqslant c + \beta \int v(w')\varphi(\mathrm{d}w')\right\} \qquad (w \in \mathsf{W}). \tag{4.5}$$

This policy tells the worker to stop when the payoff from stopping is larger than the expected payoff from continuing, assuming that v is used to value future states. We claim that σ is v-greedy. Indeed, for fixed $w \in \mathsf{W}$, combining (4.4) and (4.5) gives

$$(T_\sigma v)(w) = \max\left\{\frac{w}{1-\beta}, \, c + \beta \int v(w')\varphi(\mathrm{d}w')\right\}. \tag{4.6}$$

From (4.6) we see that $T_\tau v \leqslant T_\sigma v$ for any $\tau \in \Sigma$, so σ is v-greedy as claimed. Since the policy in (4.6) is well-defined at any $v \in V$, the ADP $(L_1(\varphi), \mathbb{T})$ is regular

Exercise 4.1.2. Starting from the usual ADP definition $Tv = \bigvee_\sigma T_\sigma v$, show that the Bellman operator of ADP $(L_1(\varphi), \mathbb{T})$ obeys

$$(Tv)(w) = \max\left\{\frac{w}{1-\beta}, \, c + \beta \int v(w')\varphi(\mathrm{d}w')\right\} \qquad (v \in V, \; w \in \mathsf{W}). \tag{4.7}$$

The expression for T in (4.7) aligns with the job search Bellman operator from Chapter 1 of Sargent and Stachurski (2025a), after replacing the finite expectation with an integral.

In the preceding analysis we used $L_1(\varphi)$ for the value space because φ is allowed to have unbounded support. Since w can be arbitrarily large, the function $T_\sigma v$ in (4.4) is unbounded. The set $L_1(\varphi)$ can handle unbounded functions. To complete this section, the next exercise looks at settings where φ has bounded support and considers how we might exploit this by selecting a smaller value space.

Exercise 4.1.3. Suppose that W is bounded and let

(i) $bm\mathsf{W}$ be the bounded Borel measurable functions on W,

(ii) $bc\mathsf{W}$ be the continuous functions in $bm\mathsf{W}$,

(iii) $ibc\mathsf{W}$ be the increasing (i.e., nondecreasing) functions in $bc\mathsf{W}$, and

(iv) $ibc\mathsf{W}_+$ be the nonnegative functions in $ibc\mathsf{W}$,

Show that if $V =$ any of these spaces paired with the pointwise order and each $T_\sigma \in \mathbb{T}$ is given by (4.4), then $(V, \mathbb{T})$ is an ADP with Bellman operator given by (4.7).

EXERCISE 4.1.4. Show that, in the setting of Exercise 4.1.3, the job search problem is an RDP. In your answer, set $V = bm\mathsf{W}$.

4.1.1.2 Optimality with IID Offers

Let's return now to using $L_1(\varphi)$ for the value space and consider optimality properties and convergence of algorithms for the job search ADP $(L_1(\varphi), \mathbb{T})$.

Proposition 4.1.1. *If Assumption 4.1.1 is in force, then*

(i) *the fundamental optimality properties hold, and*

(ii) *VFI, OPI and HPI all converge.*

Proof. Each T_σ is affine, as follows from (4.19). In particular, setting $r_\sigma := \sigma e + (1 - \sigma)c$ with $e := w/(1 - \beta)$, we can write

$$T_\sigma\, v = r_\sigma + K_\sigma\, v \quad \text{when} \quad K_\sigma\, v := (1 - \sigma)\beta\, \|v\| \mathbb{1}.$$

(Here $\|\cdot\|$ is the norm in $L_1(\varphi)$.) Observe that $0 \leqslant K_\sigma \leqslant K$ when $Kv := \beta\|v\|\mathbb{1}$. Since $\rho(K) = \beta < 1$ (see, e.g., Exercise A.4.17), the ADP $(L_1(\varphi), \mathbb{T})$ is regular. As $L_1(\varphi)$ is a Banach lattice (see §A.5.3.3), the conditions of Theorem 1.3.9 are satisfied and claims (i)–(iii) are true. □

Since the fundamental optimality properties hold, the value function $v_\top$ is a fixed point of the Bellman operator T and a policy σ is optimal if and only if it is $v_\top$-greedy, which is to say that

$$\sigma(w) = \mathbb{1}\left\{\frac{w}{1 - \beta} \geqslant c + \beta \int v_\top(w')\varphi(\mathrm{d}w')\right\}$$

for all $w \in \mathbb{R}_+$. (We are assuming that the agent accepts the job offer if indifferent.) In other words, the optimal rule is to stop if and only if

$$w \geqslant (1-\beta)\left[c + \beta \int v_\top(w')\varphi(\mathrm{d}w')\right].$$

The term on the right-hand side is called the **reservation wage**. This representation of optimal behavior is convenient because the reservation wage provides a scalar summary of the solution to the problem.

4.1.2 Rearranging the Bellman Equation

Sargent and Stachurski (2025a) show that, when the wage offer distribution has finite support, (a) it is possible to solve for the reservation wage directly, and (b) this leads to a lower-dimensional problem. The same idea works when the wage offer distribution is continuous. This section gives details.

4.1.2.1 Continuation Values

In view of (4.7), a function v satisfies the Bellman equation if

$$v(w) := \max\left\{\frac{w}{1-\beta},\ c + \beta \int v(w')\varphi(\mathrm{d}w')\right\} \quad \text{for all } w \in \mathsf{W}. \tag{4.9}$$

Taking v as given, consider the term

$$h := c + \beta \int v(w')\varphi(\mathrm{d}w'). \tag{4.10}$$

We can use h to eliminate the function v from (4.9). To do so we insert h on the right hand side, replace w with w' in (4.9), take expectations, multiply by β and add c to obtain

$$h = c + \beta \int \max\left\{\frac{w'}{1-\beta},\ h\right\}\varphi(\mathrm{d}w'). \tag{4.11}$$

This is a nonlinear equation in h, the solution of which, henceforth denoted $h_\top$, is the **optimal continuation value** of our problem. Obtaining $h_\top$ allows us to solve the dynamic programming problem, since the optimal policy can be written as

$$\sigma_\top(w) = \mathbb{1}\left\{\frac{w}{1-\beta} \geqslant h_\top\right\} \qquad (w \in \mathbb{R}_+). \tag{4.12}$$

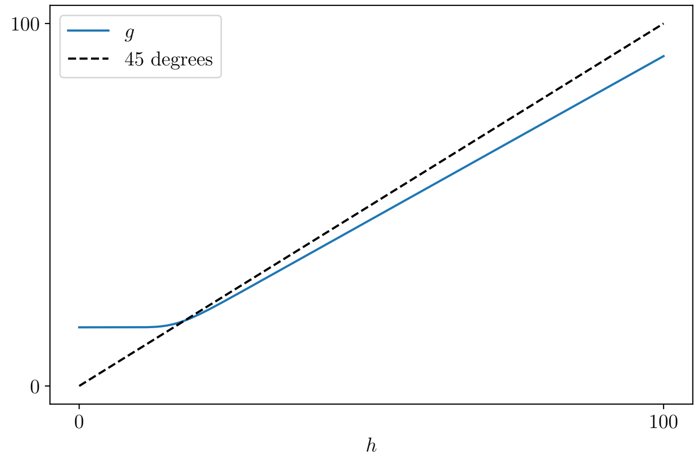


Figure 4.1: The function g from (4.14)

Another way to write the optimal policy is

$$\sigma_\top(w) = \mathbb{1}\{w \geqslant w_\top\} \quad \text{where} \;\; w_\top := (1-\beta)h_\top, \tag{4.13}$$

where the final term is the reservation wage.

In order to solve (4.11), we introduce the mapping

$$g(h) = c + \beta \int \max\left\{\frac{w'}{1-\beta}, h\right\} \varphi(\mathrm{d}w') \qquad (h \in \mathbb{R}_+). \tag{4.14}$$

It is constructed such that any solution to (4.11) is a fixed point of g and vice versa.

Exercise 4.1.5. Using the elementary bound

$$|\alpha \vee x - \alpha \vee y| \leqslant |x - y| \qquad (\alpha, x, y \in \mathbb{R}), \tag{4.15}$$

show that g is a contraction of modulus β on $\mathbb{R}_+$ under the usual Euclidean distance.

The result of Exercise 4.1.5 implies that g has a unique fixed point $h_\top$ in $\mathbb{R}_+$, which is the optimal continuation value. Figure 4.1 shows the function g when $\ln W_t = \mu + sZ_t$ for standard normal Z_t, while $\beta = 0.9$ and $c = 1.0$. The integral in (4.14) is computed by Monte Carlo. The unique fixed point is $h_\top$.

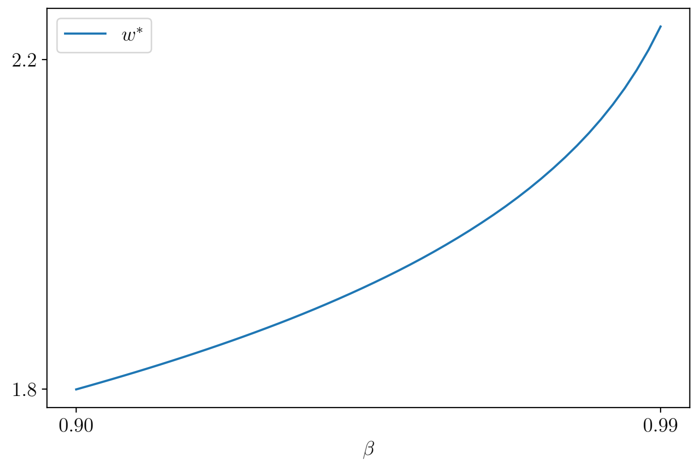


Figure 4.2: The continuation value computed directly

4.1.2.2 Reducing Dimension

One major advantage of the continuation value formulation of the problem in §4.1.2.1 is that, instead of searching for a value function $v_\top$ in the infinite-dimensional space $L_1(\varphi)$, we only need to solve for the fixed point of g in the one-dimensional space $\mathbb{R}_+$. The next exercise further reduces the search space to a bounded interval in $\mathbb{R}_+$.

Exercise 4.1.6. Find a constant K such that g maps $[0, K]$ to itself. (Hint: g is dominated by the affine function f defined by $f(h) = c + \beta \bar{w}/(1-\beta) + \beta h$, where $\bar{w} := \int w \varphi(\mathrm{d}w)$. What is the fixed point of f? Does g map this point down?)

Figure 4.2 shows the reservation wage, computed by iterating on g to obtain (an approximation to) $h_\top$ and then calculating $w_\top$ via (4.13). In the computation, c and the distribution φ are as for the last figure, while β ranges from 0.9 to 0.99.

Exercise 4.1.7. As a computational exercise, using the same specification as Figure 4.2 and $\beta = 0.98$, compute $h_\top$ as a fixed point of g in (4.14) and then $v_\top$ via $v_\top(w) = \max\{w/(1-\beta), h_\top\}$. Next, compute $v_\top$ as a fixed point of the Bellman operator (4.7). Plot both and confirm that the plots are essentially identical.

4.1.2.3 A Factored DP Perspective

Although this problem is simple, it is useful for us to see how the transformation discussed in §4.1.2.1 related to the theory on transformations of DP problems discussed in Chapter 3. For this discussion we adopt the environment of §4.1.1.2 and set

(i) $V = L_1(\varphi)$,

(ii) $\hat{V} = \mathbb{R}_+$,

(iii) $F \colon V \to \hat{V}$ with $Fv = c + \beta \int v(w')\varphi(\mathrm{d}w')$, and

(iv) $G_\sigma \colon \hat{V} \to V$ with $(G_\sigma h)(w) = \sigma(w)(w/(1-\beta)) + (1-\sigma(w))h$.

Clearly, given $h \in \hat{V}$, we can attain the bound $G_\tau h \leqslant G_\sigma h$ for all $\tau \in \Sigma$ by setting

$$\sigma(w) = \mathbb{1}\left\{\frac{w}{1-\beta} \geqslant h\right\} \qquad (w \in \mathsf{W}). \tag{4.16}$$

Hence, with $\mathbb{G} := \{G_\sigma\}_{\sigma \in \Sigma}$, the tuple $(V, \hat{V}, F, \mathbb{G})$ is a factored dynamic program.

For the primary ADP generated by $(V, \hat{V}, F, \mathbb{G})$, the policy operators have the form

$$(T_\sigma v)(w) = (G_\sigma F v)(w) = \sigma(w)\frac{w}{1-\beta} + (1-\sigma(w))\left[c + \beta \int v(w')\varphi(\mathrm{d}w')\right].$$

This is identical to (4.4), so the primary ADP is just the original job search ADP $(L_1(\varphi), \mathbb{T})$ from §4.1.1.1.

Regarding the secondary ADP generated by $(V, \hat{V}, F, \mathbb{G})$, the policy operators have the form

$$\hat{T}_\sigma h = F G_\sigma h = c + \beta \int \left[\sigma(w')\frac{w'}{1-\beta} + (1-\sigma(w'))h\right]\varphi(\mathrm{d}w').$$

The associated Bellman operator is

$$\hat{T} h = c + \beta \int \max\left\{\frac{w'}{1-\beta}, h\right\}\varphi(\mathrm{d}w').$$

On inspection, we see that the fixed point of $\hat{T}$ is a solution to (4.11). Thus, the subordinate ADP is represents the continuation value problem from §4.1.2.1.

These observations formalize the ideas expressed in §4.1.2.1, where we informally linked the original and continuation value representations of the job search problem. Now that we have a more rigorous link through the theory of factored dynamic programs, we can employ Theorem 3.2.6 to connect optimality results.

For example, that Theorem 3.2.6 tells us that a policy $\sigma \in \Sigma$ will be optimal for the job search problem when $h_\top$ is a fixed point of $\hat{T}$ and σ obeys $G_\sigma h_\top = G h_\top$. (Here G is the supremum $\bigvee_\sigma G_\sigma$.) In view of (4.16), such a σ can be found by setting

$$\sigma(w) = \mathbb{1}\left\{\frac{w}{1-\beta} \geqslant h_\top\right\} \qquad (w \in \mathsf{W}).$$

This policy aligns with, and justifies, the (informally derived) solution from (4.12).

4.1.2.4 Parametric Monotonicity

How does the solution to the job search problem vary with parameters? In terms of monotonicity, one way to answer this is to appeal to Proposition A.5.18 on page 215. Since g is an increasing contraction mapping on $\mathbb{R}_+$, this proposition implies that any parameter that shifts up the function g in (4.14) pointwise on $\mathbb{R}_+$ also shifts its fixed point up.

Example 4.1.1. The optimal continuation value $h_\top$ is increasing in c. Indeed, if $c_1 \leqslant c_2$, then

$$c_1 + \beta \int \max\left\{\frac{w'}{1-\beta}, h\right\} \varphi(\mathrm{d}w') \leqslant c_2 + \beta \int \max\left\{\frac{w'}{1-\beta}, h\right\} \varphi(\mathrm{d}w').$$

Thus, the function g shifts up everywhere when c increases and hence $h_\top$ increases with c. This is as expected, since higher unemployment compensation makes the value of continuing to the next period greater.

Exercise 4.1.8. Prove that the reservation wage $w_\top$ is increasing in unemployment compensation c. Prove also that $h_\top$ is increasing in β.

Figure 4.2 suggests that $w_\top$ is also increasing in β. Since $w_\top = (1-\beta)h_\top$, we cannot infer this directly from the fact that $h_\top$ is increasing in β. Instead, we take the fixed point equation for $h_\top$ in (4.11) and substitute $w = (1-\beta)h$, which uses the definition of the reservation wage from (4.13), to obtain a new fixed point equation $f(w) = w$ where

$$f(w) := c(1-\beta) + \beta \int \max\{w', w\}\, \varphi(\mathrm{d}w'). \tag{4.17}$$

Exercise 4.1.9. Let $\bar{w} = \int w' \varphi(\mathrm{d}w')$ be the mean wage offer. Prove the following: If $c \leqslant \bar{w}$, then the reservation wage $w_\top$ is increasing in β.

EXERCISE 4.1.10. Let τ be the first passage time to employment for an unemployed agent. That is,

$$\tau := \inf\{t \geqslant 0 : \sigma_\top(W_t) = 1\}$$

Prove that the mean first passage time $\mathbb{E}\tau$ increases with c.

How do shifts in the wage offer distribution affect the reservation wage? One observation is that a shift to a "more favorable" wage distribution should increase the reservation wage, since the agent can expect better offers. One natural way to order the set of wage distributions in terms of "favorability" is first order stochastic dominance $\preceq_F$ (see §A.5.5). We now show that the reservation wage increases as the offer distribution increases in this order.

First, let φ and ψ be two wage distributions on $\mathbb{R}_+$ with finite first moment and let $w^\varphi_\top$ and $w^\psi_\top$ be the associated reservation wages. To simplify matters we suppose that both distributions are supported on $[0, M]$. We then have the following monotonicity result:

Lemma 4.1.2. *If ψ first order stochastically dominates φ, then $w^\varphi_\top \leqslant w^\psi_\top$.*

Proof. The proof is another application of Proposition A.5.18. Let ψ and φ have the stated properties. Since the reservation wage is the fixed point of f in (4.17), it is enough to fix $w \in \mathbb{R}_+$ and show that

$$\int \max\{w', w\}\, \varphi(\mathrm{d}w') \leqslant \int \max\{w', w\}\, \psi(\mathrm{d}w')$$

Since $w' \mapsto \max\{w', w\}$ is bounded and increasing on $[0, M]$, this inequality follows directly from the definition of stochastic dominance. □

One more subtle monotonicity result for this model concerns the volatility of the wage process and its impact on the reservation wage and welfare. Intuitively, greater volatility encourages encourages patience because the option value of waiting is larger. In the next lemma, we formalize this idea using the notion of mean-preserving spreads, introduced on page 226.

Lemma 4.1.3. *If ψ is a mean-preserving spread of φ, then $w^\varphi_\top \leqslant w^\psi_\top$.*

Proof. Let ψ and φ have the stated properties and fix $w \in \mathbb{R}_+$. In view of Proposition A.5.18, it is enough to show that, under the stated assumptions, the value $f(w)$ in (4.17) increases with the mean-preserving spread, or, equivalently

$$\int \max\{w', w\}\, \varphi(\mathrm{d}w') \leqslant \int \max\{w', w\}\, \psi(\mathrm{d}w'). \tag{4.18}$$

To see that this is so, observe that, by the definition of a mean-preserving spread, there exists a pair (w', Z) such that $\mathbb{E}[Z \,|\, w'] = 0$, $w' \stackrel{d}{=} \varphi$ and $w' + Z \stackrel{d}{=} \psi$. By this fact and the law of iterated expectations,

$$\int \max\{w', w\}\, \psi(\mathrm{d}w') = \mathbb{E}\left[\max\{w' + Z, w\}\right] = \mathbb{E}\left[\mathbb{E}\left[\max\{w' + Z, w\} \,|\, w'\right]\right].$$

An application of Jensen's inequality now produces

$$\int \max\{w', w\}\, \psi(\mathrm{d}w') \geqslant \mathbb{E} \max\left\{\mathbb{E}[w' + Z \,|\, w'], w\right\}.$$

Using $\mathbb{E}[w' \,|\, w'] = w'$ and $\mathbb{E}[Z \,|\, w'] = 0$ confirms (4.18). □

4.1.3 Job Search with Correlated Wage Draws

In our simplistic model of job search we have so far assumed that wage offer draws are IID. Now let's allow these offers to have a Markov structure:

Assumption 4.1.2. The wage sequence (W_t) is P-Markov (see §A.5.4) on Borel set $\mathsf{W} \subset \mathbb{R}_+$, where P is a Markov operator on W. The Markov operator P has a stationary distribution φ on W with finite mean, so that $\int w\varphi(\mathrm{d}w) < \infty$.

4.1.3.1 An ADP Representation

As before, Σ is the set of all Borel measurable functions σ mapping W to $\{0, 1\}$. Each policy operator T_σ is adjusted to

$$(T_\sigma v)(w) = \sigma(w) \frac{w}{1-\beta} + (1 - \sigma(w)) \left[c + \beta \int v(w') P(w, \mathrm{d}w')\right].$$

We can write T_σ more succinctly as

$$T_\sigma v = \sigma e + (1 - \sigma)(c + \beta P v) \quad \text{when} \quad e(w) := \frac{w}{1-\beta}, \tag{4.19}$$

with products such as σe defined pointwise.

EXERCISE 4.1.11. Prove that T_σ is an order preserving self-map on $L_1(\varphi)$.

EXERCISE 4.1.12. Given $v \in L_1(\varphi)$, show that the policy $\sigma \in \Sigma$ given by

$$\sigma(w) := \mathbb{1}\left\{\frac{w}{1-\beta} \geqslant c + \beta \int v(w')P(w, \mathrm{d}w')\right\} \qquad (w \in \mathsf{W}). \tag{4.20}$$

is v-greedy (i.e., $T_\tau v \leqslant T_\sigma v$ for all $\tau \in \Sigma$). Show, in addition, that the ADP Bellman operator $T = \bigvee_\sigma T_\sigma$ corresponding to $(L_1(\varphi), \mathbb{T})$ obeys

$$(Tv)(w) = \max\left\{\frac{w}{1-\beta}, \, c + \beta \int v(w')P(w, \mathrm{d}w')\right\} \qquad (v \in L_1(\varphi), \; w \in \mathsf{W}). \tag{4.21}$$

It follows from Exercises 4.1.11–4.1.12 that, with $\mathbb{T}$ as all T_σ in (4.19) for some $\sigma \in \Sigma$, the pair $(L_1(\varphi), \mathbb{T})$ is a regular ADP.

EXERCISE 4.1.13. Show that every policy operator T_σ is order continuous on $L_1(\varphi)$.

EXERCISE 4.1.14. Show that the unique fixed point of T_σ in $L_1(\varphi)$ is

$$v_\sigma := (I - \beta(1-\sigma)P)^{-1}(\sigma e + (1-\sigma)c) \tag{4.22}$$

We can now state an optimality result the job search model with Markov wage draws.

Proposition 4.1.4. *If Assumption 4.1.2 holds, then the Markov job search ADP $(L_1(\varphi), \mathbb{T})$ is well-posed. Moreover,*

(i) *the fundamental optimality properties hold, and*

(ii) *VFI, OPI and HPI all converge.*

Proof. We showed above that $(L_1(\varphi), \mathbb{T})$ is regular. From (4.19) we can write $T_\sigma v = r_\sigma + K_\sigma v$ when $r_\sigma := \sigma e + (1-\sigma)c$ and $K_\sigma := (1-\sigma)\beta P$. Observe that $0 \leqslant K_\sigma \leqslant K$ when $K := \beta P$. In addition, $\rho(K) = \beta < 1$ by Lemma A.5.29. Moreover, $L_1(\varphi)$ is Dedekind complete (Lemma A.5.13). As a result, the conditions of Theorem 1.3.9 are satisfied. □

4.1.3.2 Reducing the Value Space

When solving ADPs it can be helpful to reduce the size of the value space, so that the space we need to search over when seeking $v_\top$ is smaller. This section suggests such a reduction for the job search problem. The details are left as a series of exericses.

EXERCISE 4.1.15. Let

$$\bar{v} := (I - \beta P)^{-1}(e + c) \quad \text{and} \quad V := \{v \in L_1(\varphi) \ : \ 0 \leqslant v \leqslant \bar{v}\}\,.$$

Show that, for all $\sigma \in \Sigma$, we have $v_\sigma \leqslant \bar{v}$ and $T_\sigma V \subset V$.

Since T_σ is also order preserving, $(V, \mathbb{T})$ is an ADP.

EXERCISE 4.1.16. Prove that the optimality results (i)–(iii) from §4.1.3.1 hold for $(V, \mathbb{T})$ using Theorem 1.2.14.

EXERCISE 4.1.17. Show that if W is finite, then HPI converges in finitely many steps.

4.1.3.3 A Numerical Study

Figure 4.3 shows the output of HPI under a range of parameter values. First we generate an $n \times n$ stochastic matrix P for wage offers via Tauchen's method, discretizing the AR1 process $W_{t+1} = \rho W_t + \nu \xi_{t+1}$ where (ξ_t) is IID and standard normal. We set $\beta = 0.99$, $\rho = 0.9$, $\nu = 0.2$ and $n = 500$. In the left subfigure we plot $v_\top$, computed by HPI, as well as the exit option $e(w) = w/(1-\beta)$ and the reservation wage, which is $\bar{w} := \min_{w \in \mathsf{W}} \sigma_\top(w) = 1$ when $\sigma_\top$ is $v_\top$-greedy. The reservation wage is the minimum wage off at which the unemployed worker accepts.

In the right subfigure we vary the volatility parameter ν over 0.1 0.2 and plot $\bar{w}$ as a function of ν, holding other parameters fixed. Notice that the reservation wage increases with wage offer volatility. The reason is that more volatility increases the upside of waiting, due to the possibility of high future offers. At the same time, downside risk is mitigated by the ability to reject a bad offer.

Figure 4.4 shows the first three iterates of HPI, OPI and VFI, as well as the value function $v_\top$ and the shared initial condition v. Parameter values are the same as the left hand subfigure in Figure 4.3. In the case of OPI, m is set to 10. We see that HPI converges faster than VFI in terms of reduced distance to the value function per iteration. The rate for OPI is between HPI and VFI.

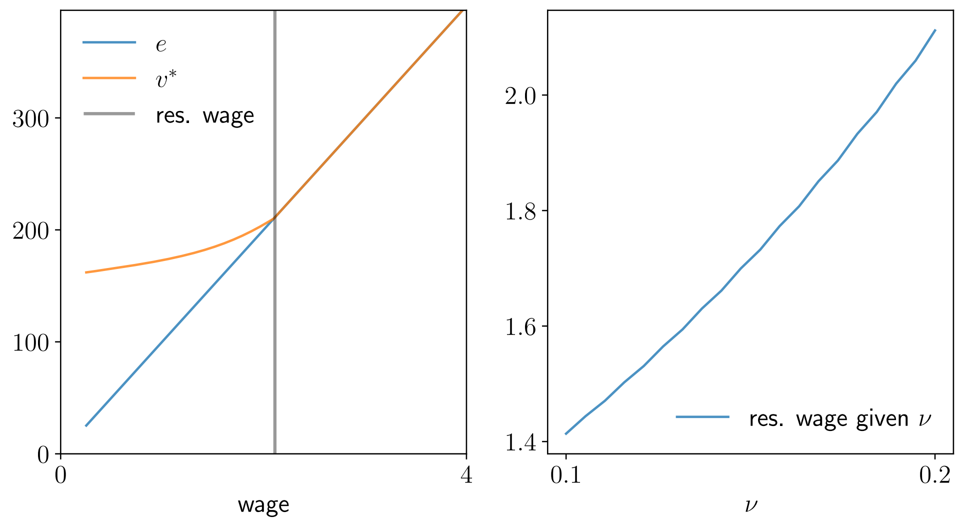


Figure 4.3: Solution to the job search problem

4.1.3.4 Persistent and Transient Components

Let's now look at a slightly more sophisticated wage offer process. In particular, we assume that (W_t) obeys

$$W_t = \exp(Z_t) + \exp(\mu + \sigma\zeta_t), \qquad \text{where} \quad Z_{t+1} = \rho Z_t + d + s\varepsilon_{t+1} \tag{4.23}$$

for some $\rho \in (-1, 1)$. The sequences $(\zeta)_{t\geqslant 1}$ and $(\varepsilon)_{t\geqslant 1}$ are both IID and standard normal. Thus, wages have a persistent component $\exp(Z_t)$ and a transient component, both of which are lognormal. The model is otherwise unchanged. The state becomes $(w, z) \in (0, \infty) \times \mathbb{R}$ and the Bellman equation is

$$v(w, z) = \max\left\{\frac{w}{1-\beta}, c + \beta\, \mathbb{E}_z v(w', z')\right\}. \tag{4.24}$$

Here $\mathbb{E}_z$ is expectation conditional on z. The expectation term can be written more explicitly as

$$\mathbb{E}_z v(w', z') = \int v\left[\exp(\rho z + d + s\varepsilon) + \exp(\mu + \sigma\zeta), \rho z + d + s\varepsilon\right]\, \varphi(\mathrm{d}\varepsilon, \mathrm{d}\zeta).$$

Here z and the parameters are fixed and φ is the $N(0, I)$ distribution on $\mathbb{R}^2$.

Rather than analyzing this model directly, let us first transform it using continuation values, analogous to the technique we used in §4.1.2. This will allow us to reduce

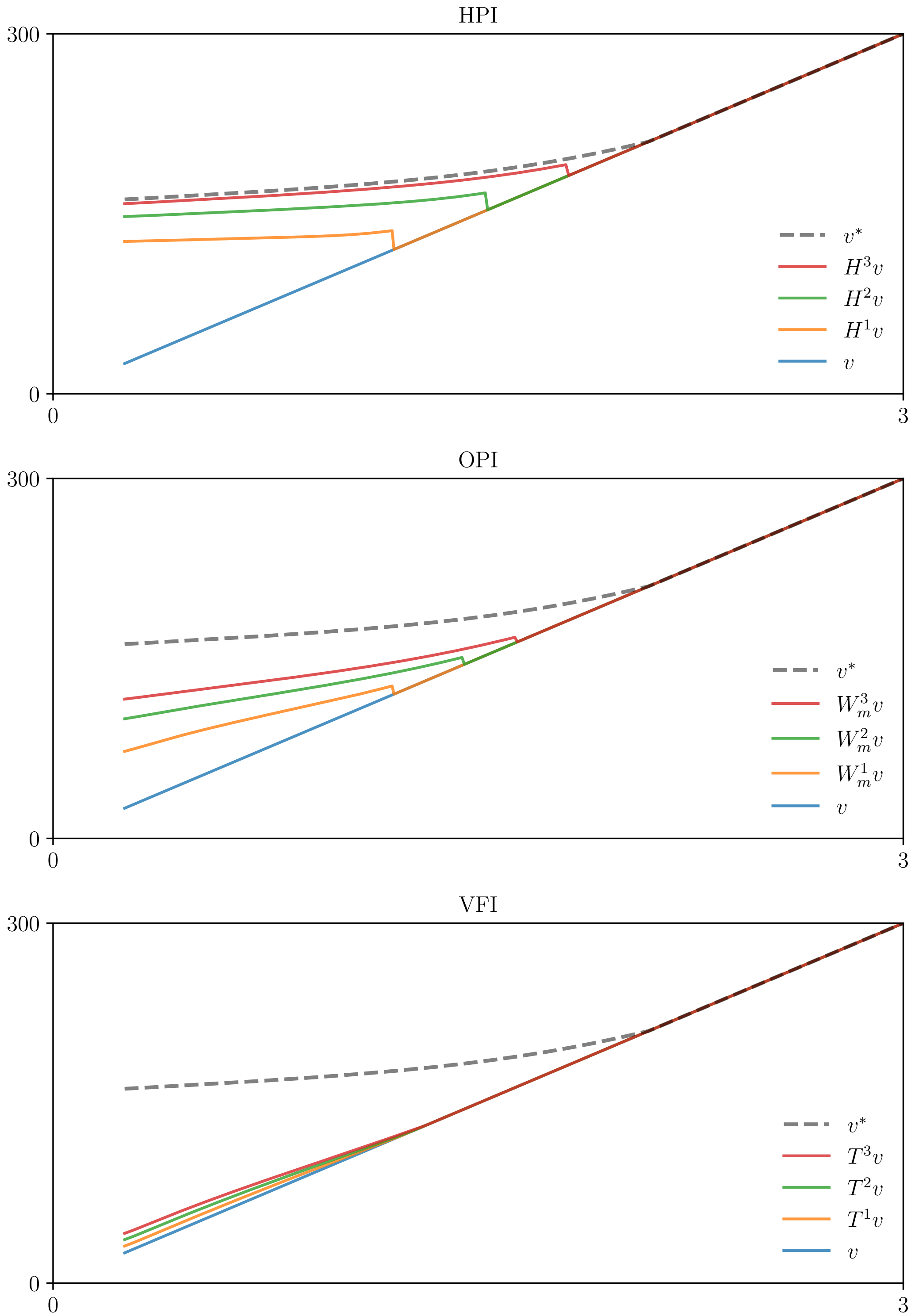


Figure 4.4: Comparison of algorithms (job search)

dimensionality, simplifing analysis and accelerating compution.

As a first step, let $h(z)$ be the continuation value from (4.24):

$$h(z) := c + \beta \, \mathbb{E}_z v(w', z') \qquad (z \in \mathbb{R}). \tag{4.25}$$

Notice that h is a function now, as opposed to the IID setting of (4.10), where the continuation value was just a constant. This is not surprising, since the current state can be used to predict future wages, which in turn determine future value.

Once we have h, the Bellman equation can be written as $v(w, z) = \max\{w/(1-\beta), \; h(z)\}$. Combining this with the definition of h, we see that

$$h(z) = c + \beta \, \mathbb{E}_z \max\left\{\frac{w'}{1-\beta}, h(z')\right\} \qquad (z \in \mathbb{R}). \tag{4.26}$$

(Note the similarity with (4.11).) The function h is defined on all of $\mathbb{R}$, since this is the domain of z. If we can obtain the solution $h_\top$ to this functional equation, we can use it to act optimally via the policy

$$\sigma_\top(w, z) = \mathbb{1}\left\{\frac{w}{1-\beta} \geqslant h_\top(z)\right\} \tag{4.27}$$

Put differently, we can stop when the current wage exceeds the reservation wage $w_\top(z) := h_\top(z)(1-\beta)$.

To make our analysis rigorous, let us construct an ADP such that the Bellman equation agrees with (4.26). To do so we take Σ to be all Borel measurable functions sending $(w', z') \in (0, \infty) \times \mathbb{R}$ to $\{0, 1\}$ and, for each $\sigma \in \Sigma$, we set

$$(\hat{T}_\sigma \, h)(z) := c + \beta \, \mathbb{E}_z \left\{\sigma(w', z') \frac{w'}{1-\beta} + (1 - \sigma(w', z'))h(z')\right\}. \tag{4.28}$$

We take φ to be the stationary distribution of (Z_t) and consider each $\hat{T}_\sigma$ as a mapping over all $h \in L_1(\varphi) := L_1(\mathbb{R}, \mathscr{B}, \varphi)$. Let $\hat{\mathbb{T}}$ be all $\hat{T}_\sigma$ given by (4.28) with $\sigma \in \Sigma$. The pair $(L_1(\varphi), \hat{\mathbb{T}})$ forms an ADP and the next exercise asks you to confirm this.

EXERCISE 4.1.18. Prove that each $\hat{T}_\sigma$ is an order preserving self-map on $L_1(\varphi)$.

EXERCISE 4.1.19. Prove: If $h \in L_1(\varphi)$, then the policy σ given by

$$\sigma(w', z') = \mathbb{1}\left\{\frac{w'}{1-\beta} \geqslant h(z')\right\} \qquad ((w', z') \in (0, \infty) \times \mathbb{R})$$

is h-greedy for $(L_1(\varphi), \hat{\mathbb{T}})$.

We can alternatively write $\hat{T}_\sigma h$ as $\hat{T}_\sigma h = m_\sigma + K_\sigma h$, where

$$m_\sigma(z) := c + \beta\, \mathbb{E}_z \left\{ \sigma(w', z') \frac{w'}{1-\beta} \right\} \quad \text{and} \quad (K_\sigma h)(z) := \mathbb{E}_z\, (1 - \sigma(w', z'))h(z').$$

Each K_σ is a positive linear operator on $L_1(\varphi)$ and, moreover, for the positive linear operator K defined by

$$(Kh)(z) := \beta \mathbb{E}_z\, h(z') = \beta \mathbb{E} h(\rho z + d + s\varepsilon_{t+1}),$$

we have $0 \leqslant K_\sigma h \leqslant Kh$. By Lemma A.5.29, the spectral radius of K equals β, which is strictly less than one. Hence, by Theorem 1.3.9, $(L_1(\varphi), \hat{\mathbb{T}})$ is well-posed, the fundamental optimality properties hold, and VFI, OPI and HPI all converge.

Let's now characterize the Bellman operator, which is defined on $L_1(\varphi)$ by $\hat{T}h = \bigvee_\sigma \hat{T}_\sigma h$. Recall that a policy σ is h-greedy if and only if $\hat{T}_\sigma h = \hat{T}h$ (Lemma 1.2.1). Using this fact and Exercise 4.1.19, you will find it straightforward to solve the next exercise.

EXERCISE 4.1.20. Given $h \in L_1(\varphi)$, prove that $\hat{T}h$ obeys

$$(\hat{T}h)(z) = c + \beta\, \mathbb{E}_z \max \left\{ \frac{w'}{1-\beta}, h(z') \right\} \qquad (z \in \mathbb{R}). \tag{4.29}$$

EXERCISE 4.1.21. Prove that $\hat{T}$ is a contraction of modulus β on $L_1(\varphi)$.

Since $L_1(\varphi)$ is complete, Banach's contraction mapping theorem implies that $\hat{T}$ has a unique fixed point $h_\top$ in $L_1(\varphi)$.

EXERCISE 4.1.22. Let c_a and c_b be two levels of unemployment compensation satisfying $c_a \leqslant c_b$. Let $\hat{T}^a$ and $\hat{T}^b$ be the corresponding continuation value operators, so that

$$(\hat{T}^i h)(z) = c + \beta\, \mathbb{E}_z \max \left\{ \frac{w'}{1-\beta}, h(z') \right\} \qquad (i \in \{a, b\},\ z \in \mathbb{R}).$$

Let h_a and h_b be their respective fixed points. Show that $h_a \leqslant h_b$ pointwise on $\mathbb{R}$.

EXERCISE 4.1.23. Suppose the agent seeks to maximize lifetime value $\mathbb{E} \sum_{t=0}^{\infty} \beta^t u(y_t)$, where y_t is earnings at time t and u is a utility function. Letting $u(c) = \ln c$, write down the modified Bellman equation and the $\hat{T}$ operator (4.29). How does the reservation wage change?

4.1.4 Firm Exit

We consider a firm exit problem originally studied by Jovanovic (1982) and Hopenhayn and Prescott (1992) and extended by many authors (see, e.g., Alessandria et al. (2021) or Sterk et al. (2021)). We follow Jovanovic (1982) but admit the following extensions to the original decision problem: (i) firm profits depend on an aggregate shock, as well as a firm-specific shock and a cross-sectional distribution, (ii) the interest rate is allowed to vary over time, and (iii) the outside option of the firm is permitted to depend on aggregates and a cross-sectional distribution.

4.1.4.1 Set Up

Let $\pi(s, \mu, z)$ be current profit for the firm, where s is an firm-specific state variable that takes values in set S, μ is a cross-sectional distribution taking values in set D, and z is an aggregate shock taking values in set Z. At the start of each period, the firm receives current profit and then decides whether to (a) exit and receive outside option $q(\mu', z')$ at the start of next period, or (b) continue. Here and below, primes denote next period values.

The decision of whether or not to continue depends on the cost of capital, which we refer to henceforth as the interest rate and denote by $r(\mu, z)$. This rate is allowed to depend on the cross-section μ and the aggregate shock. The firm is risk-neutral and discounts future payoffs with discount factor $\beta(\mu, z) := 1/(1 + r(\mu, z))$.

Let $\mathsf{X} = \mathsf{S} \times \mathsf{D} \times \mathsf{Z}$ and let $\mathscr{B}$ be a σ-algebra over X that makes π, q, β and the transition probabilies measurable. We write the dynamics as $x' = P(x, \cdot)$, meaning that P is a stochastic kernel on $(\mathsf{X}, \mathscr{B})$ and the next period state $x' = (s', \mu', z')$ given current state $x = (s, \mu, z)$ is drawn from distribution $P(x, \cdot)$. A policy is a Borel measurable map σ from X to $\{0, 1\}$ with $\sigma(x) = 1$ indicating the decision to exit at state x and $\sigma(x) = 0$ indicating the decision to continue. Let Σ be the set of all policies.

Assumption 4.1.3. The Markov operator P has unique stationary distribution φ on $(\mathsf{X}, \mathscr{B})$. The functions π, q and β are nonnegative, measurable, and φ-integrable.

Assumption 4.1.3 tells us that π, q and β are elements of the positive cone in $L_1(\varphi) := L_1(\mathsf{X}, \mathscr{B}, \varphi)$. Note that we them to be both discontinuous and unbounded above. We endow $L_1(\varphi)$ with the φ-a.e. pointwise order $\leqslant$, so that $f \leqslant g$ means $\varphi\{f > g\} = 0$.

4.1.4.2 ADP Representation

Analogous to (4.3), at current state x, a fixed policy σ yields lifetime firm value $v_\sigma(x)$, where v_σ satisfies the recursion

$$v_\sigma(x) = \pi(x) + \beta(x) \int [\sigma(x')q(x') + (1 - \sigma(x'))v_\sigma(x')]P(x, \mathrm{d}x') \quad (x \in \mathsf{X}).$$

Equivalently, v_σ is a fixed point of T_σ defined at $v \in L_1(\varphi)$ by

$$T_\sigma v = \pi + K(\sigma q + (1 - \sigma)v) \tag{4.30}$$

where the operator

$$(Kv)(x) := \beta(x) \int v(x')P(x, \mathrm{d}x') \qquad (v \in L_1(\varphi), \ x \in \mathsf{X}).$$

discounts future cash flows.

Assumption 4.1.4. K maps $L_1(\varphi)$ to itself and the spectral radius obeys $\rho(K) < 1$.

Under Assumptions 4.1.3 and 4.1.4, each T_σ is a self-map on $L_1(\varphi)$. Since K is a positive operator, each T_σ is order preserving. Hence, letting $\mathbb{T}$ be all policy operators of the form (4.30), with σ ranging over the policy set Σ, the pair $(L_1(\varphi), \mathbb{T})$ is an ADP.

Exercise 4.1.24. Show that $(L_1(\varphi), \mathbb{T})$ is regular. In particular, show that, given $v \in L_1(\varphi)$, the policy $\sigma = \mathbb{1}\{q \geqslant v\}$ is v-greedy. (Hint: See (4.5) and the subsequent proof that the job search ADP is regular.)

The ADP Bellman equation is $v = Tv := \bigvee_\sigma T_\sigma v$, where, in the present setting, the supremum is taken in $L_1(\varphi)$ under the φ-almost everywhere pointwise order $\leqslant$.

Exercise 4.1.25. Show that, given $v \in L_1(\varphi)$, we have

$$Tv = \pi + K(q \vee v). \tag{4.31}$$

It follows from (4.31) that the Bellman equation for the ADP is $v = \pi + K(q \vee v)$. Rewriting this expression using the definition of K, we get

$$v(x) = \pi(x) + \beta(x) \int \max\{q(x'), v(x')\} P(x, \mathrm{d}x') \tag{4.32}$$

This agrees with the original Bellman equation from Jovanovic (1982), modulo the extensions we have added.

4.1.4.3 Optimality

We can now state the following result for the firm exit ADP $(L_1(\varphi), \mathbb{T})$.

Proposition 4.1.5. *If Assumptions 4.1.3–4.1.4 hold, then so do the fundamental optimality properties, and VFI, HPI, and OPI all converge.*

Proof. Since each T_σ is affine, we can prove Proposition 4.1.5 by checking the conditions of Theorem 1.3.9 on page 34. The ADP $(L_1(\varphi), \mathbb{T})$ is regular (by Exercise 4.1.24 and affine, with $r_\sigma := \pi + K\sigma q$ and $K_\sigma := K(1-\sigma)$. Since r_σ lies in $L_1(\varphi)$, since $0 \leqslant K_\sigma \leqslant K$ for all $\sigma \in \Sigma$, and since, by assumption, $\rho(K) < 1$, the conditions of Theorem 1.3.9 all hold. □

Proposition 4.1.5 implies that the value function $v_\top$ solves the Bellman equation (4.32), and that $v_\top$ can be computed, at least approximately, by VFI, HPI or OPI. Moreover, policies are optimal if and only if they are $v_\top$-greedy, which, by Exercise 4.1.24, translates to $\sigma = \mathbb{1}\{q \geqslant v_\top\}$.

The next exercise is useful if we wish to implement HPI.

Exercise 4.1.26. Show that, when Assumption 4.1.4 holds, the unique fixed point of T_σ in $L_1(\varphi)$ is

$$v_\sigma := (I - K(1-\sigma))^{-1}(\pi + K\sigma q). \tag{4.33}$$

Exercise 4.1.27. Prove Proposition 4.1.5 using Theorem 1.3.8 instead.

4.2 Extensions

Roadmap.

4.2.1 Nonlinear Discounting

In this section we consider again the job search model from §4.1.3.1, leaving Σ unchanged while modifying each policy operator T_σ to

$$(T_\sigma v)(w) = \sigma(w)\frac{w}{1-\beta(w)} + (1-\sigma(w))\left[c + \int \beta[v(w')]P(w, \mathrm{d}w')\right].$$

The only difference with the original policy operators in §4.1.3.1 is that we have included nonlinear discounting, so that β is now a function of future payoffs, rather than a constant.

This form of nonlinear discounting was introduced and studied in Jaśkiewicz et al. (2014) and Bäuerle et al. (2021). One interesting motivation is **magnitude effects**, under with discount rates decrease with the size of the reward (i.e., large rewards are discounted less, so the discount factor is higher). Evidence in favor of such effects has been found by Green et al. (1997) and other researchers.

In studying this nonlinear discounting model, we continue to assume that the wage sequence (W_t) is P-Markov (see §A.5.4) on Borel set $\mathsf{W} \subset \mathbb{R}_+$. To make the analysis more straightforward, we will assume that $\mathsf{W} = [w_1, w_2]$ where $0 < c < w_1 < w_2$. For the discount factor function we set

$$\beta(w) := bF(w, \lambda) \quad \text{where } b \in (0,1) \text{ and } F(w, \lambda) = 1 - \exp(-\lambda w).$$

For the value space we take

$$V := [0, \bar{v}] \quad \text{where} \quad \bar{v} := \frac{c + w_2}{1 - b}.$$

In this expression, V is understood as a subset of $bm\mathsf{W}$, the Banach lattice of bounded measurable real-valued functions on W. As usual, V is endowed with the pointwise partial order $\leqslant$ and the supremum norm.

EXERCISE 4.2.1. Prove that each T_σ maps V into itself.

Since every T_σ is order-preserving, we see that $(V, \mathbb{T})$ is an ADP when $\mathbb{T} := \{T_\sigma : \sigma \in \Sigma\}$. Extending our earlier analysis, a policy σ is v-greedy whenever

$$\sigma(w) := \mathbb{1}\left\{\frac{w}{1 - \beta(w)} \geqslant c + \int \beta[v(w')]P(w, \mathrm{d}w')\right\} \qquad (w \in \mathsf{W}).$$

Since such a policy exists, the ADP $(V, \mathbb{T})$ is regular.

EXERCISE 4.2.2. Show that $(V, \mathbb{T})$ is order continuous.

Obtaining optimality results for this model is not entirely trivial because T_σ is not, in general, contracting. This failure of contractivity is due to the fact that β can be steep close to zero, as shown in Figure 4.5 when $b = 0.99$.

At the same time, β is concave, which gives us some hope that we can use the concavity-based optimality result in Theorem 1.3.10. To apply this theorem it suffices to show

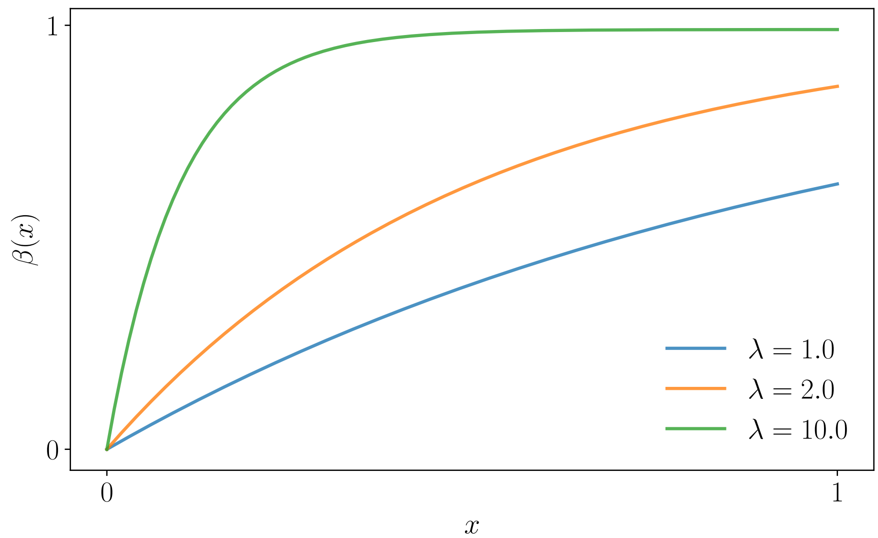


Figure 4.5: The discount function β for different choices of λ

that $(V, \mathbb{T})$ is regular, order continuous, and, for given σ,

- the operator T_σ is concave on V and $T_\sigma\, 0 \geqslant \varepsilon \bar{v}$ for some $\varepsilon \in (0, 1)$.

Concavity of T_σ follows easily from concavity of β and monotonicty of the integral. Also,

$$(T_\sigma\, v)(w) \geqslant \min \left\{ \frac{w}{1 - \beta(w)},\ c + \int \beta[v(w')] P(w, \mathrm{d}w') \right\} \geqslant w_1 \wedge c = c > 0.$$

Hence we can take an $\varepsilon \in (0, 1)$ with $T_\sigma\, 0 \geqslant \varepsilon \bar{v}$. Thus, all of the conditions of Theorem 1.3.10 are satisfied.

4.2.2 Nonlinear Expectations

In the last section we modified the job search model to include nonlinear discounting. Here we drop nonlinear discounting but assume instead that the job seeker uses power-transformed expectations, so that

$$T_\sigma\, v = \sigma e + (1 - \sigma)(c + \beta R v)$$

(compare with (4.19)), where

$$(Rv)(w) := \left(\int v^{\gamma}(w') P(w, \mathrm{d}w') \right)^{1/\gamma} \qquad (w \in \mathsf{W}, \ \gamma \neq 0).$$

The operator R is called a **Kreps–Porteus** certainty equivalent operator. The value γ parameterizes risk attitudes for the unemployed worker with respect to intertemporal gambles and can be positive or negative. As before, $e(w) := w/(1 - \beta)$ is the stopping reward.

The operator T_σ and the operator R act on the set

$$V := [c, \bar{v}] \quad \text{where} \quad \bar{v} := \frac{c + w_2}{1 - \beta}.$$

As in §4.2.1, V regarded as an order interval in the Banach lattice $bm\mathsf{W}$ and $0 < c < w_1 < w_2$. The constant β lies in $(0, 1)$.

We note the following facts regarding R:

(i) R is order preserving on V and maps any constant function to itself.

(ii) R is concave when $\gamma \leqslant 1$ and convex when $\gamma \geqslant 1$.

Fact (i) is easy to check and fact (ii) is taken from Lemma 7.3.1 of Sargent and Stachurski (2025a). (While that discussion treats the finite-state case, the same arguments hold here.)

Letting $\mathbb{T} := \{T_\sigma : \sigma \in \Sigma\}$, we aim to show that $(V, \mathbb{T})$ is an ADP and, moreover, that the conditions of Theorem 1.3.10 hold. As a first step, we observe that, for fixed $\sigma \in \Sigma$ and for sufficiently small positive ε,

$$(T_\sigma\, c)(w) \geqslant \min \left\{ \frac{c}{1 - \beta}, \ c + \beta c \right\} \geqslant c + \varepsilon(\bar{v} - c).$$

Also, for sufficiently small positive ε,

$$(T_\sigma\, \bar{v})(w) \leqslant \max \left\{ \frac{w_2}{1 - \beta}, \ c + \beta \bar{v} \right\} \leqslant \max \left\{ \bar{v} - \frac{c}{1 - \beta}, \ \bar{v} - w_2 \right\} \leqslant \bar{v} + \varepsilon(\bar{v} - c).$$

Since R and hence T_σ is order preserving, these facts tell us that T_σ maps V to itself, so $(V, \mathbb{T})$ is an ADP.

Exercise 4.2.3. Show that $(V, \mathbb{T})$ is order continuous.

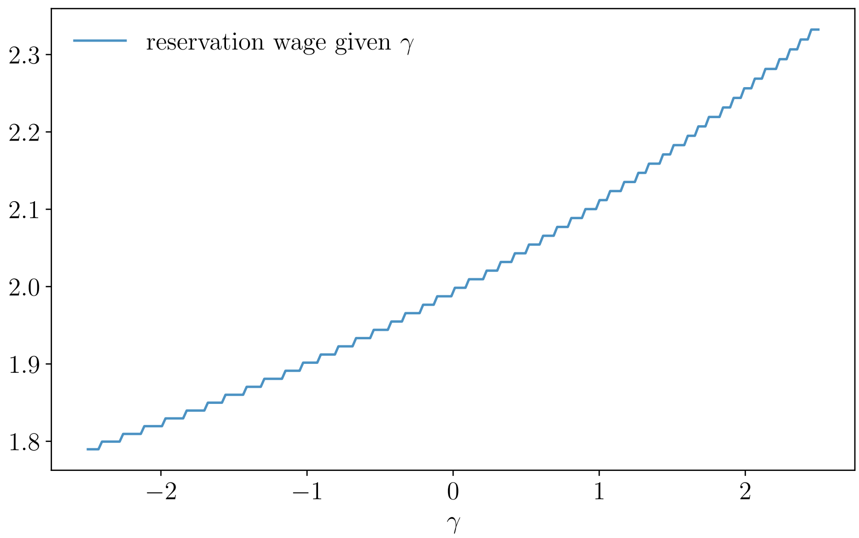


Figure 4.6: The reservation wage as a function of γ

Combining this order continuity and the last two ε bounds with fact (ii) above, we see that the conditions of Theorem 1.3.10 are satisfied for every $\gamma \in \mathbb{R} \setminus \{0\}$. As a result, the fundamental optimality properties hold and VFI, OPI, and HPI all converge.

Figure 4.6 shows the reservation wage $\bar{w}$ as a function of γ. The figure was computed as follows. Fixing γ, we calculated $v_\top$ via VFI, set

$$\sigma(w) := \mathbb{1}\left\{\frac{w}{1-\beta} \geqslant c + \beta\left(\int [v_\top(w')]^\gamma P(w, \mathrm{d}w')\right)^{1/\gamma}\right\}$$

and then set $\bar{w} := \min_{w \in \mathsf{W}} \sigma(w) = 1$. Aside from γ, the parameters used in the calculations were the same as those given in §4.1.3.3. The figure shows that the reservation wage increases with γ, as the job seeker becomes progressively less risk-averse. Decreasing risk aversion means that gambles over future payoffs become more attractive, which favors continuing over stopping. This encourages the job seeker to increase the reservation wage.

4.2.3 Job Search with Learning

Next we consider a variation of the job search model from §6.6 of Ljungqvist and Sargent (2018). The framework is the IID setting of §4.1.1.1, apart from the fact that the wage offer distribution φ is unknown to the worker. Instead, the agent learns about φ by starting with a prior belief and then successively updating her beliefs based on observed wage offers.

4.2.3.1 The Model

The structure of information is as follows: The worker knows there are two possible offer distributions, with densities f and g. At the start of time, nature selects φ to be either f or g, the wage distribution from which the entire sequence $(W_t)_{t \geqslant 0}$ will be drawn. This choice is not observed by the worker, who puts prior probability π_0 on f being chosen. In other words, the worker's initial guess of φ is $\pi_0 f(w) + (1 - \pi_0) g(w)$. Beliefs subsequently update according to Bayes' rule. Thus, the agent, having observed W_{t+1}, updates π_t to π_{t+1} via

$$\pi_{t+1} = \frac{f(W_{t+1})\pi_t}{f(W_{t+1})\pi_t + g(W_{t+1})(1 - \pi_t)}. \tag{4.34}$$

This expression is more easily understood as an application of Bayes rule if we write it as

$$\mathbb{P}\{\varphi = f \mid W_{t+1}\} = \frac{\mathbb{P}\{W_{t+1} \mid \varphi = f\}\mathbb{P}\{\varphi = f\}}{\mathbb{P}\{W_{t+1}\}}$$

and use the law of total probability to obtain the denominator:

$$\mathbb{P}\{W_{t+1}\} = \sum_{\psi \in \{f, g\}} \mathbb{P}\{W_{t+1} \mid \varphi = \psi\}\mathbb{P}\{\varphi = \psi\}.$$

We assume that the wage offer is always bounded above by a finite constant M.

Assumption 4.2.1. The densities f and g are positive on $(0, M)$ and zero elsewhere.

Using (4.34), we can formulate and ADP representation of the optimal stopping problem. Dropping time subcripts, let $\varphi_\pi := \pi f + (1 - \pi) g$ represent the estimate of the wage offer distribution given belief π and let

$$\kappa(w, \pi) := \frac{\pi f(w)}{\pi f(w) + (1 - \pi) g(w)} \qquad (w \in (0, M) \; \pi \in (0, 1)).$$

In particular, $\kappa(w, \pi)$ is the updated value π' of π having observed draw w. The state is $(w, \pi) \in (0, M) \times (0, 1)$ and π is referred to as the **belief state**.

By analogy with the original IID job search policy operators (see (4.4)), the policy operators for this learning search problem take the form

$$(T_\sigma \, v)(w, \pi) = \sigma(w, \pi) \frac{w}{1-\beta} + (1 - \sigma(w, \pi)) \left[c + \beta \int v_\top(w', \kappa(w', \pi)) \, \varphi_\pi(w') \, \mathrm{d}w' \right].$$

Each T_σ acts on $v \in V$, which we define as the set of bounded Borel measurable functions on $(0, M) \times (0, 1)$. Let $\mathbb{T} = \{T_\sigma : \sigma \in \Sigma\}$. Evidently $(V, \mathbb{T})$ is an ADP.

Exercise 4.2.4. Given $v \in V$, show that σ defined by

$$\sigma(w) = \mathbb{1} \left\{ \frac{w}{1-\beta} \geqslant c + \beta \int v_\top(w', \kappa(w', \pi)) \, \varphi_\pi(w') \, \mathrm{d}w' \right\}$$

is v-greedy for $(V, \mathbb{T})$.

Exercise 4.2.5. Show that the Bellman equation for this ADP is given by

$$v(w, \pi) = \max \left\{ \frac{w}{1-\beta}, \, c + \beta \int v(w', \kappa(w', \pi)) \, \varphi_\pi(w') \, \mathrm{d}w' \right\}. \tag{4.35}$$

4.2.3.2 An Efficient Solution Method

Rather than tackling $(V, \mathbb{T})$ directly, we will introduce a variation with a lower dimensional state space. This lower dimensional problem is significantly easier to work with. To begin, fix $v \in V$ and let $\omega(\pi)$ be the corresponding reservation wage at belief state π, which is the wage level at which the worker is indifferent between accepting and rejecting. This value satisfies

$$\frac{\omega(\pi)}{1-\beta} = c + \beta \int v(w', \kappa(w', \pi)) \, \varphi_\pi(w') \, \mathrm{d}w'. \tag{4.36}$$

We combine (4.35) and (4.36) to obtain

$$v(w, \pi) = \max \left\{ \frac{w}{1-\beta}, \, \frac{\omega(\pi)}{1-\beta} \right\}$$

and then use this expression to eliminate v in (4.36), obtaining

$$\omega(\pi) = (1-\beta)c + \beta \int \max \left\{ w', \omega[\kappa(w', \pi)] \right\} \, \varphi_\pi(w') \, \mathrm{d}w'. \tag{4.37}$$

Equation (4.37) can be understood as a functional equation in ω. Equivalently, the map ω is the fixed point of the operator $\hat{T}$ given by

$$(\hat{T}\omega)(\pi) = (1-\beta)c + \beta \int \max\{w', \omega[\kappa(w', \pi)]\} \, \varphi_\pi(w') \, \mathrm{d}w'. \tag{4.38}$$

When this fixed point is well-defined we call it the **optimal reservation wage function**. The value $\omega(\pi)$ will indicate the smallest wage offer at which the worker is willing to accept, given her current belief state π.

EXERCISE 4.2.6. Prove that $\hat{T}$ is a contraction of modulus β on $\hat{V} := bm(0,1)$.

EXERCISE 4.2.7. Prove that $\hat{T}$ maps $bc(0,1)$ to itself.

4.2.3.3 Parametric Monotonicity

Let's try computing the optimal reservation wage function using the ideas described above. The wage offer distributions are set to

$$f = \text{Beta}(4,2) \quad \text{and} \quad g = \text{Beta}(2,4), \tag{4.40}$$

as shown in Figure 4.7. The other parameters are $c = 0.1$ and $\beta = 0.95$. Since $\hat{T}$ is a contraction of modulus β on $\hat{V}$, a unique solution $\omega_\top$ to the reservation wage functional equation exists in $\hat{V}$ and $\hat{T}^k\omega \to \omega_\top$ uniformly as $k \to \infty$, for any $\omega \in \hat{V}$. Figure 4.8 shows the result of this iteration, the optimal reservation wage, as a function of π, the belief state.

Note that the optimal reservation wage function $\omega_\top$ in Figure 4.8 is increasing in π. This result seems reasonable: In Figure 4.7, the density f puts more mass on higher draws, so, as our belief shifts toward f, our reservation wage should increase. The next proposition gives conditions for such monotoncity.

Proposition 4.2.1. *If f and g have the monotone likelihood ratio property, then $\omega_\top$ is increasing in π.*

Proof. Let $ibm(0,1)$ be all increasing functions in $bm(0,1)$. AS $ibm(0,1)$ is closed in $bm(0,1)$ (see, e.g., Exercise A.1.17), it suffices to show that $\hat{T}$ is invariant on $ibm(0,1)$. So pick any $\omega \in ibm(0,1)$. Since $\hat{T}$ maps $bm(0,1)$ to itself, we need only show that $\hat{T}\omega$ is increasing. For this it suffices to show that, with

$$h(w', \pi) := \omega\left[\frac{\pi f(w')}{\pi f(w') + (1-\pi)g(w')}\right]$$

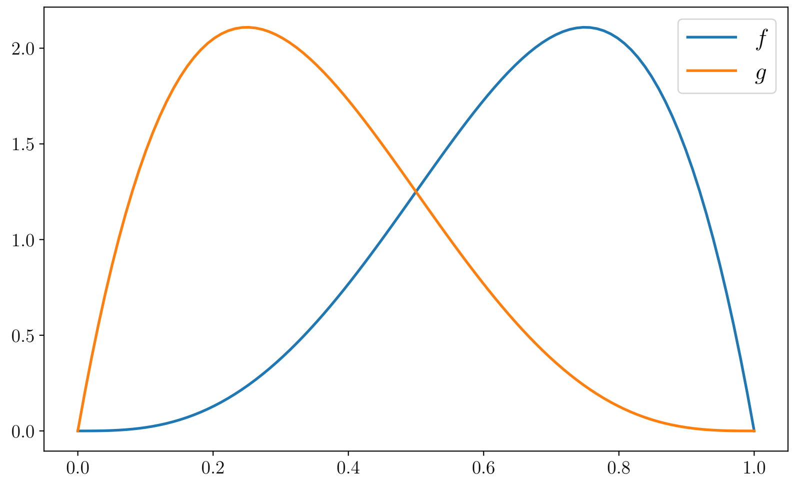


Figure 4.7: The two unknown densities f and g

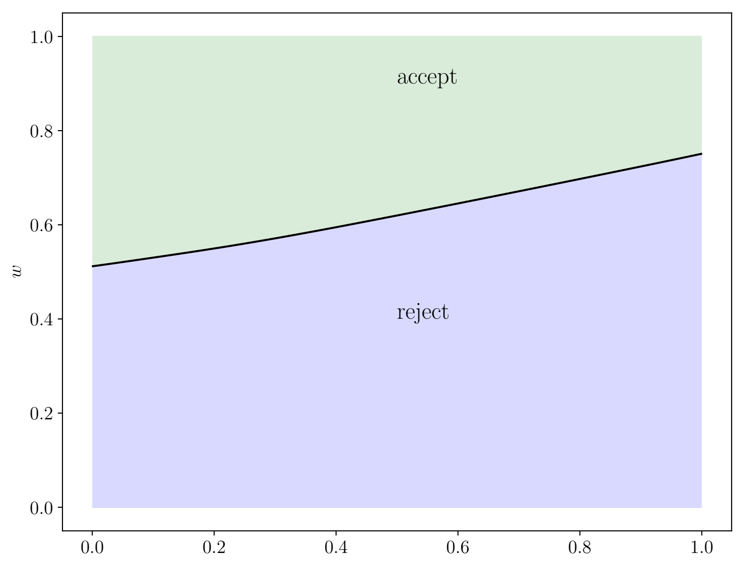


Figure 4.8: Optimal reservation wage function $\omega_{\top}$

the function

$$\pi \mapsto \int \max\{w', h(w', \pi)\} \; \varphi_\pi(w') \; \mathrm{d}w'$$

is increasing. This will be true if we can establish that (i) h is increasing in both π and w', and (ii) the map $\pi \mapsto \varphi_\pi$ is isotone with respect to $\preceq_{\mathrm{F}}$. To see that (i) holds, write h as

$$h(w', \pi) = \omega \left[\frac{1}{1 + [(1-\pi)/\pi][g(w')/f(w')]} \right]$$

Since ω is increasing, this expression is increasing in π. Also, f and g are assumed to have the monotone likelihood ratio property, which means that $g(w')/f(w')$ is decreasing in w', and hence $h(w', \pi)$ is increasing in w'. Thus, condition (i) is established.

Condition (ii) follows from Proposition A.5.31 on page 226, along with the result of Exercise 4.2.8. □

Exercise 4.2.8. Let F and G be two distributions on $\mathbb{R}$ with $G \preceq_{\mathrm{F}} F$. Let H_α be the convex combination defined by

$$H_\alpha := \alpha F + (1-\alpha) G \qquad (0 \leqslant \alpha \leqslant 1)$$

Show that $\alpha \leqslant \beta$ implies $H_\alpha \preceq_{\mathrm{F}} H_\beta$.

Exercise 4.2.9. Show that f and g in (4.40) have the monotone likelihood ratio property. Hint: the Gamma function is increasing over the interval $[2, 4]$.

4.2.4 Renewal Problems

Roadmap.

4.2.4.1 Job Search with Separation

We consider a version of the job search model from §4.1.3.1 where separation can occur. In particular, an existing match between worker and firm dissolves with probability α every period. Note that this discussion extends a treatment of a similar model in a finite-state setting from Chapter 3 of Sargent and Stachurski (2025a).

To simplify the discussion, we assume that the set of possible wage offers $\mathsf{W} \subset \mathbb{R}_+$ is bounded above by some constant M. The state space for the problem is $\mathsf{X} := \{e, u\} \times \mathsf{W}$, with a typical element (s, w) denoting employment status s (here e means employed and u means unemployed), and current offer w. A policy is a Borel measurable map $\sigma \colon \mathsf{W} \to \{0, 1\}$, where, as usual, $\sigma(w) = 0$ means "reject the current offer" and $\sigma(w) = 1$ means "accept."

The wage offer sequence (W_t) is assumed to be P-Markov on W. The value space V will be all bounded and Borel measurable $v \colon \mathsf{X} \to \mathbb{R}$ and we endow V with the supremum norm and the pointwise partial order.

The policy operators take the form

$$(T_\sigma v)(e, w) = w + \beta \left[\alpha \int v(u, w') P(w, \mathrm{d}w') + (1 - \alpha) v(e, w) \right] \tag{4.41}$$

and

$$(T_\sigma v)(u, w) = \sigma(w) v(e, w) + (1 - \sigma(w)) \left[c + \beta \int v(u, w') P(w, \mathrm{d}w') \right]. \tag{4.42}$$

The right-hand side of the first expression is the current value of being employed with offer w in hand, given the continuation values embodied in v. The right-hand side of the second expression is the current value of being employed with offer w in hand, conditional on using policy σ.

We can solve this problem directly by setting up the corresponding ADP, we can also start by simplifying the value space in a way we now describe. This will make the analysis easier and help with computation. The first step is to regard (4.41) as a fixed point problem, replacing $(T_\sigma v)(e, w)$ with $v(e, w)$ on the left hand side and treating $v(u, \cdot)$ as given. Simple algebra then gives

$$v(e, w) = \frac{1}{1 - \beta(1 - \alpha)} \left[w + \alpha\beta \int v(u, w) P(w, \mathrm{d}w') \right]. \tag{4.43}$$

Let's write this in operator notation. In doing so, we will rewrite $v(u, \cdot)$ as v_u and $v(e, \cdot)$ as v_e. Setting

$$h(w) := \frac{1}{1 - \beta(1 - \alpha)} w, \quad \text{and} \quad \gamma := \frac{\alpha\beta}{1 - \beta(1 - \alpha)},$$

we have $v_e = h + \gamma P v_u$. We substitute this expression into (4.42) to get

$$T_\sigma v_u = \sigma(h + \gamma P v_u) + (1 - \sigma)(c + \beta P v_u), \tag{4.44}$$

We take $bm\mathsf{W}$ as the value space and let $\mathbb{T} = \{T_\sigma : \sigma \in \Sigma\}$. As before, Σ is all Borel measurable maps from W to $\{0, 1\}$. Recalling that W is bounded above, one can easily confirm that T_σ maps $bm\mathsf{W}$ to itself. The maps J and K are positive linear operators, so T_σ is order preserving. Hence $(bm\mathsf{W}, \mathbb{T})$ is an ADP.

Given $v_u \in bm\mathsf{W}$, set $v_e = h + \gamma P v_u$ and consider the policy σ defined by

$$\sigma(w) := \mathbb{1}\left\{v_e(w) \geqslant c + \beta \int v_u(w')P(w, \mathrm{d}w')\right\} \quad \text{for all } w \in \mathsf{W}. \tag{4.45}$$

We claim that σ is v_u-greedy. Indeed, for such a σ and any alternative policy s we have

$$T_s\, v_u = s(h + \gamma P v_u) + (1 - s)(c + \beta P v_u) \leqslant (h + \gamma P v_u) \vee (c + \beta P v_u) = T_\sigma\, v_u.$$

The expression for σ in (4.45) is natural because it tells the worker to accept employment whenever its value is higher than the expected present value of continuing, given the continuation value for unemployment associated with v_u.

EXERCISE 4.2.10. Prove the following: There exists a $\lambda \in (0,1)$ such that T_σ is a contraction of modulus λ on $bm\mathsf{W}$ for all $\sigma \in \Sigma$.

Regarding optimality, we have the following result.

Proposition 4.2.2. *The ADP $(bm\mathsf{W}, \mathbb{T})$ is well-posed. Moreover,*

(i) *the fundamental optimality properties hold, and*

(ii) *VFI, OPI and HPI all converge.*

Proof. We showed above that every $v_u \in bm\mathsf{W}$ has at least one greedy policy, so the ADP is regular. The claims in Proposition 4.2.2 now follow from Exercise 4.2.10 and Theorem 1.3.5. □

The value function $v^u_\top$ for an unemployed worker satisfies the recursion

$$v^u_\top(w) = \max\left\{v^e_\top(w),\; c + \beta \sum_{w' \in \mathsf{W}} v^u_\top(w')P(w, w')\right\} \qquad (w \in \mathsf{W}), \tag{4.46}$$

where $v^e_\top$ is the value function for an employed worker, that is, the lifetime value of a worker who starts the period employed at wage w. Value function $v^e_\top$ satisfies

$$v^e_\top(w) = w + \beta\left[\alpha \sum_{w'} v^u_\top(w')P(w, w') + (1 - \alpha)v^e_\top(w)\right] \qquad (w \in \mathsf{W}). \tag{4.47}$$

This equation states that value accruing to an employed worker is current wage plus the discounted expected value of being either employed or unemployed next period.

We claim that, when $0 < \alpha, \beta < 1$, the system (4.46)–(4.47) has a unique solution $(v^e_\top, v^u_\top)$ in $V \times V$.

Substituting into (4.46) yields

$$v^u_\top(w) = \max\left\{\frac{1}{1-\beta(1-\alpha)}\left(w + \alpha\beta(Pv^u_\top)(w)\right),\ c + \beta\,(Pv^u_\top)(w)\right\}. \tag{4.48}$$

EXERCISE 4.2.11. Prove that there exists a unique $v^u_\top \in V$ that solves (4.48). Propose a convergent method for computing both $v^u_\top$ and $v^e_\top$.

The stopping and continuation values are given by

$$s_\top(w) := \frac{1}{1-\beta(1-\alpha)}\left(w + \alpha\beta(Pv^u_\top)(w)\right) \quad \text{and} \quad h_\top(w) := c + \beta\,(Pv^u_\top)(w)$$

respectively, for each $w \in \mathsf{W}$. The value function $v^u_\top$ is the pointwise maximum (i.e., $v^u_\top = s_\top \vee h_\top$). The worker's optimal policy while unemployed is

$$\sigma_\top(w) := \mathbb{1}\{s_\top(w) \geqslant h_\top(w)\}.$$

4.2.4.2 Optimal Harvests

Consider again the optimal harvest problem from §2.2.1.6, which we structured as an RDP with $V = bm\mathsf{X}$, $\Gamma(x) = \{0, 1\}$ for all x, and

$$B(s, p, a, v) = a(ps - m(s)) - (1-a)c + \beta \int v(q(s)(1-a), p')\varphi(\mathrm{d}p').$$

The pair (s, p) is the state and the state space for the model is $\mathsf{X} := [0, \bar{s}] \times [0, \bar{p}]$. The function q is a self-map on $[0, \bar{s}]$. We continue to assume that m and c are continuous real-valued functions on $[0, \bar{s}]$, and that the firm discounts the future at rate $\beta < 1$. If $a = 1$, then the firm chooses to harvest, whereas $a = 0$ indicates the decision to continue without harvesting.

We suppose first that $q(0) = 0$. In other words, the plantation never regenerates, so harvesting terminates the decision process. The dynamic program is now an optimal stopping problem. The Bellman equation can be written as

$$v(s, p) = \max\left\{ps - m,\ -c + \beta \int v(q(s), p')\varphi(\mathrm{d}p')\right\}.$$

We write the continuation value as

$$h(s) := -c + \beta \int v(q(s), p') \varphi(\mathrm{d}p')$$

This is lower dimensional than v so computing it directly will be advantageous. To this end, we use h to eliminate v from the Bellman equation, first by writing

$$v(s, p) = \max \{ps - m, \; h(s)\}$$

and then further manipulating to obtain the functional equation

$$h(s) = -c + \beta \int \max \{p's - m, \; h(s)\} \, \varphi(\mathrm{d}p') \tag{4.49}$$

Next we introduce an operator Q such that fixed points of Q coincide with solutions to (4.49). It takes the form

$$Qh(s) = -c + \beta \int \max \{p's - m, \; h(s)\} \, \varphi(\mathrm{d}p') \tag{4.50}$$

Let $\mathsf{S} = [0, \bar{s}]$. We claim that Q is a contraction map on $bc\mathsf{S}$ under d_∞. To show this, we use Jensen's inequality and the bound (4.15) on page 119 to obtain, for arbitrary $s \in \mathsf{S}$ and $g, h \in bc\mathsf{S}$,

$$\begin{aligned} |Qg(s) - Qh(s)| &\leqslant \beta \int |\max \{p's - m, \; h(s)\} - \max \{p's - m, \; g(s)\}| \, \varphi(\mathrm{d}p') \\ &\leqslant \beta \int |g(s) - h(s)| \, \varphi(\mathrm{d}p'). \end{aligned}$$

Hence

$$|Qg(s) - Qh(s)| \leqslant \beta \, \|g(s) - h(s)\|_\infty.$$

Taking the supremum over s completes the proof.

Exercise 4.2.12. Show that, in the current setting, the fixed point of Q decreases everywhere when maintenance cost c increases. Use this fact to show that, should we observe two firms, both starting from the same biomass and facing the same price sequence but with different maintenance cost, the firm with higher maintenance costs harvest no later than (i.e., before or at the same time as) the firm with low maintenance costs.

Next we assume that $q(0) > 0$, so the plantation regenerates after each harvest.

EXERCISE 4.2.13. Write down the Bellman operator for this problem. Show that this operator is a contraction of modulus β on $bc\mathsf{X}$ under the supremum distance.

4.3 Chapter Notes

Discuss the connection with Sargent and Stachurski (2025a).

This discounting condition in Assumption 4.1.4 is similar to (but weaker than) discounting restrictions found in Hansen and Scheinkman (2012) and Borovička and Stachurski (2020).

Appendices

Appendix A

Mathematical Background

This chapter provides a relatively extensive description of the mathematical tools we employ in the book. Readers with comprehensive mathematical backgrounds can use it as a reference. Readers from fields outside of mathematics might find it useful to review the whole chapter. Either way, all readers should, at minimum, be familiar with the material in §A.1 before starting Chapter 1.

A.1 Foundations

In this first section of the chapter, we cover several foundational ideas in analysis, including metrics and partial orders.

A.1.1 Properties of the Real Line

Let's start with the real line, which has a natural metric and a natural order. Both the metric and the order have useful completeness properties. Later, we investigate conditions under which such properties extend to more general spaces.

A.1.1.1 Min, Max, Sup and Inf

A point $u \in \mathbb{R}$ is called an **upper bound** of a set $A \subset \mathbb{R}$ if $a \leqslant u$ for all $a \in A$. Let $U(A)$ be the set of upper bounds of A. If $s \in U(A)$ and $s \leqslant u$ for all $u \in U(A)$, then s is called the **supremum** of A and we write $s = \sup A$. At most one such supremum s exists. (Why?) If s is in $U(A)$ then the following are equivalent:

(i) $s = \sup A$

(ii) for all $\varepsilon > 0$, there exists a point $a \in A$ with $a > s - \varepsilon$

Exercise A.1.1. Prove the last claim. Prove also that $\sup(0, 1] = \sup(0, 1) = 1$.

Theorem A.1.1. *If $A \subset \mathbb{R}$ is nonempty and bounded above, then $\sup A$ exists in $\mathbb{R}$.*

Theorem A.1.1 is essentially axiomatic, and is equivalent to "completeness" of $\mathbb{R}$ – which we discuss in §A.1.1.2.

For $A \subset \mathbb{R}$, a **lower bound** of A is any number ℓ such that $\ell \leqslant a$ for all $a \in A$. If $i \in \mathbb{R}$ is an lower bound for A and also satisfies $i \geqslant \ell$ for every lower bound ℓ of A, then i is called the **infimum** of A and we write $i = \inf A$. At most one such i exists, and every nonempty subset of $\mathbb{R}$ bounded from below has an infimum.

We adopt the following conventions:

- If A is not bounded above, then $\sup A := +\infty$.
- If A is not bounded below, then $\inf A := -\infty$.
- If $A = \emptyset$, the $\sup A = -\infty$ and $\inf A = +\infty$.

A number m contained in a subset A of $\mathbb{R}$ is called the **maximum** of A and we write $m = \max A$ if $a \leqslant m$ for every $a \in A$. It is called the **minimum** of A if $a \geqslant m$ for every $a \in A$.

Exercise A.1.2. Show that, if A is a closed and bounded subset of $\mathbb{R}$, then A has both a maximum and a minimum.

Exercise A.1.3. Prove the following statements:

(i) If $A \subset B$, then $\sup A \leqslant \sup B$.

(ii) If $s = \sup A$ and $s \in A$, then $s = \max A$.

(iii) If $i = \inf A$ and $i \in A$, then $i = \min A$.

Given an arbitrary set D and $f \in \mathbb{R}^D$, we set

$$\sup_{x \in D} f(x) := \sup\{f(x) : x \in D\} \quad \text{and} \quad \max_{x \in D} f(x) := \max\{f(x) : x \in D\}.$$

A point $x^* \in D$ is called a

- **maximizer** of f on D if $x^* \in D$ and $f(x^*) \geqslant f(x)$ for all $x \in D$, and a
- **minimizer** of f on D if $x^* \in D$ and $f(x^*) \leqslant f(x)$ for all $x \in D$.

Equivalently, $x^* \in D$ is a maximizer of f on D if $f(x^*) = \max_{x \in D} f(x)$, and a minimizer if $f(x^*) = \min_{x \in D} f(x)$. We define

$$\operatorname*{argmax}_{x \in D} f(x) := \{x^* \in X : f(x^*) \geqslant f(x) \text{ for all } x \in D\}.$$

The set $\operatorname{argmin}_{x \in D} f(x)$ is defined analogously.

Exercise A.1.4. Regarding the order structure of $\mathbb{R}$, the following relationships are sometimes helpful: Given $x, y \in \mathbb{R}$ and $a \in \mathbb{R}_+$,

(i) $x + y = x \vee y + x \wedge y$

(ii) $|x - y| = x \vee y - x \wedge y$

(iii) $|x - y| = x + y - 2(x \wedge y)$

(iv) $|x - y| = 2(x \vee y) - x - y$

(v) $a(x \vee y) = (ax) \vee (ay)$

(vi) $a(x \wedge y) = (ax) \wedge (ay)$

Prove (i) and (ii).

We will make use of the following lemma:

Lemma A.1.2. *If f, g map nonempty set X to $\mathbb{R}$ and both are bounded above, then*

$$|\sup f - \sup g| \leqslant \sup |f - g|.$$

Proof. We have $f = f - g + g \leqslant |f - g| + g$ and hence $\sup f \leqslant \sup |f - g| + \sup f$. Rearranging gives one side of the inequality in Lemma A.1.2. The other side is obtained by reversing the roles of f and g. □

A.1.1.2 Completeness of the Real Line

Recall that a sequence $(x_n) \subset \mathbb{R}$ is called **Cauchy** if, for all $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ with $|x_n - x_m| < \varepsilon$ whenever $n, m \geqslant N$.

Exercise A.1.5. Prove: if $u_n = 1/n$, then (u_n) is Cauchy.

EXERCISE A.1.6. Prove that every convergent sequence in $\mathbb{R}$ is Cauchy.

The converse is also true:

Theorem A.1.3. *A real sequence converges in $\mathbb{R}$ if and only if it is Cauchy.*

The statement that every Cauchy sequence in $\mathbb{R}$ converges should be understood as an axiomatic property of $\mathbb{R}$. It states that, once the irrational numbers are mixed in with the rational numbers, there are "no more gaps" in the real line. This is called the **completeness** property of $\mathbb{R}$ and follows from its definition. A clear discussion can be found in Bartle and Sherbert (2011).

A.1.2 Partial Orders

Partially ordered spaces are the natural habitat of dynamic programs, since we aim to find policies that optimal at all states (rather than just maximizing a real-valued objective). In this section we introduce the key ideas needed for the book.

A.1.2.1 Partially Ordered Sets

The pair $(V, \preceq)$ is called a **partially ordered set** – or **poset** – if V is any nonempty set and $\preceq$ is a relation $\preceq$ on $V \times V$ such that, for any u, v, w in V,

$u \preceq u$ (reflexivity)

$u \preceq v$ and $v \preceq u$ implies $u = v$ and (antisymmetry)

$u \preceq v$ and $v \preceq w$ implies $u \preceq w$ (transitivity)

The relation $\preceq$ is called a **partial order** on V. We often write V instead of $(V, \preceq)$ when $\preceq$ is understood. We sometimes say that w **dominates** v when $v \preceq w$.

Example A.1.1. The usual order $\leqslant$ is a partial order on $\mathbb{R}$.

Example A.1.2. If $\mathcal{S}$ is any collection of sets, then $\subset$ is a partial order on $\mathcal{S}$. For example, if $E, F \in \mathcal{S}$, then $E \subset F$ and $F \subset E$ implies $E = F$.

Example A.1.3. For $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$, we write $u \leqslant v$ if $u_i \leqslant v_i$ for $i = 1, \ldots, n$. It is simple to confirm that $\leqslant$ is a partial order on $\mathbb{R}^n$.

A subset C of a poset $(V, \preceq)$ is called a **chain** in V if either $u \preceq v$ or $v \preceq u$ for all $u, v \in C$. A poset $(V, \preceq)$ is called **totally ordered** if V itself is a chain.

Example A.1.4. $(\mathbb{R}, \leqslant)$ is totally ordered, while $\{(q,q) : q \in \mathbb{Z}\}$ is a chain in $\mathbb{R}^2$.

In applications, one of the most important notions of partial order is the pointwise partial order. To define it we let U, V be nonempty sets with V partially ordered by $\preceq$. Let V^U be a set of maps from U to V. For each $f, g \in V^U$, we set

$$f \preceq g \quad \iff \quad f(u) \preceq g(u) \text{ for all } u \in U. \tag{A.1}$$

Then $\preceq$ is a partial order on V^U, usually called the **pointwise order** on V^U.

EXERCISE A.1.7. Confirm that $\preceq$ is a partial order on V^U.

One very common special case is where V is a subset of $\mathbb{R}^{\mathsf{X}}$ for some nonempty set X. In this setting, *we always write the pointwise partial order as* $\leqslant$. In particular, for arbitrary $u, v \in \mathbb{R}^{\mathsf{X}}$ we write $u \leqslant v$ if and only if $u(x) \leqslant v(x)$ for all $x \in \mathsf{X}$. The partial order in Example A.1.3 is a special case, when $\mathsf{X} = \{1, \ldots, n\}$.

A.1.2.2 Bounds

Let V be a poset. $I \subset V$ is called an **order interval** in V if there exists an a, b in V with $a \preceq b$ and

$$I = [a, b] := \{v \in V : a \preceq v \preceq b\}.$$

Example A.1.5. If $(bc\mathsf{X}, \leqslant)$ is all bounded continuous real-valued functions on $\mathsf{X} \subset \mathbb{R}$ and $I = [g, h]$ for some $g, h \in bc\mathsf{X}$, then I is the order interval in $bc\mathsf{X}$ containing all $f \in bc\mathsf{X}$ such that $g(x) \leqslant f(x) \leqslant h(x)$ for all x in X.

Given a poset V and a subset A of V, we call

- $u \in V$ an **upper bound** of A if $a \preceq u$ for all a in A and
- $\ell \in V$ a **lower bound** of A if $a \succeq \ell$ for all a in A.

A subset A of poset V is called **bounded above** (resp., **bounded below**) if the set of upper bounds (resp., lower bounds) of A is nonempty (i.e., there exists at least one $v \in V$ with $a \preceq v$ for all $a \in A$). A is called **order bounded** in V if A is both bounded above and bounded below. Obviously, A is order bounded in V if and only if there exists an order interval $I \subset V$ such that $A \subset I$.

Example A.1.6. For the poset $(\mathbb{R}^n, \leqslant)$, a set $A \subset \mathbb{R}^n$ is order bounded if and only if it is bounded; that is, if and only if there exists an $M \in \mathbb{N}$ with $\|a\| \leqslant M$ for all $a \in A$.

EXERCISE A.1.8. Prove the claim in Example A.1.6.

A.1.2.3 Greatest and Least Elements

Given poset V and $A \subset V$, we say that

- $g \in V$ is the **greatest element** of A if $g \in A$ and $a \in A \implies a \preceq g$; and
- $\ell \in V$ is the **least element** of A if $\ell \in A$ and $a \in A \implies \ell \preceq a$.

In other words, a greatest element of A is an upper bound of A that is also contained in A, while a least element of A is a lower bound of A also contained in A.

Example A.1.7. Continuing Example A.1.2, if $\wp(A)$ is the set of all subsets of set A, then $\subset$ is a partial order on $\wp(A)$. Since $B \subset A$ for all $B \in \wp(A)$, we see that A is the greatest element of $\wp(A)$. The least element is $\varnothing$.

EXERCISE A.1.9. Prove: A subset A of a poset V can have at most one greatest element and at most one least element.

Not all subsets of partially ordered sets have greatest elements. For one example, observe that $\mathbb{N} \subset (\mathbb{R}, \leqslant)$ has no greatest element. In this case the set of upper bounds is empty, so finding a greatest element is impossible. We can also have situations where the set of upper bounds is nonempty but no greatest element exists.

Example A.1.8. In Figure A.1, a point $y \in \mathbb{R}^2$ obeys $y \leqslant x$ only when y lies to the southwest of x (below and to the left of the dashed lines). Thus, x is not a greatest element of the circle C shown in the figure. Some thought will convince you that no other point in C is a greatest element of C.

If a poset V is itself has a greatest element, that element is sometimes called the **top** of V. A least element of V is sometimes called the **bottom** of V.

Example A.1.9. If C is all continuous $f : \mathbb{R} \to [0, 1]$, then C is order bounded with top $\equiv 1$ and bottom $\equiv 0$.

A.1.2.4 Suprema and Infima

Let A be a subset of poset V and let $U(A)$ be the set of all upper bounds of A in V. We call $s \in V$ the **supremum** of A if s is a least element of $U(A)$. Since least elements are unique (Exercise A.1.9), subsets of V can have at most one supremum. When it exists, the supremum of A is denoted by $\bigvee A$. Also,

- if $A = \{a_i\}_{i \in I}$ for some index set I, we write $\bigvee A$ as $\bigvee_i a_i$.

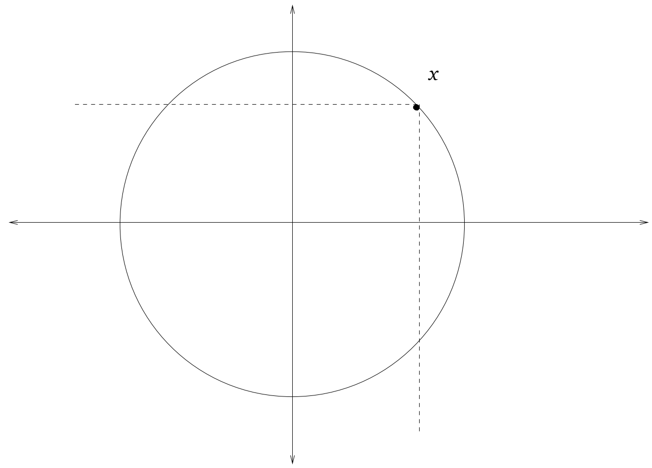


Figure A.1: The unit circle in $\mathbb{R}^2$ has no greatest element

- Given u and v in V, the supremum $\bigvee\{u, v\}$ is also called the **join** of u and v, and is written $u \vee v$.

Example A.1.10. If $V = (\mathbb{R}, \leqslant)$, then the notion of supremum for $A \subset \mathbb{R}$ reduces to the usual one from real analysis (see §A.1.1.1). In this setting (and only in this setting), we also use the traditional notation $\sup A$ to indicate the supremum of A.

It is well known that every $A \subset \mathbb{R}$ that is bounded above has a supremum (see Theorem A.1.1). This is not the case for arbitrary partially ordered sets.

Example A.1.11. Let C be the continuous functions from $[0, 1]$ into $\mathbb{R}$ paired with the pointwise order $\leqslant$. Consider the sequence of functions $F = \{f_n\}_{n \geqslant 2}$ illustrated in Figure A.2, where $f_n(x) = 0$ when $0 \leqslant x \leqslant 1/2$, $f_n(x) = n(x - 1/2)$ when $1/2 \leqslant x \leqslant 1/n + 1/2$ and $f_n(x) = 1$ otherwise. If $g \in U(F)$, the set of upper bounds of F in C, then, by continuity and the upper bound property, it must be that $g = 1$ on $[1/2, 1]$. Given any $g \in U(F)$, we can always take a $g' \in U(F)$ with $g' \leqslant g$ and $g'(x) < g(x)$ at at least one x. Hence $U(F)$ has no least element and, as a result, F has no supremum in C.

Exercise A.1.10. Let (u_n), (v_n) and (w_n) be sequences in a poset V with $u_n \preceq v_n \preceq w_n$ for all n. Suppose $\bigvee_n u_n = \bigvee_n w_n =: s$. Prove that $\bigvee_n v_n$ exists and is equal to s.

Figure A.3 provides a visualization of $u \vee v$ and $u \wedge v$ when $V = (\mathbb{R}^2, \leqslant)$. Figure A.4 provides a visualization of $f \vee g$ and $f \wedge g$ when $V = (\mathbb{R}^{\mathsf{X}}, \leqslant)$ for some subset X of the reals. In both cases, $\leqslant$ is the pointwise partial order.

Given A contained in poset V, an element of V is called the **infimum** of A if it is a greatest element of the set of lower bounds of A. The infimum of A is typically denoted $\bigwedge A$. If $V \subset \mathbb{R}$ with the usual order $\leqslant$, then we also use the notation $\inf A$. Also,

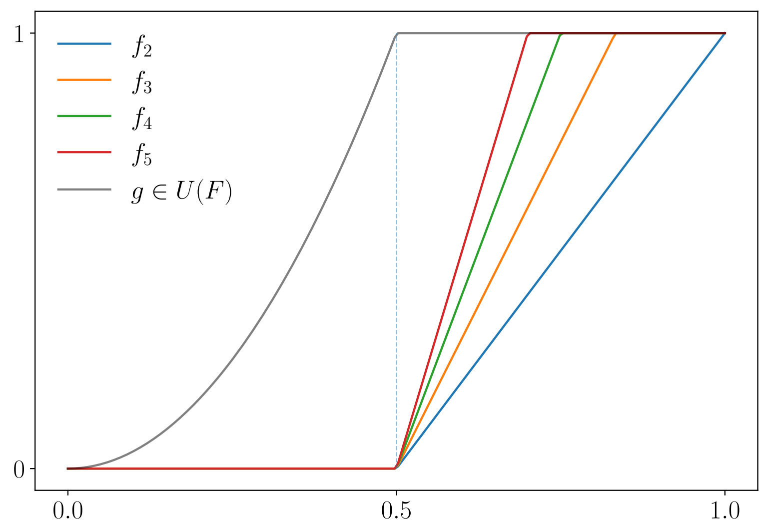


Figure A.2: The chain of functions F from Example A.1.11

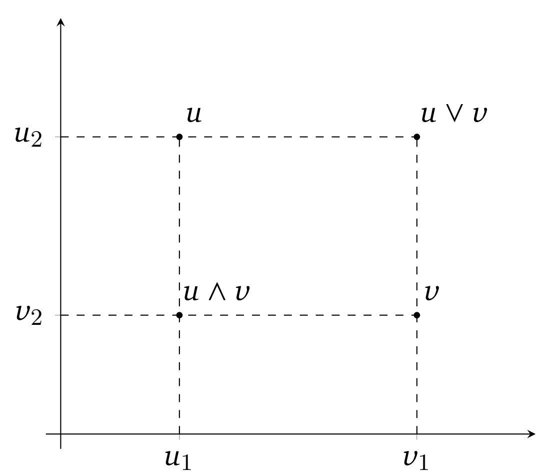


Figure A.3: The points $u \vee v$ and $u \wedge v$ in $\mathbb{R}^2$

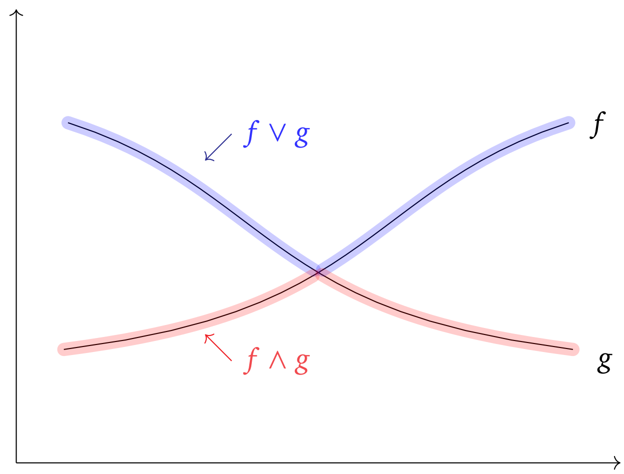


Figure A.4: Functions $f \vee g$ and $f \wedge g$ when defined on a subset of $\mathbb{R}$

- if $A = \{a_i\}_{i \in I}$ for some index set I, we sometimes write $\bigwedge A$ as $\bigwedge_i a_i$.
- Given u and v in V, the infimum $\bigwedge\{u, v\}$ is also called the **meet** of u and v, and is written $u \wedge v$.

EXERCISE A.1.11. Let X be any nonempty set, let $V \subset \mathbb{R}^{\mathsf{X}}$, and consider $(V, \leqslant)$ as a poset when $\leqslant$ is the pointwise partial order. Let G be a nonempty subset of V and let s and i be given by

$$s(x) := \sup_{g \in G} g(x) \quad \text{and} \quad i(x) := \inf_{g \in G} g(x) \qquad (x \in \mathsf{X}) \tag{A.2}$$

(The sup and inf on the right-hand side are follow the rules in §A.1.1, with s taking values in $(-\infty, +\infty]$ and i taking values in $[-\infty, \infty)$.) Prove that

(i) If $s \in V$, then $\bigvee G$ exists in V and $\bigvee G = s$.

(ii) If $i \in V$, then $\bigwedge G$ exists in V and $\bigwedge G = s$.

EXERCISE A.1.12. Let C be all continuous functions from $[0, 1]$ to itself paired with the pointwise order. Provide an example of a $G \subset C$ such that $\bigvee G$ exists in C and yet $\bigvee G$ is not equal to s, the pointwise supremum in (A.2).

EXERCISE A.1.13. Let V be any poset. Prove the following:

(i) If $b := \bigvee \varnothing$ exists in V, then V has a least element and that least element is b.

(ii) If b is the least element of V, then $\bigvee \varnothing = b$ in V.

(Analogous statements are true for the infimum of $\varnothing$ and the greatest element of V.)

EXERCISE A.1.14. Let $V = (V, \preceq)$ be any poset and let $I = [a, b]$ be an order interval in V. Suppose that $D \subset I$ and $s := \bigvee D$ exists in $(V, \preceq)$. Show that s is the supremum of D in $(I, \preceq)$; that is, in the poset I with the partial order inherited from V.

A.1.2.5 Order Duals

Given partially ordered set V, let $V^\partial = (V, \preceq^\partial)$ be the **order dual** (also called the **dual**), so that, for $u, v \in V$, we have $u \preceq^\partial v$ if and only if $v \preceq u$. We use $\bigvee^\partial A$ to denote the supremum of $A \subset V^\partial$ in V^∂ and $\bigwedge^\partial$ for the infimum.

EXERCISE A.1.15. Fix $A \subset V$ and prove the following.

(i) If $\bigvee A$ exists in V, then $\bigwedge^\partial A$ exists in V^∂ and $\bigwedge^\partial A = \bigvee A$.

(ii) If $\bigwedge A$ exists in V, then $\bigvee^\partial A$ exists in V^∂ and $\bigvee^\partial A = \bigwedge A$.

A.1.2.6 Monotone Sequences

Let V be any poset. A sequence $(v_n)_{n \geqslant 1}$ in V is called **increasing** if $v_n \preceq v_{n+1}$ for all $n \in \mathbb{N}$, and **decreasing** if $v_{n+1} \preceq v_n$ for all $n \in \mathbb{N}$. We write

- $v_n \uparrow v$ when (v_n) is increasing and $\bigvee_n v_n = v$ and
- $v_n \downarrow v$ when (v_n) is decreasing and $\bigwedge_n v_n = v$.

These symbols generalize standard notation for convergence of monotone sequences in $\mathbb{R}$. For example, if (u_n) is increasing in $\mathbb{R}$ and its limit is u, then one writes $u_n \uparrow u$. This is a special case of the usage above, since u is also the supremum of (u_n) under the standard order on $\mathbb{R}$.

EXERCISE A.1.16. Let (u_n) and (v_n) be sequences in V. Prove the following:

(i) If $v_n \uparrow v$ and $v_n \preceq u_n \preceq v$ for all n, then $\bigvee_n u_n = v$.

(ii) If $v_n \downarrow v$ and $v \preceq u_n \preceq v_n$ for all n, then $\bigwedge_n u_n = v$.

In some settings, the order theoretic concepts $\uparrow$ and $\downarrow$ have simple pointwise characterizations. The next lemma gives one such characterization for $\uparrow$ (and an analogous result holds for $\downarrow$). In the statement of the lemma, $V \subset \mathbb{R}^{\mathsf{X}}$ for some nonempty X and $\leqslant$ is the pointwise partial order. Also, we say that V is **closed under pointwise suprema** if, for every increasing $(v_n) \subset V$ that is bounded above, the pointwise supremum $s(x) = \sup_n v_n(x)$ is an element of V.

Lemma A.1.4. *Let $V \subset \mathbb{R}^{\mathsf{X}}$ be closed under pointwise suprema, let (v_n) be a sequence in V and let v be an element of V. In this setting,*

$$v_n(x) \uparrow v(x) \text{ in } \mathbb{R} \text{ for all } x \in \mathsf{X} \quad \iff \quad v_n \uparrow v.$$

Proof. ($\Rightarrow$) Let (v_n) be increasing, let v be in V and suppose that $v_n(x) \uparrow v(x)$ in $\mathbb{R}$ for all $x \in \mathsf{X}$. By Exercise A.1.11, $\bigvee_n v_n$ exists in V and equals v. Since (v_n) is increasing, we have $v_n \uparrow v$. ($\Leftarrow$) Suppose that $v_n \uparrow v$ for some $v \in V$. Fix $x \in \mathsf{X}$ and note that $v_n(x)$ is increasing and bounded above by $v(x)$. Hence the pointwise supremum function $s(x) = \sup_n v_n(x)$ exists in $\mathbb{R}^{\mathsf{X}}$ and $s \leqslant v$. Since V is closed under pointwise suprema, we also have $s \in V$. Since $v_n \leqslant s \leqslant v$ for all n and $\bigvee_n v_n = v$, we see that $s = v$. This means that, for any $x \in \mathsf{X}$, we have $\sup_n v_n(x) = v(x)$. Hence $v_n(x) \uparrow v(x)$. □

A.1.2.7 Order Preserving Maps

A self-map S from poset $V = (V, \preceq)$ to poset $U = (U, \lhd)$ is called

- **order preserving** if $v, w \in V$ and $v \preceq w$ implies $Sv \lhd Sw$, and
- **order reversing** if $v, w \in V$ and $v \preceq w$ implies $Sw \lhd Sv$.

Example A.1.12. Let $\leqslant$ be the pointwise partial order on $\mathbb{R}^n$. If A is an $n \times n$ matrix with nonnegative values and b is a vector in $\mathbb{R}^n$, then the affine operator S sending v to $Av + b$ is order preserving on $(\mathbb{R}^n, \leqslant)$. Indeed, if $u \leqslant v$, then $u - v \leqslant 0$ and hence $A(u - v) \leqslant 0$. It follows that $Au - Av \leqslant 0$ and, therefore, $Su - Sv \leqslant 0$.

In the definition of order preserving above, one common setting is when $U = \mathbb{R}$ with its standard order. In this case, the mapping S is often called **increasing**. We will also use this terminology. The result in the next exercise uses the fact that the standard order on $\mathbb{R}$ is closed (i.e., preserved under limits).

EXERCISE A.1.17. Let X be any poset set and let $b\mathsf{X}$ be the bounded real-valued functions on X paired with the supremum norm. Let $ib\mathsf{X}$ be the set of increasing functions in $b\mathsf{X}$. Prove that $ib\mathsf{X}$ is closed in $b\mathsf{X}$.

EXERCISE A.1.18. Let $(V, \preceq)$, be a partially ordered set and let $\mathscr{S}$ be the set of all order preserving self-maps on V. Let $\preceq$ be the the pointwise order on $\mathscr{S}$ (i.e., $S \preceq T$ if $Sv \preceq Tv$ for all $v \in V$). In this setting, prove the following statements:

(i) If $S \in \mathscr{S}$, then $S^k \in \mathscr{S}$ for all $k \in \mathbb{N}$.

(ii) If $S, T \in \mathscr{S}$ and $S \preceq T$, then $S^k \preceq T^k$ for all $k \in \mathbb{N}$.

A.1.2.8 Order Stability

In §1.1.2.2 we discussed the fact that contractivity of the Bellman operator plays a significant role in the proof of Bellman-type optimality results for the optimal savings problem. Contractivity is a metric property that has no immediate counterpart in an abstract partially ordered set. This motivates us to introduce weaker conditions on operators that are well defined in any poset and, at the same time, strong enough to generate useful optimality results. This section gives details.

Let V be a poset and let S be a self-map on V. In this setting, we call S

- **upward stable** on V if S has a unique fixed point $\bar{v}$ in V and, in addition, $v \in V$ with $v \preceq S\,v$ implies $v \preceq \bar{v}$,
- **strongly upward stable** on V if S has a unique fixed point $\bar{v}$ in V and, in addition, $v \in V$ with $v \preceq S\,v$ implies $S^n v \uparrow \bar{v}$,
- **downward stable** on V if S has a unique fixed point $\bar{v}$ in V and, in addition, $v \in V$ with $S\,v \preceq v$ implies $\bar{v} \preceq v$.
- **strongly downward stable** on V if S has a unique fixed point $\bar{v}$ in V and, in addition, $v \in V$ with $S\,v \preceq v$ implies $S^n v \downarrow \bar{v}$.

It is clear that strong upward stability implies upward stability, while strong downward stability implies downward stability.

We call S

- **order stable** on V if S is both upward and downward stable, and
- **strongly order stable** on V if S is strongly upward stable and strongly downward stable.

Figure A.5 gives an illustration of an order stable map S on $V = [0, 1]$. All points mapped up by S lie below its unique fixed point, while all points mapped down by S lie above its fixed point.

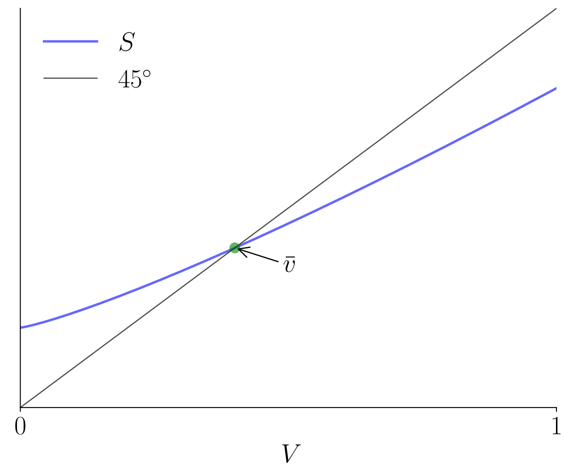


Figure A.5: An order stable map S on $[0, 1]$

EXERCISE A.1.19. Consider the self-map on $\mathbb{R}^k$ defined by $Sv = r + Av$ where $r \in \mathbb{R}^k$ and $A \in \mathbb{R}^{k \times k}$. Show that S is strongly order stable on $\mathbb{R}^k$ when $A \geqslant 0$ and $\rho(A) < 1$.

EXERCISE A.1.20. Let S be order preserving and upward stable on V, with fixed point $\bar{v}$. Prove that if $v \preceq Sv$, then $S^m v \preceq \bar{v}$ for all m.

The following result is useful when we consider minimization problems.

Lemma A.1.5. *S is order stable on V if and only if S is order stable on V^∂.*

Proof. Let S be as stated. By definition, S has a unique fixed point $\bar{v} \in V$. Hence it remains only to check that S is upward and downward stable on V^∂. Regarding upward stability, suppose $v \in V$ and $v \preceq^\partial Sv$. Then $Sv \preceq v$ and hence $\bar{v} \preceq v$, by downward stability of S on V. But then $v \preceq^\partial \bar{v}$, so S is upward stable on V^∂. The proof of downward stability is similar.

We have shown that S is order stable on V^∂ whenever S is order stable on V. The reverse implication holds because the dual of V^∂ is V. □

A.1.3 Metric Space

Add roadmap.

A.1.3.1 Definition

Let V be a nonempty set. A function $d\colon V \times V \to \mathbb{R}$ is called a **metric** on V if, for any $u, v, w \in V$,

$$
\begin{aligned}
& d(u,v) \geqslant 0, && \text{(nonnegativity)} \\
& d(u,v) = 0 \iff u = v, && \text{(identifiability)} \\
& d(u,v) = d(v,u) \text{ and} && \text{(symmetry)} \\
& d(u,v) \leqslant d(u,w) + d(w,v). && \text{(triangle inequality)}
\end{aligned}
$$

Together, the pair (V, d) is called a **metric space**. When the metric is clear from context we refer to the metric space by the symbol V alone.

Example A.1.13 (The space $b\mathsf{X}$)**.** Let X be any set. Let $b\mathsf{X}$ denote all bounded functions from X to $\mathbb{R}$. For all f, g in $b\mathsf{X}$, let

$$\|f\|_\infty := \sup_{x \in \mathsf{X}} |f(x)| \quad \text{and} \quad d_\infty(f,g) := \|f - g\|_\infty.$$

The map $f \mapsto \|f\|_\infty$ is called the **supremum norm** and d_∞ is called the **supremum distance**. The pair $(b\mathsf{X}, d_\infty)$ is a metric space. The triangle inequality holds because, given f, g, h in $b\mathsf{X}$ and $x \in \mathsf{X}$, we have (by the triangle inequality in $\mathbb{R}$),

$$|f(x) - g(x)| \leqslant |f(x) - h(x)| + |h(x) - g(x)| \leqslant d_\infty(f,h) + d_\infty(h,g).$$

The right side is an upper bound for the left side, so $d_\infty(f,g) \leqslant d_\infty(f,h) + d_\infty(h,g)$.

Exercise A.1.21. Complete the proof that d_∞ is a metric on $b\mathsf{X}$.

Example A.1.14 (The space $\ell_p(\mathsf{X})$)**.** Let X be finite or countable and fix p with $1 \leqslant p < \infty$. Let

$$\|h\|_p := \left\{ \sum_{x \in \mathsf{X}} |h(x)|^p \right\}^{1/p} \quad \text{and} \quad d_p(g,h) = \|g - h\|_p.$$

With $\ell_p(\mathsf{X}) := \left\{ h \in \mathbb{R}^\mathsf{X} : \|h\|_p < \infty \right\}$ the pair $(\ell_p(\mathsf{X}), d_p)$ is a metric space. The triangle inequality can be established via the **Hölder inequality** that states that $\|fg\|_1 \leqslant \|f\|_p \, \|g\|_q$ whenever $p, q \in [1, \infty]$ with $1/p + 1/q = 1$. In this setting the triangle inequality is also called the **Minkowski inequality**.

Example A.1.15 (Euclidean space)**.** If, in Example A.1.14, we take $\mathsf{X} = \{1, \ldots, n\}$ and $p = 2$, then $\mathbb{R}^\mathsf{X}$ is naturally identified with $\mathbb{R}^n$, the set of real-valued n vectors

$u = (u_1, \dots, u_n)$, while $d_2(u, v)$ is the ordinary Euclidean distance $(\sum_{i=1}^n (u_i - v_i)^2)^{1/2}$ between vectors u and v.

If (V, d) is a metric space and $N \subset V$, then (N, d) is also a metric space (where d in the second case is defined by restricting the original metric to $(u, v) \in N \times N$).

Example A.1.16 (The space $bc\mathsf{X}$)**.** Let X be a metric space and let $bc\mathsf{X}$ be all continuous functions in $b\mathsf{X}$. Since $bc\mathsf{X}$ is a subset of $b\mathsf{X}$, $(bc\mathsf{X}, d_\infty)$ is a metric space.

Exercise A.1.22. Let V be any nonempty set and consider the **discrete metric** on V given by $d(u, v) = \mathbb{1}\{u \neq v\}$. Prove that d is a metric on V (as suggested by the name).

A.1.3.2 Convergence

Given any point u in metric space (V, d), the **ε-ball** around u is the set

$$B_\varepsilon(u) := \{v \in V : d(u, v) < \varepsilon\}.$$

Analogous to the definition in normed linear space, we say that sequence $(u_n) \subset V$ **converges to $u \in V$** if

$$\forall \, \varepsilon > 0, \; \exists \, n_\varepsilon \in \mathbb{N} \text{ s.t. } n \geqslant n_\varepsilon \implies u_n \in B_\varepsilon(u).$$

Example A.1.17. Recall that a sequence (x_n) in $\mathbb{R}$ converges to $x \in \mathbb{R}$ if, given any $\varepsilon > 0$, there is an $N \in \mathbb{N}$ such that $|x_n - x| < \varepsilon$ for all $n \geqslant N$. This is equivalent to the statement that $x_n \to x$ in the metric space (V, d) when $V = \mathbb{R}$ and $d(x, y) = |x - y|$.

Exercise A.1.23. Let d be the discrete metric. Show that, for any $u \in V$, there exists an $\varepsilon > 0$ such that $B_\varepsilon(u) = \{u\}$. Show in addition that if (u_n) is a sequence in V converging to some point in V, then (u_n) is eventually constant.

Exercise A.1.24. Show that limits in metric spaces are unique. In other words, show that if $u_n \to u$ and $u_n \to y$ in a metric space V, then $u = y$.

Exercise A.1.25. Prove that a sequence (u_n) in a metric space V converge to a point u in V if and only if every subsequence of (u_n) also converges to u.

A metric space V is called **separable** if there exists a countable set $A \subset V$ such that, for any $v \in V$, there exists a sequence (a_n) contained in A with $a_n \to v$. For example, $\mathbb{R}$ is separable because any $v \in \mathbb{R}$ can be expressed as the limit of a rational sequence. Separability is useful in certain settings – particularly when we need to combine topology and measure (see, e.g., Theorem A.3.3). In the applications we consider, most spaces will be separable.

A.1.3.3 Open and Closed Sets

Let V be a metric space. A point $u \in A \subset V$ is called **interior** to A if there exists an $\varepsilon > 0$ such that $B_\varepsilon(u) \subset A$.

EXERCISE A.1.26. Let $V = \mathbb{R}$ and let $d(x, y) = |x - y|$. Show that 1 is interior to $A := [0, 1)$ but 0 is not. Show that $\mathbb{Q}$, the set of rational numbers in $\mathbb{R}$, contains no interior points.

EXERCISE A.1.27. Let V be arbitrary and let d be the discrete metric. Let A be any subset of V. Show that every point of A is interior to A.

A subset G of V is called **open** in V if every $u \in G$ is interior to G.

Example A.1.18. By Exercise A.1.27, every subset of a discrete metric space is open.

EXERCISE A.1.28. Let V be any metric space. Show that the ε-ball around u is open for any $u \in V$ and any $\varepsilon > 0$.

EXERCISE A.1.29. Show that

- Arbitrary unions and finite intersections of open sets in V are open in V.
- Arbitrary intersections and finite unions of closed sets in V are closed in V.

A subset F of V is called **closed** if given any sequence (u_n) satisfying $u_n \in F$ for all n and $u_n \to u$ for some $u \in V$, the point u is in F. In other words, F contains the limit points of all convergent sequences that take values in F.

Example A.1.19. Limits in $\mathbb{R}$ preserve orders, so $a \leqslant x_n \leqslant b$ for all $n \in \mathbb{N}$ and $x_n \to x$ implies $a \leqslant x \leqslant b$. Thus, any closed interval $[a, b]$ in $\mathbb{R}$ is closed in the standard (one dimensional Euclidean) metric.

Example A.1.20. As in the definition on page 163, let S be a metric space and let bcS be the set of all continuous functions in bS (see Example A.1.13 for the definition). The set bcS is a closed set in bS because uniform limits of continuous functions are continuous.

EXERCISE A.1.30. Let $\mathscr{C}^1$ denote all continuously differentiable functions f from $[-1, 1]$ to $\mathbb{R}$. As before let $d_\infty(f, g) = \sup_{x \in S} |f(x) - g(x)|$. The set $\mathscr{C}^1$ is *not* a closed subset of (bS, d_∞). To prove this, show that $d_\infty(f_n, f) \to 0$ as $n \to \infty$ when

$$f_n(x) := (x^2 + 1/n)^{1/2} \quad \text{and} \quad f(x) := |x|$$

Conclude that $(\mathscr{C}^1, d_\infty)$ is not closed.

EXERCISE A.1.31. Prove: If V is a metric space, then $G \subset V$ is open if and only if G^c is closed.

A.1.3.4 Compactness

A set D in V is called **bounded** if there exists a finite K such that $d(u, v) \leqslant K$ whenever $u, v \in D$.

EXERCISE A.1.32. Show that a subset D of V is bounded if and only if there exists an ε-ball $B_\varepsilon(u)$ such that $u \in D$ and $D \subset B_\varepsilon(u)$.

A subset K of V is called **precompact** in V if every sequence in K has a subsequence converging to some point in V. The set K is called **compact** if, in addition, the limit points always lie in K. Thus, K is compact if and only if K is closed and precompact.

EXERCISE A.1.33. Prove that every precompact subset of a metric space is bounded.

The converse is not true. For example, consider the set $b\mathbb{R}$ with the supremum distance (Example A.1.13). Let f_n be the normal density with variance 1 and mean n for each n in $\mathbb{N}$. The set $\{f_n\}_{n \in \mathbb{N}}$ is bounded, since $d_\infty(f_n, f_m) \leqslant 1$ for all n, m. But it is not precompact. For example, the sequence $\{f_n\}_{n \in \mathbb{N}}$ has not convergent subsequence. Indeed, every pair of distinct points f_n, f_m in the sequence has $d_\infty(f_n, f_m) = 1$.

(Pre)compactness resembles finiteness, as the next two exercises illustrate.

EXERCISE A.1.34. Let (V, d) be a metric space where d is the discrete metric. Show that a nonempty set $K \subset V$ is finite if and only if K is compact.

EXERCISE A.1.35. Show that (i) every subset of a precompact subset of V is also precompact in V, and (ii) the union of finitely many precompact (resp., compact) subsets of V is also precompact (resp., compact) in V

From Theorem A.4.5 on page 197, we recall that, in $\mathbb{R}^d$, the compact sets are precisely those that are both closed and bounded. In infinite dimensional space, however, this equivalence no longer holds.

Example A.1.21. Consider the space $(\ell_p(\mathsf{X}), d_p)$ defined on page 162. Let $p = 1$, let $\mathsf{X} = \{x_1, x_2, \ldots\}$ be countably infinite, and consider the sequence $\{h_n\} \subset \ell_1(\mathsf{X})$ defined by $h_n(x) = \mathbb{1}\{x = x_n\}$. For any $m \neq n$ we have $d_1(h_n, h_m) = \sum_x |h_n(x) - h_m(x)| = 2$. Clearly $\{h_n\}$ has convergent subsequence in $\ell_1(\mathsf{X})$. In particular, $\{h_n\}$ is bounded but not precompact.

A.1.3.5 Completeness

Let $V = (V, d)$ be a metric space. Analogous to the real case (see §A.1.1.2), a sequence $(u_n) \subset V$ is called **Cauchy** if, given any $\varepsilon > 0$, there exists an $n_\varepsilon \in \mathbb{N}$ such that $n, m \geqslant n_\varepsilon$ implies $d(u_n, u_m) < \varepsilon$. (V, d) is called **complete** if every Cauchy sequence in V converges in V.

Examples of complete spaces include $\mathbb{R}^n$ paired with any metric generated by a norm, the set of $n \times k$ matrices paired with any metric generated by a norm, the space $(\ell_p(\mathsf{X}), d_p)$ for countable X and $p \in [1, \infty]$, the space $(b\mathsf{X}, d_\infty)$, and the space $(bc\mathsf{X}, d_\infty)$.

EXERCISE A.1.36. Prove the following:

(i) If $V = (0, 1)$ and $d(u, y) = |u - y|$, then (V, d) is not complete.

(ii) If (V, d) is complete and $C \subset V$ is closed, then (C, d) is complete.

EXERCISE A.1.37. Let metrics d_1 and d_2 on V be **equivalent**, in the sense that there positive constants α, β such that $\alpha d_1(u, v) \leqslant d_2(u, v) \leqslant \beta d_1(u, v)$ for all $u, v \in V$. Show that (u_n) is Cauchy in (V, d_1) if and only if (u_n) is Cauchy in (V, d_2). Conclude that (V, d_1) is complete if and only if (V, d_2) is complete.

A.2 Topology

Topological spaces are a generalization of metric spaces. They are useful for two reasons. One is that there exist interesting and useful topological spaces that cannot be represented as metric spaces. The second is that, by stripping away some of the structure naturally present in metric spaces, topological arguments add simplicity and clarity to many discussions in analysis.

A.2.1 Topological Space

We begin by introducing topological spaces and investigating some of their core characteristics.

A.2.1.1 Definition and Examples

A **topological space** is a pair (V, τ) where V is a nonempty set and τ is a collection of subsets of V such that

(i) $\varnothing$ and V are both in τ,

(ii) τ is closed under finite intersections, and

(iii) τ is closed under arbitrary unions.

Statements (ii) and (iii) mean that

$$A, B \in \tau \implies A \cap B \in \tau \quad \text{and} \quad \mathscr{A} \subset \tau \implies \cup_{A \in \mathscr{A}} A \in \tau.$$

The family τ is called a **topology** on V. The elements of τ are called **open sets**. Complements of open sets are called **closed**.

Example A.2.1. Let V be any nonempty set. The set of all subsets of V is a topology on V, referred to as the **discrete topology**.

Example A.2.2. Let (V, d) be a metric space and τ let be the set of all open subsets of V (as defined in §A.1.3.3). The results in Exercise A.1.29 imply that τ is a topology on V. We call τ the topology **generated by** d.

Example A.2.3. If V is any set and $\mathscr{A}$ is any nonempty collection of subsets of V, then there exists a uniquely defined "smallest" topology τ that contains $\mathscr{A}$, constructed by taking the intersection of all topologies containing $\mathscr{A}$. (This intersection is not empty – at minimum, it contains the discrete topology – and one easily confirms that the intersection of a nonempty set of topologies is again a topology.) We call τ the topology **generated by** $\mathscr{A}$.

A subset N of a topological space $V = (V, \tau)$ is called a **neighborhood** of a point $v \in V$ if there exists a $G \in \tau$ with $x \in G \subset N$. A topological space V is called a **Hausdorff space** if, for any $u, v \in V$ with $u \neq v$, there exist neighborhoods N of u and M of v with $U \cap V = \emptyset$. All topological spaces we consider in this book are Hausdorff spaces.

EXERCISE A.2.1. Show that the topological space generated by a metric d on given set V is Hausdorff.

A.2.1.2 Nets

We briefly introduce nets, which are a generalization of a sequence. Nets are important because they characterize topologies, in a sense described below. At the same time, nets allow use to describe definitions and properties in a way that connects neatly to sequence-based definitions in metric spaces.

Let A be any nonempty set. A **preorder** on A is a relation $\preceq$ on $A \times A$ such that, for any a, b, c in A we have $a \preceq a$ (reflexivity) and $a \preceq b$ and $b \preceq c$ implies $a \preceq c$ (transitivity). Obviously any antisymmetric preorder on A is a partial order on A. A **directed set** is a nonempty set A and a preorder $\preceq$ on A such that, for any $a, b \in A$, there exists a $c \in A$ with $a \preceq c$ and $b \preceq c$.

Example A.2.4. The set of natural numbers $\mathbb{N}$ is a directed set when paired with the usual order $\leqslant$.

Example A.2.5. If L is a lattice then L is also a directed set.

Let V be any set. A **net** in V is a function from a directed set A to V, typically written as $v_\bullet$ or $(v_\alpha)_{\alpha \in A}$. We sometimes simplify the latter to (v_α). The interpretation is that $\alpha \in A$ is mapped to $v_\alpha \in V$. Obviously any sequence (v_n) in V is also a net in V.

A net $(v_\alpha)_{\alpha \in A}$ in V is said to **converge** to $v \in V$ and we write $v_\alpha \to v$ if, for any neighborhood N of v, there exists a $\beta \in A$ such that $v_\alpha \in N$ whenever $\beta \preceq \alpha$. This generalizes the concept of convergence of sequences in metric space. It is easy to check that convergent nets in V have unique limits whenever V is Hausdorff. (The converse is also true.)

The next theorem shows that nets can be used to characterize topologies.

Theorem A.2.1. *A subset C of a topological space V is closed in V if and only if every convergent net contained in C converges to an element of C.*

For a proof of Theorem A.2.1, see Theorem 2.14 of Aliprantis and Border (2006).

Let $(v_\alpha)_{\alpha\in A}$ and $(w_\beta)_{\beta\in B}$ be two nets in V. The net $(w_\beta)_{\beta\in B}$ is called a **subnet** of $(v_\alpha)_{\alpha\in A}$ if there exists an order preserving map p from B to A such that (i) $w_\beta = v_{p(\alpha)}$ for all $\alpha \in A$ and (ii) for all $\alpha \in A$, there exists a $\alpha' \in p(B)$ such that $\alpha \preceq \alpha'$.

Subnets generalize subsequences. For example, suppose that $w_n = 1/n^2$ and $v_n = 1/n$ for $n \in \mathbb{N}$, then $(w_n)_{n\in\mathbb{N}}$ is a subnet of $(v_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ (take $A = B = \mathbb{N}$ and $p(n) = n^2$).

A subset K of topological space V is called **compact** if, given any net (v_α) contained in K, there exists a subnet (w_β) of (v_α) and a point $v \in K$ such that $w_\beta \to v$. This generalizes the notion of a compact subset of a metric space, as given in §A.1.3.4.

A.2.1.3 Continuous Functions

Let V and W be topology spaces. A function $f : V \to W$ is said to be **continuous** at $v \in V$ if, for any net (v_α) in V with $v_\alpha \to v$ in V we have $f(v_\alpha) \to f(v)$ in W. If f is continuous at every $v \in V$ we simply say that f is continuous. It is well-known that f is continuous on V if and only if $f^{-1}(G)$ is open in V whenever G is open in W. (For a proof of this equivalence, see, e.g., Theorem 2.28 of Aliprantis and Border (2006).)

Example A.2.6. If V is paired with the discrete metric $\tau = \wp(V)$, then every function from V into another topological space W is continuous.

One of the most important features of continuous functions is that they carry compact sets into compact sets (see, e.g., §2.3 of Dudley (2002)):

Theorem A.2.2. *If f is a continuous function from topological space V to topological space W, then $f(K)$ is compact in W whenever K is compact in V.*

Exercise A.2.2. Prove: If T is a self-map on a Hausdorff space V with $T^m u \to u^*$ as $m \to \infty$ for some pair $u, u^* \in V$ and, in addition, T is continuous at u^*, then u^* is a fixed point of T.

A.2.1.4 Initial Topologies

Let V be a nonempty set and, for each α in index set Λ, let f_α be a function from V to topological space (W_α, τ_α). The **initial topology** generated by $\{f_\alpha\}_{\alpha\in\Lambda}$ is the topology τ on V generated (in the sense of Example A.2.3) by the family of sets

$$\mathscr{A} := \{f_\alpha^{-1}(G) : G \in \tau_\alpha,\ \alpha \in \Lambda\}.$$

Evidently each f_α is continuous with respect to τ on V. The following lemma nicely characterizes convergence with respect to the initial topology. In the statement, τ is the initial topology just described. Add ref.

Lemma A.2.3. *For net (v_α) in V and given $v \in V$, the following statements are equivalent:*

(i) *$v_\alpha \to v$ under the initial topology τ and*

(ii) *$f(v_\alpha) \to f(v)$ under τ_α for all $\alpha \in \Lambda$.*

A.2.1.5 Metrizable Spaces

A topological space (V, τ) is called **metrizable** if there exists a metric d on V such that d generates the topology τ. In metrizable spaces, sequences have the same "rights" as sequences in Euclidean space, in the sense that they determine the topology and hence other derived objects such as continuous functions. For example,

Theorem A.2.4. *Given two metrizable spaces V and W,*

(i) *a function $f : V \to W$ is* **continuous** *if and only if, for $u \in V$ and any sequence $(v_n) \subset V$ we have $f(v_n) \to f(v)$ in W whenever $v_n \to v$ in V.*

(ii) *A set C is closed in V if and only if any convergent sequence contained in C converges to an element of V.*

Exercise A.2.3. Prove the following:

(i) Every metrizable topological space is a Hausdorff space.

(ii) If d_1 and d_2 are equivalent metrics on V, then (V, d_1) and (V, d_2) generate the same topological space.

Part (ii) of Exercise A.2.3 shows why it is often nicer to discuss topologies than metrics. For example, we will see later (in §A.4.2) that all metrics on Euclidean space $\mathbb{R}^n$ generated by a norm are equivalent. Hence, while there are infinitely many norms on $\mathbb{R}^n$, they all generate the same topology. This means that, when discussing norm topologies, we can speak unambiguously about open sets, compact sets, continuous functions, etc.

A.2.1.6 Product Topologies

Let $\{(V_n, \tau_n)\}_{n\in\mathbb{N}}$ be a family of topological spaces and consider the Cartesian product $V = \prod_{n\in\mathbb{N}} V_n$. The i-th projection map on V is the function π_i sending $v = (v_n)_{n\in\mathbb{N}} \in V$ into v_i. The **product topology** on V, denoted here by τ, is the initial topology generated by the set of projection maps $\{\pi_n\}_{n\in\mathbb{N}}$. The following result is a direct consequence of Lemma A.2.3.

Lemma A.2.5. *If $(v_\alpha) = ((v_\alpha^1, v_\alpha^2, \dots))_{\alpha\in\Lambda}$ is a net in V and $v = (v^1, v^2, \dots)$ is an element of V, then $v_\alpha \to v$ in the product topology if and only if $v_\alpha^i \to v^i$ in (V_i, τ_i) for all $i \in \mathbb{N}$.*

Example A.2.7. Consider $\mathbb{R}$ with its usual topology, generated by the metric $d(u, v) = |u-v|$. Let set of n-vectors $\mathbb{R}^n$ is the Cartesian product of n copies of $\mathbb{R}$. The projections can be identified with the canonical basis vectors $e_1, \dots, e_n$, since, given $u \in \mathbb{R}^n$, the i-th projection is $\pi_i(u) = u_i = \langle u, e_i\rangle$. In view of Lemma A.2.3, a sequence (u_k) converges to $u \in \mathbb{R}^n$ in the product topology if and only if $\langle u_k, e_i\rangle \to \langle u, e_i\rangle$ in $\mathbb{R}$ for all i in $\{1, \dots, n\}$. In other words, a sequence in $\mathbb{R}^n$ converges in the product topology if and only if it converges pointwise.

Exercise A.2.4. Extending Example A.2.7, let $((\mathsf{X}_i, d_i))_{i=1}^n$ be n metric spaces and let $\mathsf{X} := \prod_i X_i$ be endowed with the product topology. Prove that a sequence $(u_k) \subset \mathsf{X}$ converges to a point $u \in \mathsf{X}$ if and only if $d_i(u_k, u) \to 0$ as $k \to \infty$ for all i in $\{1, \dots, n\}$.

A.2.1.7 Existence of Extrema

For finite subsets of $\mathbb{R}$, maxima and minima clearly exist. For infinite collections the same is not true. For example, the set $(0, 1)$ has neither a maximum nor a minimum.

Under what conditions on primitives are maxima and minima guaranteed to exist? There are multiple approaches to this issue, depending on the structure of the problem. In this section we treat one of the most fundamental, attributed to the German mathematician Karl Weierstrass (1815–1897).

Let f be a function from a metric space V to $\mathbb{R}$. Let (v_n) be an V-valued sequence and let v be a point in V. The function f is called

- **lower semicontinuous** at v if $v_n \to v$ implies $f(v) \leqslant \liminf_n f(v_n)$, and
- **upper semicontinuous** at v if $v_n \to v$ implies $f(v) \geqslant \limsup_n f(v_n)$.

If f is lower semicontinuous at every point in f, then f is called lower semicontinuous, and similarly for upper continuity.

EXERCISE A.2.5. Show that

(i) $f: V \to \mathbb{R}$ is continuous if and only if it is both lower and upper semicontinuous.

(ii) $f: V \to \mathbb{R}$ is lower semicontinuous on V if and only if, for each $\alpha \in \mathbb{R}$, the sublevel set $C_\alpha := \{x \in V : f(x) \leqslant \alpha\}$ is closed.

The next theorem generalizes the result in Exercise A.1.2. A proof can be found in Jahn (2020).

Theorem A.2.6 (Extreme Value Theorem, Weierstrass)**.** *Let K be a compact subset of V and let $f: K \to \mathbb{R}$.*

(i) *If f is lower semicontinuous, then f has a minimizer on K.*

(ii) *If f is upper semicontinuous, then f has a maximizer on K.*

In particular, if f is a continuous function from K to $\mathbb{R}$, then f has both a maximizer and a minimizer in K.

A.2.2 Stability and Contractions

One of the most important approaches to fixed points in metric space is via the theory of contractive maps. Here we review key results.

A.2.2.1 Fixed Points

Let V be any set and let S be a self-map on V. If $v \in V$ obeys $Sv = v$, then v is called a **fixed point** of S in V.

Example A.2.8. For example, if $V = \mathbb{R}$ and S is the identity, then every point in $\mathbb{R}$ is fixed under S. If, instead, $Sx = x^2$, then the set of fixed points is $\{0, 1\}$.

Now let V be a topological space. We call $S: V \to V$ **globally stable** on V if S has a unique fixed point $u^* \in V$ and $S^k u \to u^*$ as $k \to \infty$ for all $u \in V$.

EXERCISE A.2.6. Prove: If V is a Hausdorff space and $S: V \to V$ has a fixed point $u^* \in V$ with $S^k u \to u^*$ as $k \to \infty$ for all $u \in V$, then S is globally stable on V.

Lemma A.2.7. *Let V be Hausdorff and suppose that, for some $i \in \mathbb{N}$, the map S^i is globally stable on V with fixed point u^*. If, in addition, S is continuous at u^*, then S is globally stable on V with $Su^* = u^*$.*

Proof. Let S and i satisfy the stated hypotheses. Fix $u \in V$. We first claim that $S^k u \to u^*$ as $k \to \infty$. To verify this, observe that any integer m can be expressed as $ni+j$ for some integer n and some j in $0, \ldots, i-1$. Let U be a neighborhood of u^*. Globally stability of S^i allows us to choose, for each such j, an integer N_j such that $S^{ni+j}u = S^{ni}S^j u \in U$ whenever $n \geqslant N_j$. With $N = \max_j N_j$, we have $S^k u \in U$ whenever $k \geqslant Ni$, so $S^k u \to u^*$ as claimed.

Since S is continuous at u^*, Exercise A.2.2 implies that u^* is a fixed point of S. Applying Exercise A.2.6, S is globally on V. □

We will often make use of the following lemma.

Lemma A.2.8. *Let S be globally stable on topological space V with unique fixed point $v^* \in V$. If $U \subset V$ is closed in V and $SU \subset U$, then $v^* \in U$.*

Proof. Fix $u \in U$. Since S is globally stable on V, we have $v_n := S^n u \to v^*$ as $n \to \infty$. Since $SU \subset U$, we also have $v_n \in U$ for all n. As U is closed in V, these two facts imply that $v^* \in U$ (see Theorem A.2.1). □

A.2.2.2 Contractions

A self-map S on metric space $V := (V, d)$ is called **nonexpansive** if $d(Su, Sv) \leqslant d(u, v)$ for all u, v in V, and **contracting** or, more specifically, **a contraction of modulus λ** if there exists a $\lambda < 1$ such that

$$d(Su, Sv) \leqslant \lambda d(u, v) \quad \text{for all} \quad u, v \in V \tag{A.3}$$

Exercise A.2.7. Show that every contracting self-map S on V has at most one fixed point. Is the same true if S is only nonexpansive?

Exercise A.2.8. Prove: If S is nonexpansive on V, then S is continuous on V.

Theorem A.2.9 (Banach's contraction mapping theorem)**.** *If V is a complete metric space and $S \colon V \to V$ is a contraction of modulus λ on V, then S has a unique fixed point u^* in V and*

$$d(S^n u, u^*) \leqslant \lambda^n d(u, u^*) \quad \textit{for all } n \in \mathbb{N}$$

For a proof, see, for example, Aliprantis and Border (2006), Theorem 3.48.

A.2.2.3 Eventual Contractions

As stated in the next result, most of the conclusions of Banach's contraction mapping theorem carry over when S is eventually contracting. A proof can be found on p. 9 of Goebel and Kirk (1990).

Theorem A.2.10. *If V is complete, S is a self-map on V and S^k is a contraction on V for some $k \in \mathbb{N}$, then S is globally stable on V.*

As an application, consider the discrete Lyapunov equation $\Sigma = A\Sigma A^\top + M$ where all matrices are in $\mathbb{M}^{n\times n}$ and Σ is the unknown. To study this equation we introduced the Lyapunov operator $L(\Sigma) = A\Sigma A^\top + M$.

Exercise A.2.9. Prove that L is globally stable on $\mathbb{M}^{n\times n}$ when $r(A) < 1$.

Exercise A.2.10. Show that $L\Sigma^* = \Sigma^*$, where $\Sigma^* := \sum_{i\geqslant 0} A^i M (A^\top)^i$.

Exercise A.2.11. Show in addition that Σ^* is

(i) positive semidefinite whenever M is positive semidefinite and

(ii) positive definite whenever M is positive idefinite and A is nonsingular.

A.3 Measure and Integration

When solving analytical problems, we often need to place restrictions on the functions we consider. For example, it would be embarrassing if we proposed an algorithm for solving a given problem via a Taylor expansion and yet the functions we applied this algorithm to had kinks. These thoughts lead us to consider classes of "nice" functions, that are well behaved in one way or another. Linear functions on $\mathbb{R}$ are well behaved in many ways, as are polynomial functions. These, in turn, are a special case of the functions on $\mathbb{R}$ that have derivatives of every order. Such functions are special case of the "smooth" functions, which are those functions having continuous first derivative. The smooth functions are contained in the class of Lipschitz functions, which lie inside the class of everywhere continuous functions.

Need we generalize any more? Clearly the answer is affirmative. Sometimes economic variables exibit jumps. Agents make sudden changes in behavior. This means that we

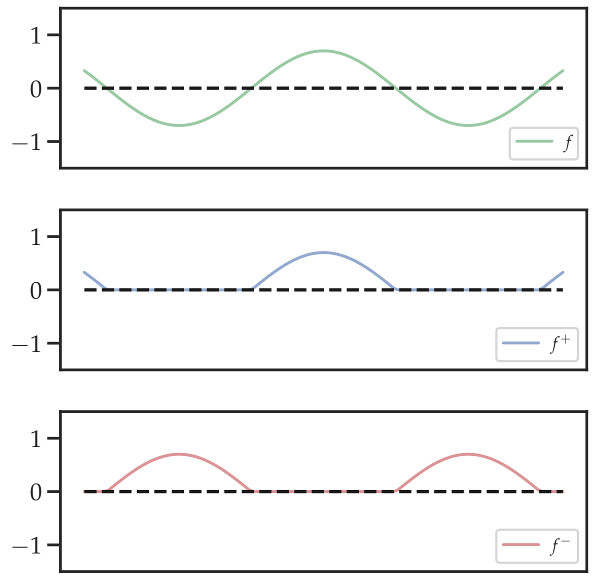


Figure A.6: Decomposition of functions

must admit discontinuities. At the same time, we do not wish to stray too far from continuity, so as to avoid dealing with the wildest functions mathematicians can dream up. This line of thought naturally leads us to the notion of measurable functions, which are defined below.

We will see that, depending on how we define them, the class of measurable functions can be very broad, encompassing all the types of functions listed above. At the same time, the class of measurable function are closed under the standard arithmetic and limiting operations. For example, sums and scalar multiples of measurable functions are measurable functions, as are pointwise limits, suprema and infima. Third, it turns out that measurable functions admit a well defined theory of integration with all the usual helpful properties.

In this section, we introduce measurable functions and then review foundational results from the theory of integration. Throughout, for real-valued f on an arbitrary domain, we set $f^+ := f \vee 0$ and $f^- := -(f \wedge 0)$. See Figure A.6 for an illustration. The function f^+ is called the **positive part** of f, while f^- is called the **negative part**.

The identity $f = f^+ - f^-$ always holds, so the pair f^+, f^- provides a decomposition of f into the difference between two nonnegative functions.

A.3.1 Measure Theory

Add roadmap.

A.3.1.1 Measurable Space

Let X be any nonempty set. A collection of subsets $\mathscr{A}$ of X is called a **σ-algebra** on X if

(i) $\mathsf{X} \in \mathscr{A}$,

(ii) $A \in \mathscr{A}$ implies $A^c \in \mathscr{A}$, and

(iii) if $\{A_n\}_{n \geqslant 1}$ is a sequence contained in $\mathscr{A}$, then $\cup_n A_n \in \mathscr{A}$.

A pair $(\mathsf{X}, \mathscr{A})$ where X is a nonempty set $\mathscr{A}$ is a σ-algebra on X is called a **measurable space**.

Points (ii) and (iii) tell us that $\mathscr{A}$ is "stable" under the taking of complements and unions. By De Morgan's law $(\cap_n A_n)^c = \cup_n A_n^c$, any σ-algebra is stable under countable intersections too. By (i) and (ii), $\varnothing \in \mathscr{A}$ also holds.

Example A.3.1. Given any set X and any subset $A \subset \mathsf{X}$, the family of sets $\mathscr{A} := \{\mathsf{X}, A, A^c, \varnothing\}$ is a σ-algebra on X.

Example A.3.2. The power set $\wp(\mathsf{X})$ is a σ-algebra on X, as is the pair $\{\varnothing, \mathsf{X}\}$.

Example A.3.3. The set of all circles in $\mathbb{R}^2$ is not a σ-algebra on $\mathbb{R}^2$. Indeed, this family is not stable under the taking of either unions or intersections. For the same reason, the set of all rectangles on $\mathbb{R}^2$ is not a σ-algebra on $\mathbb{R}^2$.

One way to define a σ-algebra is to take a collection $\mathscr{C}$ of subsets of X, and consider the smallest σ-algebra that contains this collection.

Definition A.3.1. Let $\mathscr{C}$ be any collection of subsets of X. The **σ-algebra generated by $\mathscr{C}$** is the smallest σ-algebra on X that contains $\mathscr{C}$, and is denoted by $\sigma(\mathscr{C})$.[1]

Now let X be a metric space. The family of **Borel sets** on X, denoted by either $\mathscr{B}$ or $\mathscr{B}_{\mathsf{X}}$ depending on whether or not the underlying space is clear, is defined as the σ-algebra generated by the open sets of X. Evidently $\mathscr{B}$ contains not only all the open subsets of X but also all the closed ones. From these sets we can continue taking complements and countable unions and everything we produce must be a Borel set. In fact it turns out that every set we work with in day-to-day analysis is a Borel set.

[1] More precisely, $\sigma(\mathscr{C})$ is the intersection of all σ-algebras on X that contain $\mathscr{C}$. One can show that $\sigma(\mathscr{C})$ is always a well defined σ-algebra, since the intersection is nonempty (it at least contains $\wp(\mathsf{X})$) and any intersection of σ-algebras is again a σ-algebra.

A.3.1.2 Measurable Functions

Functions are at the heart of all of our analysis. Functions without structure are difficult to manage so we put them into groups. In high school and in calculus, nice functions—the ones that we can manage—are those functions that are continuous or smooth. Working with continuous functions is convenient not only because these functions have attractive properties (think of the beautiful and powerful Mean Value Theorem) but also because they tend to reproduce themselves: Addition, subtraction, multiplication and composition of functions all preserve continuity.

As we move on to larger, more complicated problems, the set of continuous functions turns out to be too small. This is when we have to make the leap from continuous functions to *measurable functions*. Fortunately, measurable functions are also quite managable, and, just like continuous functions they tend to reproduce themselves. Let's review the key ideas.

Given two arbitrary measurable spaces $(\mathsf{X}, \mathscr{A})$ and $(\mathsf{Y}, \mathscr{B})$, a function f from X to Y is called $(\mathscr{A}, \mathscr{B})$-**measurable** if

$$f^{-1}(B) \text{ is in } \mathscr{A} \text{ whenever } B \in \mathscr{B}.$$

In other words, measurable functions are those functions that pull measurable sets back to measurable sets. If Y is a metric space and $\mathscr{B}$ is its Borel sets, then we will say that f is **Borel measurable**. It can be shown in this case (see, e.g., Çınlar (2011), Proposition 2.3) that f is Borel measurable if and only if either one of the following apparently weaker conditions are satisfied:

(i) $f^{-1}(G)$ is in $\mathscr{A}$ whenever G is open in Y

(ii) Y is a Borel subset of $\mathbb{R}$ and $f^{-1}((-\infty, \alpha))$ is in $\mathscr{A}$ for all $\alpha \in \mathbb{R}$.

From this result it is immediate that every continuous function from X to Y is also Borel measurable.

Exercise A.3.1. Let $(\mathsf{X}, \mathscr{A})$ be a measurable space and let B be a subset of X. Consider $\mathbb{1}_B$ as a map from X to $\mathbb{R}$. Show that $\mathbb{1}_B$ is a Borel measurable function if and only if $B \in \mathscr{A}$.

The definition of Borel measurability is not particularly intuitive, since the class of Borel sets is so large it's difficult to get a sense of what it does and doesn't contain. At the same time, Borel functions are ubiquitous in applied mathematics. Why?

When we solve dynamic programming problems and other problems from functional analysis, we start of by specifying a class of functions within which we hope to find a solution. What class should we use?

The class of continuous functions is the go-to class of "well behaved" functions in elementary mathematics. But while this class is closed under uniform limits (see example A.1.20), it is not closed under pointwise limits,[2] which makes it hard to work with in some instances. On the other hand, the set of Borel functions *is* closed under the taking of pointwise limits:

Lemma A.3.1. *If $(\mathsf{X}, \mathscr{A})$ is a measurable space and $\{f_n\}$ is a sequence of real valued Borel measurable functions on $(\mathsf{X}, \mathscr{A})$, then the functions*

$$f := \sup_n f_n, \quad f := \limsup_{n \to \infty} f_n, \quad \textit{and} \;\; f := \lim_{n \to \infty} f_n,$$

are all Borel measurable on $(\mathsf{X}, \mathscr{A})$ whenever they exist. The same is true if we replace sup with inf.

In fact, in our setting, the set of Borel measurable functions is precisely the smallest class of functions that contains the continuous functions and is closed under the taking of pointwise limits (see, e.g., §11.7 of Kechris (2012)).

It is also true that compositions of Borel measurable functions are also Borel measurable, and, when the functions are real-valued, that Borel measurability is preserved under algebraic operations. The next lemma gives one statement of these results:

Lemma A.3.2. *If $(\mathsf{X}, \mathscr{A})$ is a measurable space, α, β are real scalars and f and g are real-valued Borel measurable functions on $(\mathsf{X}, \mathscr{A})$, then the functions*

$$\alpha f + \beta g, \quad fg \quad \textit{and} \quad f/g \textit{ when } g \neq 0$$

are all Borel measurable functions on $(\mathsf{X}, \mathscr{A})$.

See Çınlar (2011), Chapter 1, Section 2 for proofs.

A.3.1.3 Parametric Continuity and Measurable Selections

We often wish to know whether or not continuity passes from primitives to solutions. For example, we might ask whether an equilibrium object, constructed through a process that involves optimization, varies continuously with parameters. The most commonly used theorem in this domain is **Berge's theorem of the maximum**. Here we state a version of Berge's theorem. Throughout, A and X are topological spaces.

[2]For example, the pointwise limit of the sequence of functions $\{f_n\}$ given by $f_n(x) = x^n$ on $[0, 1]$ is discontinuous.

A **correspondence** from X to A is a map Γ from X to the set of all subsets of A. A correspondence Γ from X to A is called **nonempty** if $\Gamma(x)$ is nonempty for all $x \in \mathsf{X}$. A function σ from X to A is called a **measurable selection** with respect to Γ if σ is Borel measurable and $\sigma(x) \in \Gamma(x)$ for all $x \in \mathsf{X}$.

A nonempty correspondence Γ is called

- **compact-valued** if $\Gamma(x)$ is compact for all $x \in \mathsf{X}$,
- **lower hemi-continuous** at $x \in \mathsf{X}$ if, for any $y \in \Gamma(x)$ and any (x_n) with $x_n \to x$, there exists a sequence (y_n) with $y_n \in \Gamma(x_n)$ for all n and $y_n \to y$,
- **upper hemi-continuous** at $x \in \mathsf{X}$ if, for any sequence (x_n) with $x_n \to x$ and any sequence (y_n) with $y_n \in \Gamma(x_n)$ for all n, there exists a convergent subsequence of (y_n) whose limit is in $\Gamma(x)$, and
- **continuous** on X if it is both lower and upper hemi-continuous at every $x \in \mathsf{X}$.

Exercise A.3.2. Let X and A be subsets of finite-dimensional Euclidean space and let g, h be continuous functions from X to A with $g(x) \leqslant h(x)$ for all $x \in \mathsf{X}$, where $\leqslant$ is the pointwise partial order. Let $\Gamma(x) = [g(x), h(x)]$. Prove that Γ is compact-valued and continuous on X.

Let Γ be a nonempty, compact-valued correspondence from X to A. Let q be a real valued function on $\mathsf{G} := \{(x, a) \in \mathsf{X} \times \mathsf{A} : a \in \Gamma(x)\}$ and set

$$m(x) := \max_{a \in \Gamma(x)} q(x, a) \qquad (x \in \mathsf{X}) \tag{A.4}$$

whenever the maximum is well defined.

Theorem A.3.3. *If Γ is continuous on X, q is continuous on G, and A is separable, then*

(i) *m is well defined and continuous on X, and*

(ii) *there exists a measurable selection σ such that $q(x, \sigma(x)) = m(x)$ for all x.*

If, in addition, the maximizer in (A.4) is unique at each x, then the uniquely defined measurable selection σ is continuous.

A proof of the continuity results in Theorem A.3.3 can be found in §17.5 of Aliprantis and Border (2006). Existence of a measurable selection is proved in §18.3 of the same reference.

A.3.1.4 Measures

Through the theory constructed above, we can identify broad classes of sets and functions that are relatively well behaved (e.g., Borel sets and Borel functions). This opens the way to analyzing how to (a) measure these sets and (b) integrate the functions. The first step is to introduce the notion of a **measure**, which is a map μ from a σ-algebra $\mathscr{A}$ to $[0, \infty]$ satisfying

(i) $\mu(\emptyset) = 0$ and

(ii) $\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ whenever $\{A_n\} \subset \mathscr{A}$ is disjoint.

Here disjointness of $\{A_n\}$ means that any two distinct sets in this sequence are disjoint.

Example A.3.4. Let $\mathsf{X} = \{x_1, x_2, \ldots\}$ be a countable set paired with $\wp(\mathsf{X})$, the set of all its subsets. Define $c \colon \wp(\mathsf{X}) \to \mathbb{R}_+$ by $c(A) = |A|$, where $|A|$ is the number of elements in A, with $c(A) = \infty$ if A is infinite. Some thought will convince you that c is a measure on $\wp(\mathsf{X})$. This measure is called the **counting measure**.

Example A.3.5. It can be proved (see, e.g., Dudley (2002), §3.2) that there exists exactly one measure on the Borel subsets of $\mathbb{R}^2$ that assigns area to rectangles in the usual way (i.e., area is the product of the two sides). Its is called **Lebesgue measure** and often denoted by λ. The measure λ also assigns the usual measures of area to other standard sets, such as circles. Indeed, if C is a circle with radius r, then $\lambda(C) = \pi r^2$. In other words, λ is the measure that gives us the classical notion of area in the plane, as taught in basic geometry.

Example A.3.6. Example A.3.5 introduced Lebesgue measure for $\mathbb{R}^2$. An analogous version, also called Lebesgue measure and denoted by λ, is defined on the Borel subsets of $\mathbb{R}^k$ for every $k \in \mathbb{N}$. For example, if $k = 1$ and I is an interval in $\mathbb{R}$, then $\lambda(I)$ is the length of that interval.

Returning to the general case of a measure μ on measurable space $(\mathsf{X}, \mathscr{A})$, if there exists a sequence of sets $(A_n) \subset \mathscr{A}$ with $\mu(A_n) < \infty$ for all n and $\cup_n A_n = \mathsf{X}$, then μ is called **σ-finite**. If $\mu(\mathsf{X}) < \infty$, then μ is called **finite**. If $\mu(\mathsf{X}) = 1$, then μ is called a **probability measure**.

If X is a metric space and $\mathscr{A} = \mathscr{B}$ (the Borel sets), then μ is called a **Borel measure**. If $\mathscr{A} = \mathscr{B}$ and $\mu(\mathsf{X}) = 1$, then μ is called a **Borel probability measure**. For a Borel probability measure μ, the value $\mu(B)$ usually is interpreted as the probability that, when a random element of X is selected, that element is in B.

Example A.3.7. Take the setting of example A.3.4 but now let the measure be given by $\nu(A) = \sum_{x \in A} p(x)$ instead of $|A|$, where p is a function from X to $\mathbb{R}_+$. It's not hard to see that ν defines a measure on $\wp(\mathsf{X})$. If $\sum_{x \in \mathsf{X}} p(x) < \infty$ then ν is a finite Borel measure. If the sum equals unity then ν is a Borel probability measure.

A **measure space** is a triple $(\mathsf{X}, \mathscr{A}, \mu)$ where $(\mathsf{X}, \mathscr{A})$ is a measurable space and μ is a measure on $\mathscr{A}$. If $\mu(\mathsf{X}) = 1$, then the measure space is also called a **probability space**. In this case it is common to write the measure space as $(\Omega, \mathcal{F}, \mathbb{P})$. A **random variable** on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is an $(\mathcal{F}, \mathscr{B})$-measurable map X from Ω to $\mathbb{R}$ paired with its Borel sets $\mathscr{B}$. More generally, given measurable space $(E, \mathscr{E})$, an E-valued **random element** on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is an $(\mathcal{F}, \mathscr{E})$-measurable map X from Ω to E. The **distribution** of this random element X is the probability measure P defined by

$$P(B) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in B\} \qquad (B \in \mathscr{E})$$

Here's a reassuring fact that implies Borel probability measures on standard sets like $\mathbb{R}$ are not strange creatures at all. Rather, they are fundamental objects that we've been using all along.

Theorem A.3.4 (Lebesgue-Stieltjes representation theorem). *There is a one-to-one correspondence between $\mathcal{F}$, the set of cumulative distribution functions on $\mathbb{R}$, and the set of Borel probability measures on $\mathbb{R}$. For each $F \in \mathcal{F}$, the corresponding probability measure μ satisfies*

$$\mu((a, b]) = F(b) - F(a) \text{ for all } a, b \in \mathsf{X} \text{ with } a < b$$

More generally, we have the interpretation

$$\mu(B) = \text{ probability that } x \in B \text{ when } x \text{ is drawn from } F$$

This representation in terms of probability measures is attractive because it assigns probabilities to subsets of X directly, rather than in the roundabout way that F does, and because measures can be defined in abstract settings that cdfs can't handle. Moreover, from measures we can construct a powerful theory of integration, a topic we turn to in §A.3.2.

A.3.2 Integration

[roadmap]

A.3.2.1 Abstract Integrals

When we study calculus, we learn a basic notion of integration—the "area under the curve" of a given function. The Fundamental Theorem of Calculus helps us find these areas, leading to results such as

$$\int_0^1 x^2 \, \mathrm{d}x = \frac{1}{3}. \tag{A.5}$$

We can also prove more general results, such as linearity of integration. That is,

$$\int_0^1 [\alpha f(x) + \beta g(x)] \, \mathrm{d}x = \alpha \int_0^1 f(x) \, \mathrm{d}x + \beta \int_0^1 g(x) \, \mathrm{d}x \tag{A.6}$$

for continuous functions f, g and scalars α, β.

While these results are elegant and important, we need a far more general notion of integration to make sense of all the problems treated in this text. To get a sense of why this is true, let's go right back to the objective function for the household in (1.1) on page 2. The expectation $\mathbb{E}$ used in this objective is, like all expectations, defined as a kind integral. But the formal definition is not trivial, because we are integrating over an infinity of shocks (due to the infinite horizon). How does one integrate over an infinite-dimensional space?

Let's answer this question in stages, beginning with an abstract notion of an integral. Our definition extends the elementary notion that an integral is a mapping that assigns numbers to functions (the area under their curve) and has certain monotonicity and linear properties (e.g., (A.6) above). To state the definition, we take $(\mathsf{X}, \mathscr{A})$ to be measurable space and $m\mathscr{A}_+$ to be the set of nonnegative real-valued Borel measurable functions on $(\mathsf{X}, \mathscr{A})$. We define an **integral** on $m\mathscr{A}_+$ to be a function $I \colon m\mathscr{A}_+ \to [0, \infty]$ such that

(i) $I(f) = 0$ when $f = 0$ everywhere on X,

(ii) $f_1 \leqslant f_2 \leqslant \cdots$ and $\lim_{n \to \infty} f_n = f$ implies $\lim_{n \to \infty} I(f_n) = I(f)$, and

(iii) $\alpha, \beta \geqslant 0$ and $f, g \in m\mathscr{A}_+$ implies $I(\alpha f + \beta g) = \alpha I(f) + \beta I(g)$.

The limit in (ii) is a pointwise limit, so that $\lim_{n \to \infty} f_n = f$ means $\lim_{n \to \infty} f_n(x) = f(x)$ for every $x \in \mathsf{X}$.

We can see that the standard Riemann integral taught in high school roughly fits this pattern (e.g., compare (iii) and (A.6) above). However, the details contain some crocodiles.[3] Fortunately, measure theory comes to the rescue here: we can use measure theory to define an integral obeying (i)—(iii) and implementing the Riemann integral as a special case. This integral is constructed using Lebesgue measure, as introduced in example A.3.5.

Rather than constructing this integral in isolation, let us state the following important result, proved in chapter 1 of Çınlar (2011). It states that *every* measure on a measurable space creates a unique and well defined integral.

Theorem A.3.5. *Let $(\mathsf{X}, \mathscr{A})$ be measurable space. There exists a one to one correspondence between the set of measures on $(\mathsf{X}, \mathscr{A})$ and the set of integrals on $m\mathscr{A}_+$. For any measure μ, the corresponding integral I_μ satisfies*

$$I_\mu(\mathbb{1}_B) = \mu(B) \text{ whenever } B \in \mathscr{A} \tag{A.7}$$

The value $I_\mu(f)$ is called the **integral of f under μ** and the following notation is common:

$$I_\mu(f) := \int f \, \mathrm{d}\mu := \int f(x)\mu(\mathrm{d}x).$$

Example A.3.8. Let $\mathscr{B}$ be the Borel sets on $\mathbb{R}$ and let λ be Lebesgue measure. According to theorem A.3.5, there exists an integral I_λ that associates to each Borel measurable function from $\mathbb{R}$ to $\mathbb{R}_+$ a number $I_\lambda(f)$, often written as $\int f \, \mathrm{d}\lambda$ or just $\int f(x) \, \mathrm{d}x$. If f is continuous and supported on an interval $[a, b]$, then $I_\lambda(f)$ equals $\int_a^b f(x)dx$ in the standard Riemann sense (see, e.g., Çınlar (2011), §1.4). For example, in the example in (A.5), with $f(x) = x^2$ on $[0, 1]$ and zero elsewhere, we have $\int f \, \mathrm{d}\lambda = 1/3$.

The integral I_λ introduced in example A.3.8 is called the the **Lebesgue integral**, and it extends the standard Riemann integral to a larger set of functions (the Borel measurable functions), while at the same time guaranteeing that the attractive properties (i)—(iii) in the definition of the integral will hold.

Equation (A.7) makes sense in this setting because if, say, $f = \mathbb{1}_{[a,b]}$, then

$$I_\lambda(f) = \lambda([a, b]) = b - a,$$

where the first equality is by (A.7) and the second is by the fact that Lebesgue measure assigns length to intervals. The value $b-a$ is also what we would expect for the integral,

[3]For example, limiting properties like (ii) are tricky because, for one thing, it is hard to guarantee that the limiting function f in (ii) is even Riemann integrable.

since it is the area under the curve for this simple function.[4]

Example A.3.9. Let $\mathsf{X} = \mathbb{N}$ and let c be the counting measure from example A.3.4. Let $x := \{x_n\}$ be any nonnegative sequence. We can view this sequence as a map from X to $\mathbb{R}_+$ and its integral is

$$\int x \, \mathrm{d}c = \sum_{i=1}^{\infty} x_i. \tag{A.8}$$

That is, ordinary series are just a kind of integral. We prove a generalization of (A.8) in example A.3.10 below.

Example A.3.10. As in example A.3.7, let X be countable, let $p \colon \mathsf{X} \to \mathbb{R}$ be nonnegative with $\sum_{x \in \mathsf{X}} p(x) < \infty$ and let ν be the measure on the σ-algebra $\wp(\mathsf{X})$ defined by $\nu(A) = \sum_{x \in A} p(x)$. Then, for any $f \colon \mathsf{X} \to \mathbb{R}_+$, the integral corresponding to ν is

$$\int f(x)\nu(\mathrm{d}x) = \sum_{x \in \mathsf{X}} f(x)p(x). \tag{A.9}$$

To see this, suppose first that f is zero off a finite set A contained in X. Then f can be written as

$$f(y) = \sum_{x \in A} f(x)\mathbb{1}_{\{x\}}(y) \qquad (y \in \mathsf{X}).$$

By the linearity property in part (iii) of the definition of the integral and the fact that, by definition, $\int \mathbb{1}_B \, \mathrm{d}\nu = \nu(B)$ for all $B \in \wp(\mathsf{X})$, we have

$$\int f(x)\nu(\mathrm{d}x) = \sum_{x \in A} f(x)\nu(\mathbb{1}_{\{x\}}) = \sum_{x \in A} f(x)p(x) = \sum_{x \in \mathsf{X}} f(x)p(x).$$

Hence (A.9) is valid. To handle arbitrary f, rather than just functions supported on finite sets, let f be any function from $\mathsf{X} = \{x_1, x_2, \ldots\}$ to $\mathbb{R}_+$ and let f_n be defined by $f_n(x_i) = f(x_i)\mathbb{1}\{i \leqslant n\}$. Each f_n is supported on a finite set, so

$$\int f_n \, \mathrm{d}\nu = \sum_{i \leqslant n} f_n(x_i)p(x_i) = \sum_{i \leqslant n} f(x_i)p(x_i).$$

Since $\{f_n\}$ is monotone increasing and converges to f, by part (ii) of the definition of

[4] A more general perspective on (A.7) that you might find useful is as follows. Suppose we identify measurable sets with their indicator functions. Then μ already provides us with an "integral" over the indicators in $m\mathscr{A}_+$. The map I_μ extends the reach of this function to all of $m\mathscr{A}_+$.

the integral we have

$$\int f(x)\nu(\mathrm{d}x) = \lim_{n\to\infty} \int f_n \,\mathrm{d}\nu = \sum_{i=1}^{\infty} f(x_i)p(x_i).$$

This is another way of writing (A.9).

A.3.2.2 Expectation

If μ is a probability measure and $w\colon \mathsf{X} \to \mathbb{R}$, then one often writes $\mathbb{E}w(x)$ for the integral of $w(x)$ with respect to μ. That is,

$$\mathbb{E}w(x) = \int w \,\mathrm{d}\mu$$

Here we are thinking of x as a random variable drawn from distribution μ and the integral corresponds to the **expectation** of $w(x)$ under μ.

A.3.2.3 Properties of Integrals

Define almost everywhere convergence. Isotone to increasing or order preserving. This section should include linearity of the integral.

From the properties in theorem A.3.5 we can deduce additional properties that the integral must satisfy. Before stating them, let us note that the notion of the integral extends to functions that take negative values, as well as just the nonnegative functions in $m\mathscr{A}_+$.

Indeed, if $(\mathsf{X}, \mathscr{A}, \mu)$ is a measure space and $f \in m\mathscr{A}$ is not necessarily nonnegative, then we can still decompose it into the difference between to nonnegative function via $f = f^+ - f^-$. Imposing linearity, we now set

$$\int f \,\mathrm{d}\mu := \int f^+ \,\mathrm{d}\mu - \int f^- \,\mathrm{d}\mu.$$

The only risk here is that both terms on the right equal $+\infty$, in which case the integral is not well defined. If both integrals are finite we call f **integrable** with respect to μ.

Now, let us state some general properties. In what follows, we leave $(\mathsf{X}, \mathscr{A}, \mu)$ fixed and write the integral $I_\mu(f)$ of f under μ as $\int f \,\mathrm{d}\mu$.

First, every integral is isotone (with respect to the pointwise ordering over $m\mathscr{A}$), in the sense that

$$f \leqslant g \implies \int f \,\mathrm{d}\mu \leqslant \int g \,\mathrm{d}\mu. \tag{A.10}$$

To see this, observe that $g-f$ is nonnegative (and measurable) and hence $\int (g-f)\,\mathrm{d}\mu$ is well defined and nonnegative. Now, using the linearity in part (iii) of theorem A.3.5, we have

$$\int g \,\mathrm{d}\mu = \int (g - f + f)\,\mathrm{d}\mu = \int (g-f)\,\mathrm{d}\mu + \int f \,\mathrm{d}\mu \geqslant \int f \,\mathrm{d}\mu.$$

Next we turn to limiting results. A battery of useful limit theorems exist for the integral we have defined. In our statements of these results, $(\mathsf{X}, \mathscr{A}, \mu)$ is any measure space and f and f_n are $(\mathscr{A}, \mathscr{B})$-measurable functions from X to $\mathbb{R}$ for all $n \in \mathbb{N}$.

Theorem A.3.6 (Fatou's lemma). *If $f_n \geqslant 0$ for all n and $f = \liminf f_n$, then*

$$\int f \,\mathrm{d}\mu \leqslant \liminf \int f_n \,\mathrm{d}\mu.$$

Fatou's lemma (upgraded here to a theorem) says, in essence, that the integral induced by any measure μ is lower semicontinuous on $m\mathscr{A}_+$. Proofs of Fatou's lemma and the following limit theorems can be found in §4.3 of Dudley (2002).

Theorem A.3.7. *Let $\lim_{n\to\infty} f_n = f$ hold μ-almost everwhere on X. If either*

(i) $-\infty < \int f_1 \,\mathrm{d}\mu$ *and* $f_n \leqslant f_{n+1}$ *for all* $n \in \mathbb{N}$, *or*

(ii) *there exists a* $g \in m\mathscr{A}_+$ *with* $\int g \,\mathrm{d}\mu < \infty$ *and* $|f_n| \leqslant g$ *for all* $n \in \mathbb{N}$, *then*

$$\lim \int f_n \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu. \tag{A.11}$$

The first implication (i.e., (i) $\implies$ (A.11)) is called the **monotone convergence theorem**. The second is called the **dominated convergence theorem**.

A.3.3 Product Spaces

This needs to be added, since we discuss product spaces and product measures – for example, in Exercise A.5.18.

A.3.4 Conditioning

Next we review prediction based on conditional expectations. Conditional expectations are themselves a cornerstone of economic theory and empirics, since they describe optimal forecasts based on limited information. We discussed conditional expectations in the context of Markov chains in Chapter **??**. Here we provide a brief treatment of the general setting that suffices for what follows.

A.3.4.1 Definition

Let Y and the elements of $\mathscr{G} := \{X_1, \ldots, X_k\}$ be scalar random variables. Consider the problem of predicting Y given $\mathscr{G}$. That is, we wish to form a prediction of the value that Y will take once $X_1, \ldots, X_k$ are known, without any additional information on the state of the world. Another way to say this is that we seek a (nonrandom) function $f\colon \mathbb{R}^k \to \mathbb{R}$ such that

$$\hat{Y} := f(X_1, \ldots, X_k) \text{ is a good predictor of } Y.$$

To find such an f we must define what "good" means. The most common definition in the present context is that **mean squared error** $\mathbb{E}[(\hat{Y} - Y)^2]$ is small. Thus, we have a minimization problem in function space (the set from which f is chosen). Based on projection arguments, it can be shown that there exists an essentially unique $\hat{f}$ in the set of functions from $\mathbb{R}^k$ to $\mathbb{R}$ that solves

$$\hat{f} = \operatorname*{argmin}_f \mathbb{E}[(Y - f(X_1, \ldots, X_k))^2]. \tag{A.12}$$

(See, e.g., Çınlar (2011).) We call the resulting variable

$$\hat{Y} := \hat{f}(X_1, \ldots, X_k)$$

the **conditional expectation** of Y given $\mathscr{G}$. Common alternative notations for $\hat{Y}$ include

$$\mathbb{E}_{\mathscr{G}} Y =: \mathbb{E}[Y \,|\, \mathscr{G}] =: \mathbb{E}[Y \,|\, X_1, \ldots, X_k].$$

In the present context, $\mathscr{G}$ is often called an **information set**.

A.3.4.2 Properties

In the next proposition, a random variable Y is called **$\mathscr{G}$-measurable** if there exists a function f such that $Y = f(X_1, \ldots, X_k)$. Intuitively, Y is perfectly predictable given the data in $\mathscr{G}$.

Proposition A.3.8. *Let X and Y be random variables with finite first moment and let $\mathscr{G}$ and $\mathscr{H}$ be information sets. The following properties hold:*

(i) $\mathbb{E}_{\mathscr{G}} X$ *is* $\mathscr{G}$*-measurable*

(ii) *If* $\mathscr{G} \subset \mathscr{H}$*, then* $\mathbb{E}_{\mathscr{G}}[\mathbb{E}_{\mathscr{H}} Y] = \mathbb{E}_{\mathscr{G}} Y$ *and* $\mathbb{E}[\mathbb{E}_{\mathscr{G}} Y] = \mathbb{E} Y$.

(iii) *If* Y *is independent of the variables in* $\mathscr{G}$*, then* $\mathbb{E}_{\mathscr{G}} Y = \mathbb{E} Y$.

(iv) *If* Y *is* $\mathscr{G}$*-measurable, then* $\mathbb{E}_{\mathscr{G}} Y = Y$.

(v) *If* X *is* $\mathscr{G}$*-measurable, then* $\mathbb{E}_{\mathscr{G}}[XY] = X\mathbb{E}_{\mathscr{G}} Y$.

(vi) $\mathbb{E}_{\mathscr{G}}[\alpha X + \beta Y] = \alpha \mathbb{E}_{\mathscr{G}} X + \beta \mathbb{E}_{\mathscr{G}} Y$ *for all* α, β *in* $\mathbb{R}$.

Property (i) states that the linearity of expectations is preserved under conditioning. Property (ii) is called the **law of iterated expectations**, and is shared by all projections. Property (v) is sometimes called **conditional determinism**, since X can be treated like a constant when it is pinned down by the information set. A full proof of Proposition A.3.8 can be found in Çınlar (2011).

A.3.4.3 Vector-Valued Conditional Expectations

If $Y = (Y_1, \ldots, Y_m)$ is a vector, then the conditional expectation of Y given information set $\mathscr{G}$ is the vector containing the conditional expectation of each element (similar to ordinary vector expectations). Thus, written as column vectors,

$$\mathbb{E}_{\mathscr{G}} \begin{pmatrix} Y_1 \\ \vdots \\ Y_m \end{pmatrix} = \begin{pmatrix} \mathbb{E}_{\mathscr{G}} Y_1 \\ \vdots \\ \mathbb{E}_{\mathscr{G}} Y_m \end{pmatrix}.$$

EXERCISE A.3.3. Prove: $\mathbb{E}_{\mathscr{G}}[AY + b] = A\mathbb{E}_{\mathscr{G}}[Y] + b$ for any $b \in \mathbb{R}^n$ and $k \times n$ matrix A.

A.4 Vector Spaces and Norms

We humans have natural geometric intuition about the space $\mathbb{R}^n$ when $n = 3$. If this intuition can be expressed algebraically, then $\mathbb{R}^3$ results often extend to $\mathbb{R}^n$ for arbitrary $n \in \mathbb{N}$ – and also to more general collections of objects, such as matrices, complex numbers and real-valued functions, provided that these collections are assigned some basic algebraic structure analogous to that enjoyed by vectors in $\mathbb{R}^3$.

Of course we need to formalize what "analogous" means by codifying the properties that we need the algebraic operations to satisfy. This leads to the concept of (abstract) vector space. In this section we recall the definition of such spaces and review key properties.

A.4.1 Vector Space

We begin with linear algebraic properties in abstract sets that generalize the idea of adding and scalar multiplying vectors in $\mathbb{R}^n$. Then we discuss properties of subsets of and maps over these abstract "vector spaces."

A.4.1.1 Definition and Properties

A **vector space** (also called a **linear space**) is a triple $(E, +, \cdot)$ where E is a nonempty set, $+$ is a map from $E \times E$ to E called **addition** and $\cdot$ is a map from $\mathbb{R} \times E$ to E called **scalar multiplication**, such that for all $u, v, w \in E$ and $\alpha, \beta \in \mathbb{R}$,

(i) $u + (v + w) = (u + v) + w$

(ii) $u + v = v + u$

(iii) there exists an element $0 \in E$, called the **origin**, s.t. $u + 0 = u$ for all $u \in E$

(iv) for all $u \in E$, there exists a $v \in E$ such that $u + v = 0$

(v) $\alpha \cdot (\beta \cdot u) = (\alpha \cdot \beta) \cdot u$

(vi) $1 \cdot u = u$

(vii) $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$

(viii) $(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u$

In practice, the $\cdot$ symbol is usually omitted, so $\alpha u := \alpha \cdot u$. In the present context, the values $\alpha, \beta, \ldots$ are often called **scalars**. Also, the origin, which shares the symbol 0 with the zero element from $\mathbb{R}$, is sometimes referred to as the **additive identity**.[5]

[5] Some authors would call what we have described as a *real* vector space, which can then be extended to the notion of complex vector spaces. We have no need for this extension here, so we drop the adjective "real."

Example A.4.1. The obvious example of a vector space is $\mathbb{R}^n$ with the usual notions of addition and scalar multiplication. The origin in (iii) is the n-vector of zeros, while v in (iv) is $-u$. All of the axioms are satisfied under this identification. (It would be highly surprising if this was not true, since $\mathbb{R}^n$ is the model for the axioms.)

Example A.4.2. Another example of a vector space is $\mathbb{R}^{n \times k}$, the space of real $n \times k$ matrices, under the usual rules

$$A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}) \quad \text{and} \quad \alpha B = \alpha(b_{ij}) = (\alpha b_{ij}).$$

The origin is the $n \times k$ matrix of zeros.

Example A.4.3. The set $\mathbb{R}^\mathsf{X}$ of real-valued functions on an arbitrary nonempty set X is a real vector space when paired with the usual notions of addition and scalar multiplication of functions: for $f, g \in \mathbb{R}^\mathsf{X}$ and $\alpha \in \mathbb{R}$, the functions $f + g$ and αf are defined by

$$(f + g)(x) := f(x) + g(x) \quad \text{and} \quad (\alpha f)(x) = \alpha f(x).$$

The zero element is $f \equiv 0$. Axioms (i)—(viii) are easily verified.

The vector space $\mathbb{R}^n$ is a special case of Example A.4.3, obtained when $\mathsf{X} = [n]$.

A.4.1.2 Convexity

Given vector space E, set $C \subset E$ is called **convex** if $u, v \in C$ and $\alpha \in [0, 1]$ implies $\alpha u + (1 - \alpha)v \in C$. In other words, C is closed under the taking of convex combinations.

Exercise A.4.1. Consider the set $\mathscr{D}(\mathsf{X})$ of **distributions** on a countable set X, defined as all $\varphi \in \mathbb{R}^\mathsf{X}$ such that $\varphi(x) \geqslant 0$ for all $x \in \mathsf{X}$ and $\sum_x \varphi(x) = 1$. Show that $\mathscr{D}(\mathsf{X})$ is a convex subset of $\mathbb{R}^\mathsf{X}$.

Exercise A.4.2. Let X be any set. For $g, h \in \mathbb{R}^\mathsf{X}$, the **order interval** $[g, h]$ in $\mathbb{R}^\mathsf{X}$ is all $f \in \mathbb{R}^\mathsf{X}$ such that $g \leqslant f \leqslant h$ pointwise on X. Show that $[g, h]$ is convex.

When E is any vector space, a nonempty subset C of E is called a **cone** in E if

(i) C is convex,

(ii) $x \in C$ and $-x \in C$ implies $x = 0$ and

(iii) $\alpha x \in C$ whenever $x \in C$ and $\alpha \geqslant 0$.

(Some authors refer to C as a "pointed convex cone.")

Add an exercise involving cones.

A.4.1.3 Linear Maps and Subspaces

Analogous to the case of $\mathbb{R}^n$, a **linear subspace** of vector space E is a set $S \subset E$ satisfying

$$\alpha, \beta \in \mathbb{R} \text{ and } u, v \in S \implies \alpha u + \beta v \in S. \tag{A.13}$$

EXERCISE A.4.3. Prove that the following sets are all linear subspaces of $\mathbb{R}^{n \times k}$.

(i) The set of all diagonal matrices.

(ii) The set of all upper triangular matrices.

(iii) The set of all symmetric matrices.

The proof of the next is also a useful exercise:

Proposition A.4.1. *If $(E, +, \cdot)$ is a vector space and S is a linear subspace of E, then $(S, +, \cdot)$ is itself is a vector space.*

Example A.4.4. Since $\mathbb{R}^{\mathsf{X}}$ is a vector space, to check whether or not $S \subset \mathbb{R}^{\mathsf{X}}$ is a vector space we just need to test whether (A.13) holds.

Example A.4.5. Let $b\mathsf{X}$ be the set of all bounded functions in $\mathbb{R}^{\mathsf{X}}$. This set is a linear subspace of $\mathbb{R}^{\mathsf{X}}$. Indeed, if f and g are bounded on X, then so is $\alpha f + \beta g$ for any scalars α and β, as follows from the triangle inequality. Now Proposition A.4.1 implies that $b\mathsf{X}$ is a real vector space in its own right.

Example A.4.6. Let $c\mathsf{X}$ be the set of continuous functions from metric space X to $\mathbb{R}$. Condition (A.13) holds for $c\mathsf{X}$ when treated as a subset of $\mathbb{R}^{\mathsf{X}}$, since continuity is preserved under addition and scalar multiplication. Hence $c\mathsf{X}$ is a linear subspace of $\mathbb{R}^{\mathsf{X}}$ and a vector space in its own right.

EXERCISE A.4.4. Let $\mathscr{P}_n$ be the set of all order n polynomials on $\mathbb{R}$, with each $p \in \mathscr{P}_n$ having the form $p(x) = \sum_{i=1}^{n-1} c_i x^i$ for some $(c_i) \in \mathbb{R}^n$. Show that $\mathscr{P}_n$ is a linear subspace of the set of all real valued functions on $\mathbb{R}$, and hence a real vector space in its own right.

A **linear operator** from vector space E into vector space F is a map $A \colon E \to F$ satisfying

$$\alpha, \beta \in \mathbb{R} \text{ and } u, v \in F \implies A(\alpha u + \beta v) = \alpha A u + \beta A v. \tag{A.14}$$

Example A.4.7. A matrix $A \in \mathbb{R}_{n \times k}$ is a linear operator from $\mathbb{R}^k$ to $\mathbb{R}^n$ when identified with the map $x \mapsto Ax$.

It can in fact be shown that every linear operator from $\mathbb{R}^k$ to $\mathbb{R}^n$ can be represented by an $n \times k$ matrix.

Example A.4.8. Let X be countable and consider the operator P mapping $h \in b\mathsf{X}$ into Ph in $b\mathsf{X}$ defined by

$$(Ph)(x) = \sum_{x' \in \mathsf{X}} p(x, x')h(x'), \tag{A.15}$$

where p is nonnegative and obeys $\sum_y p(x, x') = 1$ for all $x \in \mathsf{X}$.

Exercise A.4.5. Show that P (i) maps $b\mathsf{X}$ to itself and (ii) is linear on $b\mathsf{X}$.

The "kernel function" p in (A.15) operator can be identified with a matrix in $\mathbb{R}^{n \times n}$ when $|\mathsf{X}| = n \in \mathbb{N}$. No such identification exists when $|\mathsf{X}| = \infty$.

A.4.1.4 Bases and Dimension

A **linear combination** of vectors $u_1, \ldots, u_k$ in E is a vector of the form $\alpha_1 u_1 + \cdots + \alpha_k u_k$ where $\alpha_1, \ldots, \alpha_k$ are scalars. A set $S \subset E$ is called **linearly independent** if, for any finite set $\{u_1, \ldots, u_k\} \subset S$, we have

$$\alpha_1 u_1 + \cdots + \alpha_k u_k = 0 \text{ implies } \alpha_1 = \cdots = \alpha_k = 0.$$

A **basis** of a linear subspace S of E is a linearly independent subset B of S that spans S (i.e., each $u \in S$ can be expressed as a finite linear combination of elements of B).

Theorem A.4.2. *For a vector space E, the following statements are true:*

(i) *E has at least one basis.*

(ii) *If E has a basis with n elements, then every basis of E has n elements.*

(iii) *If E has an infinite basis, then every basis is infinite.*

A proof can be found in Jänich (1994). In case (ii), we say that E is **n-dimensional**. E is called **finite-dimensional** if E is n-dimensional for some $n \in \mathbb{N}$. In case (iii), we call E **infinite-dimensional**.

Example A.4.9. If X is finite, then $\mathbb{R}^{\mathsf{X}}$ is finite dimensional with dimension $|\mathsf{X}|$. A basis is provided by the functions f_i defined by $f_i(x) = \mathbb{1}\{x = i\}$. Any $g \in \mathbb{R}^{\mathsf{X}}$ can be expressed as a linear combination of these basis vectors via $g(x) = \sum_{i \in \mathsf{X}} g(i)\mathbb{1}\{x = i\}$.

Example A.4.10. $\mathbb{R}^\infty := \mathbb{R}^\mathbb{N}$ denotes the set of all real sequences $(x_i)_{i\in\mathbb{N}}$. This set is a vector space, as shown in Example A.4.3. Extending the canonical basis vectors, we can provide a basis for $\mathbb{R}^\infty$ via

$$\begin{aligned} e_1 &= (1, 0, 0, 0, \ldots) \\ e_2 &= (0, 1, 0, 0, \ldots) \\ e_3 &= (0, 0, 1, 0, \ldots) \\ &\vdots \end{aligned}$$

Hence $\mathbb{R}^\infty$ is infinite-dimensional.

Exercise A.4.6. Show that the matrix space $\mathbb{R}^{n\times k}$ has dimension nk.

A.4.2 Normed Vector Space

In this section we recall basic definitions and properties concerning normed vector space and linear operators acting on such space.

A.4.2.1 Norms on Vector Space

Given vector space E, a map $\|\cdot\|\colon E \to \mathbb{R}$ is called a **norm** on E if, for any $\alpha \in \mathbb{R}$ and any $u, v \in E$,

(a) $\|u\| \geqslant 0$ (nonnegativity)

(b) $\|u\| = 0 \iff u = 0$ (positive definiteness)

(c) $\|\alpha u\| = |\alpha|\|u\|$ and (positive homogeneity)

(d) $\|u + v\| \leqslant \|u\| + \|v\|$ (triangle inequality)

The pair $(E, \|\cdot\|) = ((E, +, \cdot), \|\cdot\|)$ is called a **normed vector space** (or **normed linear space**). When $\|\cdot\|$ is understood we refer to the space using the symbol E.

Example A.4.11. Euclidean vector space is the canonical example: The mapping defined on $\mathbb{R}^k$ by $\|u\| = \sqrt{\langle u, u\rangle}$ with $\langle u, u\rangle = \sum_{i=1}^k u_i^2$ is a norm on $\mathbb{R}^k$. The triangle inequality can be proved via the Cauchy–Schwarz inequality. The objective of studying normed linear spaces, as defined above, is to extend this canonical example and leverage its analysis to more general settings.

Example A.4.12. Let $b[0,1]$ be the vector space of continuous functions from $[0,1]$ to $\mathbb{R}$. Fix $p \geqslant 1$ and let

$$\|g\|_p := \left\{ \int_0^1 |g(t)|^p \, \mathrm{d}t \right\}^{1/p}, \tag{A.16}$$

It can be shown that $\|\cdot\|_p$ is a norm on $c[0,1]$. For example,

Exercise A.4.7. Prove the claim that $\|\cdot\|_p$ is a norm on $c[0,1]$ when $p = 1$.

Consider a normed vector space $(E, \|\cdot\|)$ with origin 0. A subset S of E is called **bounded** if there exists an $M \in \mathbb{N}$ such that $\|u\| \leqslant M$ for all $u \in S$.

Exercise A.4.8. Prove: Every linear subspace of E contains the origin. Also, if S is a bounded linear subspace of S, then $S = \{0\}$.

If $(E, \|\cdot\|)$ is a normed linear space, then (E, d) is a metric space when $d(u, v) := \|u - v\|$. All metric concepts extended to $(E, \|\cdot\|)$. For example, $G \subset E$ is said to be open in E if G is open in (E, d). A sequence (v_n) in E is said to converge to $v \in E$ and we write $v_n \to v$ if $d(v_n, v) \to 0$.

In a similar vein, if N is any subset of E, then (N, d) is also a metric space. Many metric spaces we encounter are of this type.

Example A.4.13. Let S be a subset of $\mathbb{R}$ and let ibS be all increasing functions in bS. This set is not a linear subspace of bS and hence $(ibS, \|\cdot\|_\infty)$ is not a normed linear space. However, ibS becomes a metric space when paired with d_∞.

Let $(E, \|\cdot\|)$ be a vector space and let $\|\cdot\|'$ be any other norm on E. If E is finite-dimensional, then these two norms are known to be **equivalent** (see, e.g., Aliprantis and Burkinshaw (1998), Theorem 27.6), which means that there exist finite positive constants A, B such that $\|x\| \leqslant A\|x\|'$ and $\|x\|' \leqslant B\|x\|$ for all $x \in E$.

Exercise A.4.9. In the setting described above, Prove that $(v_k) \subset E$ is convergent (resp., Cauchy) in $(E, \|\cdot\|)$, then (v_k) is convergent (resp., Cauchy) in $(E, \|\cdot\|')$.

A.4.2.2 Isometries

Suppose we wish to shown that a normed linear space $(E, \|\cdot\|_E)$ has some desirable property. One way to do this is to attack the proof directly. Another is to show that $(E, \|\cdot\|_E)$ is "similar" to another normed linear space that has the property in question.

To clarify the second approach, let E and W be two vector spaces. A map $T\colon E \to W$ is called an **isomorphism** if T is a linear bijection. If such a map exists, then the two spaces are said to be **isomorphic**.

Example A.4.14. Let $\mathscr{P}_n$ be the set of n-th order polynomials on $\mathbb{R}^n$, as defined in Exercise A.4.4. Consider the map T from $\mathscr{P}_n$ to $\mathbb{R}^n$ that sends $p(x) = \sum_{i=1}^{n-1} c_i x^i$ into its coefficient vector $(c_i) \in \mathbb{R}^n$. If $q(x) = \sum_{i=1}^{n-1} d_i x^i$, then the coefficient vector of $p+q$, is $(c_i + d_i)$, so

$$T(p+q) = (c_i + d_i) = (c_i) + (d_i) = Tp + Tq.$$

A simple extension to this argument shows that T is a linear operator. If two coefficient vectors differ then the polynomials differ at at least one point, so T is one-to-one. Clearly T is onto. Hence $\mathscr{P}_n$ and $\mathbb{R}^n$ are isomorphic.

Exercise A.4.10. Show that the property of being isomorphic is an equivalence relation on the set of all real vector spaces.

Exercise A.4.11. Prove: if E and W are isomorphic and $\dim E = n$, then $\dim W = n$.

Let $(E, \|\cdot\|_E)$ and $(W, \|\cdot\|_W)$ be two normed linear spaces. We call T an **isometric isomorphism** between these two spaces if T is an isomorphism between E and W and, in addition, $\|Tu\|_W = \|u\|_E$ for all $u \in E$. In this case, the two spaces are said to be **isometrically isomorphic**.

Exercise A.4.12. Let $(E, \|\cdot\|_E)$ and $(W, \|\cdot\|_W)$ be isometrically isomorphic under $T\colon E \to W$. Prove the following statements:

(i) If (v_k) converges in $(E, \|\cdot\|_E)$, then (Tv_k) converges in $(W, \|\cdot\|_W)$.

(ii) If (v_k) is Cauchy in $(E, \|\cdot\|_E)$, then (Tv_k) is Cauchy in $(W, \|\cdot\|_W)$.

(iii) If $(E, \|\cdot\|_E)$ is complete, then so is $(W, \|\cdot\|_W)$.

Theorem A.4.3. *If E is a real vector space with $\dim E = n$, then there exists a norm $\|\cdot\|$ on E such that $(E, \|\cdot\|)$ is isometrically isomorphic to $(\mathbb{R}^n, \|\cdot\|_1)$.*

To construct the norm $\|\cdot\|$ in Theorem A.4.3, let $B := \{b_1, \ldots, b_n\}$ be a basis of E, so that each $v \in E$ has a unique representation of the form $v = \sum_{i=1}^n v_i b_i$ for some $(v_1, \ldots, v_n) \in \mathbb{R}^n$. To each such v, we assign the norm

$$\|v\| := \sum_i |v_i|.$$

Let T map each v into its basis representation $(v_i) \in \mathbb{R}^n$.

EXERCISE A.4.13. Complete the proof of Theorem A.4.3 by showing that

(i) T is an isomorphism from E to $\mathbb{R}^n$ and

(ii) $\|v\| = \|Tv\|_1$ for all $v \in E$, where $\|\cdot\|_1$ is the ℓ_1 norm on $\mathbb{R}^n$.

A.4.2.3 Completeness

Completeness is essential to many important theorems in applied analysis. Fortunately, the completeness of $\mathbb{R}$ is inhereted by many useful spaces. For example,

Theorem A.4.4. *Every finite-dimensional normed vector space is complete.*

The proof is a solved exercise. Throughout, $\|\cdot\|_1$ is the ℓ_1 norm.

EXERCISE A.4.14. Prove that the space $(\mathbb{R}^n, \|\cdot\|_1)$ is complete. Then, using this fact and results stated above, prove Theorem A.4.4.

A complete normed vector space is called a **Banach space**. There are many other important Banach spaces, beyond the finite-dimensional ones.

Example A.4.15. Recall that in Example A.1.13 on page 162, we imposed a distance on f, g in $b\mathsf{X}$ via

$$d_\infty(f, g) := \|f - g\|_\infty \quad \text{where} \quad \|f\|_\infty := \sup_{x \in \mathsf{X}} |f(x)|$$

The pair $(b\mathsf{X}, \|\cdot\|_\infty)$ forms a Banach space. The completeness of this space is inherited from the completeness of $\mathbb{R}$ (see, e.g., section 3.2 of Aliprantis and Border (2006)).

Example A.4.16. We previously discussed the fact that $bc\mathsf{X}$ is a closed subset of $b\mathsf{X}$, and that closed subsets of complete metric spaces are complete. Hence $(bc\mathsf{X}, \|\cdot\|_\infty)$ forms a Banach space.

Example A.4.17. Following Example A.1.14 on page 162, we define

$$\|h\|_p := \left\{ \sum_{x \in \mathsf{X}} |h(x)|^p \right\}^{1/p} \quad \text{and} \quad \ell_p(\mathsf{X}) := \left\{ h \in \mathbb{R}^{\mathsf{X}} \,:\, \|h\|_p < \infty \right\}.$$

The pair $(\ell_p(\mathsf{X}), \|\cdot\|_p)$ is a Banach space. See §A.4.2.5 for more details.

A.4.2.4 Compactness

Let $(E, \|\cdot\|)$ be a normed vector space. A function (and hence a sequence) taking values in E is called bounded if its range is a bounded set.

EXERCISE A.4.15. Show that every convergent sequence in $(E, \|\cdot\|)$ is bounded.

A subset K of E is called **closed** if the limit of every convergent sequence in K lies in K. A subset K of E is called **compact** if every sequence in K has a subsequence converging to some point in K.

Compact sets share many properties with finite sets. For example, just as a function on a finite set F has a minimizing value and a maximizing value in that set, a continuous function on a compact set also has this property—a fact we will use many times. See §A.2.1.7 for more discussion.

Since all equivalent metrics induce the same precompact sets and the same bounded sets, *any* metric induced by a norm on a finite dimensional vector space has the property that its precompact and bounded sets coincide. More generally, the class of open sets, closed sets, bounded sets, compact sets and precompact sets in finite dimensional normed linear space does not depend on the particular norm being used. The next theorem states one of these facts for the record.

Theorem A.4.5 (Bolzano–Weierstrass)**.** *A subset of a finite dimensional normed vector space is compact if and only if it is closed and bounded.*

In infinite dimensional spaces, this one-to-one pairing between closed bounded sets and compact sets breaks down. In fact, the closed unit ball of a normed vector space E is compact if and only if E is finite-dimensional. When we shift to working with infinite-dimensional spaces, characterizing compact sets becomes a major task.

A.4.2.5 L_p Spaces

We know from Lemma A.3.2 that, given a measurable space $(\mathsf{X}, \mathscr{A})$, the class of Borel measurable functions from X to $\mathbb{R}$ will be closed under addition and scalar multiplication, and hence forms a vector subspace of $\mathbb{R}^{\mathsf{X}}$. It is, therefore, a vector space in its own right (see Proposition A.4.1). But we still need a norm.

To this end, let μ be a σ-finite measure on $(\mathsf{X}, \mathscr{A})$ and let $p \geqslant 1$. Let f be $\mathscr{A}$-measurable and consider the possibly infinite number

$$\|f\|_p := \int |f|^p \, \mathrm{d}\mu.$$

This looks like a norm but it isn't one yet. One issue is that it might be infinite. We can resolve this easily by defining $\mathcal{L}_p(\mathsf{X}, \mathscr{A}, \mu)$ to be the set of all Borel measurable $f \colon \mathsf{X} \to \mathbb{R}$ such that $\|f\|_p < \infty$. So $\| \cdot \|_p$ is finite on this set by construction.

However, there is still one more problem: We may have $f \neq 0$ and yet $\|f\|_p = 0$. This is because a function that is equal to zero everywhere except a set E such that $\mu(E) = 0$ has integral zero. Indeed, for such a function f,

$$\int f \,\mathrm{d}\mu = \int \mathbb{1}_E f \,\mathrm{d}\mu + \int \mathbb{1}_{E^c} f \,\mathrm{d}\mu = 0 + \int \mathbb{1}_{E^c} 0 \,\mathrm{d}\mu = 0$$

For example, when $\mathsf{X} = \mathbb{R}$ and μ is Lebesgue measure, the function $\mathbb{1}_{\mathbb{Q}}$ integrates to zero.

Apart from that, $\|\cdot\|_p$ has the other properties of a norm on $\mathcal{L}_p(\mathsf{X}, \mathscr{A}, \mu)$. For example, the triangle inequality holds as a result of the **Minkowski inequality**, which, in the present setting, states that, for $f, g \in \mathcal{L}_p(\mathsf{X}, \mathscr{A}, \mu)$,

$$\left\{ \int |f + g|^p \,\mathrm{d}\mu \right\}^{\frac{1}{p}} \leqslant \left\{ \int |f|^p \,\mathrm{d}\mu \right\}^{\frac{1}{p}} + \left\{ \int |g|^p \,\mathrm{d}\mu \right\}^{\frac{1}{p}} \tag{A.17}$$

For these reason $\| \cdot \|_p$ is referred to as a **seminorm**.

From our seminorm we can create something approaching a metric via

$$d_p(f, g) = \|f - g\|_p \quad (f, g \in \mathcal{L}_p(\mathsf{X}, \mathscr{A}, \mu))$$

It isn't quite a metric because $d_p(f, g) = 0$ does not imply $f = g$, since $\| \cdot \|_p$ is only a seminorm. Typically we refer to d_p as a **pseudometric**.

To turn a seminorm into a norm and a pseudometric into a metric, we regard all points at zero distance from each other as the same point. Formally, we partition the original space into *equivalence classes* of points at zero distance from one another, and consider the set of these classes as a new space. The distance between any two equivalence classes is just the distance between arbitrarily chosen members of each class. This value does not depend on the particular members chosen.

The normed linear space derived from the $\mathcal{L}_p(\mathsf{X}, \mathscr{A}, \mu)$ is traditionally denoted $L_p(\mathsf{X}, \mathscr{A}, \mu)$. Since two functions in $\mathcal{L}_p(\mathsf{X}, \mathscr{A}, \mu)$ are at zero distance if and only if they are equal μ-almost everywhere, the new space $L_p(\mathsf{X}, \mathscr{A}, \mu)$ consists precisely of equivalences classes of functions that are equal μ-almost everywhere.

Theorem A.4.6. *The space $L_p(\mathsf{X}, \mathscr{A}, \mu)$ paired with the norm $\| \cdot \|_p$ is a Banach space.*

Scheffés identity provides a useful quantitative interpretation of d_1 distance between densities: For any densities f and g on $(\mathsf{X}, \mathscr{A}, \mu)$, we have

$$\|f - g\|_1 = 2 \times \sup_{B \in \mathscr{A}} \left| \int_B f \, \mathrm{d}\mu - \int_B g \, \mathrm{d}\mu \right| \tag{A.18}$$

Finally, **Scheffés lemma** is useful for testing L_1 convergence:

Lemma A.4.7 (Scheffé). *If (f_n) and f are in $L_1(\mathsf{X}, \mathscr{A}, \mu)$ and $f_n \to f$ μ-almost everywhere as $n \to \infty$, then*

$$\int |f_n - f| \, \mathrm{d}\mu \to 0 \textit{ if and only if } \int |f_n| \, \mathrm{d}\mu \to \int |f| \, \mathrm{d}\mu.$$

In the case where f_n and f are densities, Scheff'e's lemma tells us that $f_n \to f$ in L_1 if and only if $f_n \to f$ almost everywhere.

A.4.3 Bounded Linear Operators

If E and F are normed linear spaces, then the **operator norm** of A is defined as

$$\|A\| := \sup_{\|u\|=1} \|Au\|. \tag{A.19}$$

(Here $\|u\|$ is the norm of u in E and $\|Au\|$ is the norm of Au in F.) When $\|A\|$ is finite, A is called a **bounded linear operator**. The set of all bounded linear operators from E to F will be denoted $\mathcal{B}(E, F)$. If $E = F$ then we write $\mathcal{B}(E)$. Every $A \in \mathcal{B}(E, F)$ is continuous, since, for $u_n \to u$ in E we have

$$\|Au_n - Au\| \leqslant \|A\| \|u_n - u\| \to 0.$$

The converse is also true: every continuous linear operator from E to F is bounded – see §2.7 of Kreyszig (1978) for a proof of this fact, as well as Theorem A.4.8 below.

Theorem A.4.8. *If F is a finite-dimensional, then every linear operator from E to F is bounded.*

Exercise A.4.16. Prove that $\|A\|$ equals the supremum of $\|Au\|/\|u\|$ over all $u \neq 0$.

As suggested by the name, the operator norm is a norm on $\mathcal{B}(E, F)$. The details are left as an exercise.

The operator norm is **submultiplicative**: If $A, B \in \mathcal{B}(E)$, then $\|AB\| := \|A \circ B\| \leqslant \|A\| \cdot \|B\|$. Iteratively applying the submultiplicative property gives $\|A^i\| \leqslant \|A\|^i$ for any $i \in \mathbb{N}$ and $A \in \mathcal{B}(E)$, where A^i is the i-th composition of A with itself.

Once we have a norm on $\mathcal{B}(E, F)$, we have an induced metric given by $d(A, B) = \|A - B\|$, and $\mathcal{B}(E, F)$ will be a Banach space whenever this metric is complete.

Theorem A.4.9. *If F is a Banach space, then $\mathcal{B}(E, F)$ is also a Banach space.*

Let E be a Banach space and let A be an element of $\mathcal{B}(E)$. A complex scalar λ is called an **eigenvalue** of $A \in \mathcal{B}(E)$ if there exists a nonzero vector e such that $Ae = \lambda e$. The **spectrum** of A, typically denoted $\sigma(A)$, is the set of all scalar λ such that $\lambda I - A$ fails to be bijective on E. Any eigenvalue λ lies in $\sigma(A)$ because if $Ae = \lambda e$ for some nonzero e, then $\lambda I - A$ maps e to 0, while also mapping 0 to 0. Hence $\lambda I - A$ is not bijective. For $A \in \mathcal{B}(E)$, the **spectral radius** of A is defined as

$$\rho(A) := \sup\{|\lambda| : \lambda \in \sigma(A)\} \tag{A.20}$$

It is well known (see, e.g., Kreyszig (1978), §7.3) that

(i) $\rho(A) \leqslant \|A\|$, where $\|\cdot\|$ is the operator norm, and

(ii) $\|A^k\|^{1/k} \to \rho(A)$ as $k \to \infty$ (**Gelfand's formula**).

EXERCISE A.4.17. Let φ be a probability measure on measurable space $(\mathsf{X}, \mathscr{A})$ and let β be a positive real number. Let K be the linear operator on $L_1(\varphi) := L_1(\mathsf{X}, \mathscr{A}, \varphi)$ defined by $Kv = \beta\|v\|\mathbb{1}$ for all $v \in L_1(\varphi)$. Prove that $\rho(K) = \beta$.

The following theorem is essential for many results in the book.

Theorem A.4.10 (Neumann series lemma)**.** *If E is a Banach space, A is an element of $\mathcal{B}(E)$ and $\rho(A) < 1$, then $I - A$ is nonsingular and*

$$(I - A)^{-1} = \sum_{j=0}^{\infty} A^j$$

(The infinite sum is defined as the limit of the partial sums in E. Hence the infinite sum exists if and only if the partial sums converge in E.)

Proof. First observe that the sequence $B_n := \sum_{i=0}^{n} A^i$ is Cauchy when $\rho(A) < 1$. Indeed, using the operator norm,

$$\|B_k - B_{k+n}\| \leqslant \left\| \sum_{i \geqslant k}^{\infty} A^i \right\| \leqslant \sum_{i \geqslant k}^{\infty} \|A^i\|.$$

The final term will converges to zero in k if $\sum_{i=0}^{\infty} \|A^i\|$ is finite. By the root test for convergence of series, this will be true whenever we have $\limsup_{i\to\infty} \|A^i\|^{1/i} < 1$. We know this is true by the hypothesis $\rho(A) < 1$ and Gelfand's formula.

Since $\mathcal{B}(E)$ is complete under the operator norm, this Cauchy property implies that the limit $\sum_{i=0}^{\infty} A^i$ exists. Moreover, $(I - A)\sum_{i=0}^{\infty} A^i = I$, since

$$\left\|(I-A)\sum_{i=0}^{\infty} A^i - I\right\| = \lim_{n\to\infty}\left\|(I-A)\sum_{i=0}^{n} A^i - I\right\| = \lim_{n\to\infty}\left\|A^{n+1}\right\|$$

and the right hand side converges to zero by $\rho(A) < 1$ and Gelfand's formula. □

Corollary A.4.11. *Fix $A \in \mathcal{B}(E)$ and $b \in E$, where E is a Banach space. If $\rho(A) < 1$, then the operator T on E defined by $Tv = Av + b$ is globally stable, with unique fixed point in E given by*

$$v^* := (I - A)^{-1} b = \sum_{t \geqslant 0} A^t b.$$

Exercise A.4.18. Prove Corollary A.4.11.

Exercise A.4.19. Let $\mathbb{M}$ be the set of $k \times k$ real matrices and define L by $LP = A^\top P A$ at each $P \in \mathbb{M}$. Prove the following statements.

(i) L is in $\mathcal{B}(\mathbb{M})$, the set of bounded linear operators on $\mathbb{M}$.

(ii) $\rho(A) < 1$ implies $\rho(L) < 1$, where ρ is the spectral radius on $\mathbb{M}$ and $\mathcal{B}(\mathbb{M})$ respectively, and

(iii) Given $B \in \mathbb{M}$, the condition $\rho(A) < 1$ implies that the operator $T := \mathbb{M} \to \mathbb{M}$ defined by $TP = B + LP$ has the unique fixed point $P^* := \sum_{t\geqslant 0} L^t B$.

[Hint: $\|A^\top\| = \|A\|$ when $\|\cdot\|$ is the operator norm on $\mathcal{B}(\mathbb{R}^k)$.]

Exercise A.4.20. Following on from Exercise A.4.19, prove in addition that, when $\rho(A) < 1$ and $b \in \mathbb{R}^k$, the unique solution to $x = b + Lx$ in $\mathbb{R}^k$ is

$$x^* := \sum_{t\geqslant 0} d_t \quad \text{where } d_0 = b \text{ and } d_{t+1} = A^\top d_t A.$$

A.5 Order

Roadmap.

A.5.1 Order Continuity and Order Completeness

When studying the real line $\mathbb{R}$, we can define completeness either as existence of limits for Cauchy sequences (Theorem A.1.3) or as existence of suprema for bounded above sets (Theorem A.1.1). The first idea can be extended to metric spaces by generalizing the concept of Cauchy sequences (see §A.1.3.5). The second can be extended to posets by analogy with existence of suprema. The aim of this section is to describe this second concept of completeness.

A.5.1.1 Lattices and Chains

When we discuss posets, there are multiple notions of completeness, with each one determined by the classes of sets that are required to have suprema. Specifically, a nonempty poset V is called

- a **lattice** if every finite subset of V has both a supremum and an infimum in V,
- a **complete lattice** if every subset of V has a supremum and an infimum in V,
- **chain complete** if every chain in V has a supremum and an infimum in V, and
- **σ-chain complete** if every at most countable chain in V has a supremum and an infimum in V.[6]

Here's a diagram that illustrates the relationships

$$\text{complete lattice} \begin{array}{l} \nearrow\!\!\!\nearrow \ \text{lattice} \\ \searrow\!\!\!\searrow \ \text{chain complete} \implies \sigma\text{-chain complete} \end{array}$$

In the definitions above, finite sets are understood to be nonempty (i.e., in one-to-one correspondence with $\{1, \ldots, n\}$ for some $n \in \mathbb{N}$). Also, "at most countable" means empty, finite or countable. The fact that the empty set is included has significance, as the next lemma illustrates.

Lemma A.5.1. *If V is σ-chain complete, then V is order bounded.*

Proof. This follows from Exercise A.1.13 on page 157, since $\varnothing$ is a chain in V. □

[6] It's unfortunate that the symbol σ does double duty in this book, representing countability (as in σ-continuous and σ-algebra) as well as, in the ADP setting a typical policy σ in the index set Σ. We require this extra duty in order to maintain consistency with the literature (e.g., Zaanen (2012) and Sargent and Stachurski (2025a)).

Example A.5.1. If $V = [g, h]$ is all $f \in \mathbb{R}^{\mathsf{X}}$ with $g \leqslant f \leqslant h$, then $(V, \leqslant)$ is a complete lattice. To see this, observe first that if F is any nonempty subset of V, then $s(x) := \sup_{f \in F} f(x)$ and $i(x) := \inf_{f \in F} f(x)$ are well-defined elements of V, and also the supremum and infimum of F in V by Exercise A.1.11. Second, if F is empty, then we take $\bigvee F = g$ and $\bigwedge F = h$, which are again the supremum and infimum of F in V (see Exercise A.1.13). Either way, every subset of V has a supremum and an infimum.

Exercise A.5.1. Given an example of a lattice that is not chain complete.

Many readers will be familiar with the Knaster–Tarski fixed point theorem, which shows that the set of fixed points of an order preserving self-map on a complete lattice is nonempty. The following theorem is closely related. For ease of reference, we also refer to it as the Knaster–Tarski fixed point theorem. In the statement, V is a nonempty poset.

Theorem A.5.2 (Knaster–Tarski). *If V is chain complete and S is an order preserving self-map on V, then S has a fixed point in V.*

Proof. See, for example, Theorems 8.11 and 8.22 of Davey and Priestley (2002). □

Theorem A.5.2 shows that chain completeness is related to existence of fixed points. The next lemma shows that chain completeness is also related to order stability.

Lemma A.5.3. *Let S be an order-preserving self-map on V. If S has at most one fixed point in V and V is chain complete, then S is order stable on V.*

Proof. First suppose that V is chain complete, with greatest element $\top$ and least element $\perp$. Fix $v \in V$ with $Sv \preceq v$. Since $I := [\perp, v]$ is itself chain complete, and since S maps I to itself and is order preserving, the Knaster–Tarski fixed point theorem implies that S has a fixed point $\bar{v}$ in I. By assumption, $\bar{v}$ is the only fixed point of S in V. Moreover, $\bar{v} \in I$, so $\bar{v} \preceq v$. This proves downward stability of S on V. A similar argument proves upward stability. □

A **sublattice** of a lattice V is a subset S of V with the property that $u \vee v$ and $u \wedge v$ are in S whenever $u, v \in S$.

Exercise A.5.2. Let X be a metric space and let $bc\mathsf{X}$ be the bounded continuous functions from X to $\mathbb{R}$. Prove that $bc\mathsf{X}$ is a sublattice of $\mathbb{R}^{\mathsf{X}}$.

A.5.1.2 Dedekind Completeness

Consider the canonical partially ordered set $(\mathbb{R}^k, \leqslant)$. This set is not σ-chain complete: for example, letting $\mathbb{1}$ be a vector of ones, the increasing sequence $(v_n) = (n\mathbb{1})$ has no supremum. At the same time, $(\mathbb{R}^k, \leqslant)$ certainly has some completeness properties. For example, it follows easily from Exercise A.1.11 that every bounded above subset of $\mathbb{R}^k$ has a supremum, and every bounded below subset of $\mathbb{R}^k$ has an infimum. This motivates the following definitions:

A partially ordered set V is called **Dedekind complete** if, for any nonempty $A \subset V$,

(i) A is bounded above $\implies$ A has a supremum in V and

(ii) A is bounded below $\implies$ A has an infimum in V.

V is called **σ-Dedekind complete** if, for any nonempty finite or countable $A \subset V$,

(i) A is bounded above $\implies$ A has a supremum in V and

(ii) A is bounded below $\implies$ A has an infimum in V.

Example A.5.2. If X is any nonempty set, then $(\mathbb{R}^{\mathsf{X}}, \leqslant)$ is Dedekind complete. Indeed, if $G \subset \mathbb{R}^{\mathsf{X}}$ is nonempty and bounded above, then $s(x) = \sup_{g \in G} g(x)$ exists in $\mathbb{R}$ at each $x \in \mathsf{X}$. Hence, by Exercise A.1.11, the supremum $\bigvee G$ exists in $\mathbb{R}^{\mathsf{X}}$ (and equals s). A similar argument shows that any nonempty bounded below subset of $\mathbb{R}^{\mathsf{X}}$ has an infimum.

Example A.5.3. The set of continuous functions from $[0, 1]$ into $\mathbb{R}$ paired with the pointwise order is a lattice but not σ-Dedekind complete. You can verify this by rereading Example A.1.11 on page 155.

There are natural connections between Dedekind (resp., σ-Dedekind) completeness and chain (resp. σ-chain) completeness. Here is one simple result.

Lemma A.5.4. *Let V be any poset.*

- *If $I = [a, b] \subset V$ and V is Dedekind complete, then $(I, \preceq)$ is chain complete.*
- *If $I = [a, b] \subset V$ and V is σ-Dedekind complete, then $(I, \preceq)$ is σ-chain complete.*

Proof. For part (i), let $I = [a, b]$, V be as stated and let A be a subset of I. On one hand, if A is nonempty, then, by Dedekind completeness, $s := \bigvee A$ exists in V. By Exercise A.1.14, s is the supremum of A in $(I, \preceq)$. A similar argument shows that A has an infimum in $(I, \preceq)$. On the other hand, if $A = \varnothing$, then a is an upper bound of $\varnothing$ (vacuously – see Exercise A.1.13) and $a \preceq v$ for every upper bound v of $\varnothing$ in I (in fact for every $v \in I$). Hence a is the supremum of A in $(I, \preceq)$. A similar argument shows that b is the infimum of $\emptyset$ in $(I, \preceq)$.

The proof of part (ii) is very similar to the proof of part (i). □

A.5.1.3 Order Continuity

We call a map S from poset V to poset W **order continuous** on V if

$$Sv_n \uparrow Sv \quad \text{whenever } v_n \uparrow v.$$

In other words, if $(v_n) \subset V$ with $v_n \uparrow v \in V$, then $\bigvee_n Sv_n$ exists in W and equals Sv.

Example A.5.4. Fix $V \subset \mathbb{R}^k$ and let F be a self-map on V. If F is order preserving and continuous with respect to the ordinary Euclidean topology (also called norm continuous), then F is order continuous on V. This holds because convergence of a sequence (v_n) to v in the Euclidean topology is equivalent to pointwise convergence, which, for an increasing sequence, is equivalent to $v_n \uparrow v$ in V (Lemma A.1.4). Thus, since F is order preserving and norm continuous, $v_n \uparrow v$ implies $Fv_n \uparrow Fv$.

Remark A.5.1. The definition of order continuity varies across subfields of mathematics. The notion we use here is relatively weak but all we will need for our analysis. (In some sources, what we call order continuity is referred to as σ-order continuity, or countable order continuity. Since we use no other notions of order continuity, we maintain the simpler name.)

In the next lemma, V and W are arbitrary posets.

Lemma A.5.5. *Every order continuous map from V to W is order preserving.*

Proof. Let S be an order continuous map from V to W. Fix $v, v' \in V$ with $v \preceq v'$. Let (v_n) be such that $v_1 = v$ and $v_n = v'$ for all $n > 1$. Evidently $\bigvee_n v_n = v'$. Since S is order continuous, the supremum $\bigvee Sv_n$ exists in W and equals Sv'. This tells us that the supremum of $\{Sv, Sv'\}$ is Sv'. Hence $Sv \preceq Sv'$ and S is order preserving. □

Next we state a variation on the Tarski–Kantorovich fixed point theorem.

Theorem A.5.6 (Tarski–Kantorovich I)**.** *Let S be order continuous self-map on V. If V is σ-Dedekind complete and there exist elements $v_a \preceq v_b$ in V with $v_a \preceq Sv_a$ and $Sv_b \preceq v_b$, then there exists a $\bar{v} \in V$ such that $S^n v_a \uparrow \bar{v}$ and, moreover, $S\bar{v} = \bar{v}$.*

Proof. Let S, V be as stated. Fix $v_a \preceq v_b$ in V with $v_a \preceq Sv_a$ and $Sv_b \preceq v_b$. The map S is order continuous and hence order preserving, so the sequence $(v_n) := (S^n v_a)$ is increasing. As the set V is σ-Dedekind complete and the sequence is bounded above by v_b, the suprema $\bigvee_{n \geqslant 1} v_n$ and $\bigvee_{n \geqslant 1} Sv_n$ exist in V. If $\bar{v} := \bigvee_n v_n$, then, by order continuity, $S\bar{v} = S \bigvee_{n \geqslant 1} v_n = \bigvee_{n \geqslant 1} Sv_n = \bigvee_{n \geqslant 2} v_n = \bar{v}$. Hence $S\bar{v} = \bar{v}$. We have also shown that $S^n v_a \uparrow \bar{v}$. □

Here's a more standard version of the Tarski–Kantorovich fixed point theorem.

Theorem A.5.7 (Tarski–Kantorovich II). *If S is an order continuous self-map on V and V is σ-chain complete, then S has a fixed point in V.*

Proof. By σ-chain completeness, V has a least element v_a and greatest element v_b. For these elements we have $v_a \preceq Sv_a$ and $Sv_b \preceq v_b$. An essentially identical argument to the one in the proof of Theorem A.5.6 shows that $(S^n v_a)$ has a supremum and that supremum is a fixed point of S. □

The next lemma is analogous to Lemma A.5.3, which studied order stability in the chain complete case.

Lemma A.5.8. *Let S be a self-map on σ-chain complete poset V. If S has at most one fixed point in V and S is order continuous, then S is order stable on V.*

Proof. The proof is almost identical to that of Lemma A.5.3, with the only significant difference being that the Tarski–Kantorovich theorem (page 205) is used to obtain the fixed point. □

A.5.2 Ordered Vector Space

Next we add algebraic structure to posets. The combination of algebraic operations and order will allow us to develop sharp sufficient conditions for dynamic programs and convergence of algorithms.

A.5.2.1 Definition and Properties

Let $E = (E, +, \cdot)$ be a vector space with origin 0 (see §A.4.1) and let $\leqslant$ be a partial order on E. We call $(E, \leqslant)$ an **ordered vector space** if the order is preserved under addition and nonnegative scalar multiplication; that is, if

(i) $u \leqslant v$ implies $u + b \leqslant v + b$ for any $b \in E$, and

(ii) $u \leqslant v$ and $\alpha \in \mathbb{R}$ with $0 \leqslant \alpha$ implies $\alpha u \leqslant \alpha v$.

The **positive cone** of E, typically denoted by E_+, is all $v \in E$ with $0 \leqslant v$.

Example A.5.5. $\mathbb{R}^n$ is an ordered vector space under the pointwise order $\leqslant$, with positive cone equal to the set of nonnegative vectors in $\mathbb{R}^n$.

EXERCISE A.5.3. Let $\mathcal{S}$ be the vector space of all symmetric $n \times n$ matrices (with addition and scalar multiplication defined in the obvious way) and let $\mathcal{N}$ be the negative semidefinite matrices in $\mathcal{S}$. As in §2.3.3.3, we impose the Loewner partial order, writing $A \preccurlyeq B$ when $A - B \in \mathcal{N}$. Show that $(\mathcal{S}, \preccurlyeq)$ is an ordered vector space.

EXERCISE A.5.4. Let X be any nonempty set and let $\mathbb{R}^{\mathsf{X}}$ be the vector space of real-valued functions on X. Let $\leqslant$ be the pointwise partial order. Show that $(\mathbb{R}^{\mathsf{X}}, \leqslant)$ is an ordered vector space.

EXERCISE A.5.5. Let $(E, \leqslant)$ be an ordered vector space and fix $u, v, w \in E$. Prove that

(i) $u \leqslant 0$ and $v \leqslant 0$ implies $u + v \leqslant 0$,

(ii) $u \leqslant v$ implies $-v \leqslant -u$,

(iii) $(u \vee v) + w = (u + w) \vee (v + w)$, and

(iv) $\alpha(u \vee v) = (\alpha u) \vee (\alpha v)$ whenever $\alpha \geqslant 0$.

Using the definition in §A.1.2.4, if (v_n) is a sequence in ordered vector space E and $v \in E$, then the statement $v_n \uparrow v$ means that (v_n) is increasing and $\bigvee_n v_n = v$.

EXERCISE A.5.6. Prove that

(i) if $u_n \uparrow 0$ and $v_n \uparrow 0$, then $u_n + v_n \uparrow 0$, and

(ii) if $u_n \uparrow u$ and $b \in E$, then $u_n + b \uparrow u + b$.

Lemma A.5.9. *Let (u_n) and (v_n) be sequences in ordered vector space E and let α, β be nonnegative constants. The following implications hold:*

(i) *If $u_n \uparrow u$ and $v_n \uparrow v$, then $\alpha u_n + \beta v_n \uparrow \alpha u + \beta v$.*

(ii) *If $u_n \uparrow u$, then $-u_n \downarrow -u$.*

Proof. These claims follow from Theorem 10.2 of Zaanen (2012). They can also be obtained by applying and extending the results in Exercise A.5.6. □

In some settings, a partial order is introduced into a vector space E by first choosing a (pointed convex) cone C on E (see §A.4.1.2) and stating that $u \leqslant v$ if and only if $v - u \in C$. The following discussion clarifies this idea.

EXERCISE A.5.7. With $\leqslant$ defined as above, show that $(E, \leqslant)$ is an ordered vector space and that C is the positive cone of $(E, \leqslant)$.

EXERCISE A.5.8. Continuing the previous exercise, show that if E is a normed linear space and C is closed in E, then $\leqslant$ is a closed partial order (see page 214).

EXERCISE A.5.9. Show conversely that if $(E, \leqslant)$ is an ordered vector space, then the positive cone in E is a (pointed convex) cone.

A.5.2.2 Operators on Ordered Vector Space

A linear operator T mapping ordered vector space E to itself is called **positive** if T is invariant on the positive cone; that is, if $u \in E$ and $u \geqslant 0$ implies $Tu \geqslant 0$.

Example A.5.6. Let $\leqslant$ be the pointwise order on $\mathbb{R}^n$ and let A be an $n \times n$ matrix. We identify A with the linear operator $\mathbb{R}^n \ni x \mapsto Ax \in \mathbb{R}^n$. This operator is positive if and only if all elements of A are nonnegative.

(In the canonical example given above, *positive* operators are identified with *nonnegative* matrices. Unfortunately, this notational inconsistency is deeply embedded in the existing literature so we must accept it.)

EXERCISE A.5.10. Prove that a linear operator mapping E to itself is positive if and only if it is order preserving.

Let E be an ordered vector space and let $A \colon E \to E$ be a linear operator. Recalling the definition in §A.5.1.3, A is order continuous on E when $(v_n) \subset E$ and $v_n \uparrow v \in E$ implies $Av_n \uparrow Av$. By Lemma A.5.5, every order continuous linear operator is order preserving – and hence positive. The next exercise can be completed using Lemma A.5.9.

EXERCISE A.5.11. Show that, for a positive linear operator $A \colon E \to E$, the following statements are equivalent:

(i) A is order continuous on E.

(ii) $Av_n \downarrow Av$ whenever $(v_n) \subset E$ and $v_n \downarrow v \in E$.

(iii) $Av_n \downarrow 0$ whenever $(v_n) \subset E$ and $v_n \downarrow 0$.

A self-map S on a convex subset C of ordered vector space $E := (E, \leqslant)$ is called **convex** on C if

$$S(\lambda v + (1-\lambda)v') \leqslant \lambda Sv + (1-\lambda)Sv' \text{ whenever } v \leqslant v' \in C \text{ and } 0 \leqslant \lambda \leqslant 1$$

The map S is called **concave** on C if

$$\lambda Sv + (1-\lambda)Sv' \leqslant S(\lambda v + (1-\lambda)v') \text{ whenever } v \leqslant v' \in C \text{ and } 0 \leqslant \lambda \leqslant 1$$

EXERCISE A.5.12. Let $(\mathsf{X}, \mathscr{A})$ be a measurability space and let $C = [0, b]$ be an order interval of $bm\mathsf{X}$. Suppose that $Sv = u + f(Kv)$ maps C to itself, where $u \in bm\mathsf{X}$, K is a linear operator on $bm\mathsf{X}$ and $f \colon \mathbb{R}_+ \to \mathbb{R}_+$ is concave. The function f is applied pointwise, so that $f(Kv)(x) = f((Kv)(x))$. Prove that S is concave on C.

A.5.2.3 Riesz Space

Next we introduce Riesz spaces, which are ordered vector spaces with lattice structure. This structure allows for the introduction of a notion of absolute value, which behaves similarly to the pointwise absolute value over vectors in $\mathbb{R}^n$. Absolute value in turn helps us clarify and quantify the actions of operators, providing new opportunities for establishing optimality conditions in dynamic programs.

An ordered vector space E is called a **Riesz space** if E is a lattice. With $\vee$ and $\wedge$ as the lattice operations and $u, v, w \in E$, the following properties always hold:

(i) $u \wedge v = -((-u) \vee (-v))$ and $u \vee v = -((-u) \wedge (-v))$.

(ii) $(u \wedge v) + w = (u + w) \wedge (v + w)$ and $(u \vee v) + w = (u + w) \vee (v + w)$.

These facts can be easily verified and other related results are found in Chapter 2 of Zaanen (2012).

EXERCISE A.5.13. Show that if E is an ordered vector space and E is closed under $\vee$ (i.e., $u, v \in E$ implies $u \vee v \in E$), then E is a Riesz space.

Example A.5.7. If X is any nonempty set and $\leqslant$ is the pointwise order, then $(\mathbb{R}^{\mathsf{X}}, \leqslant)$ is a Riesz space. Indeed, $(\mathbb{R}^{\mathsf{X}}, \leqslant)$ is an ordered vector space (Exercise A.5.4) and, given $f, g \in \mathbb{R}^{\mathsf{X}}$, the pointwise maximum $x \mapsto \max\{f(x), g(x)\}$ is the supremum $f \vee g$ of $\{f, g\}$. This can be checked directly or by referring back to Exercise A.1.11.

Lemma A.5.10. *Let V be a linear subspace of a Riesz space $E = (E, \leqslant)$. If V is a sublattice of E, then $(V, \leqslant)$ is a Riesz space.*

Proof. It is clear that any linear subspace of an ordered vector space is again an ordered vector space. Moreover, since V is a sublattice of $(E, \leqslant)$, it follows that $(V, \leqslant)$ is itself a lattice. Thus, $V = (V, \leqslant)$ is a Riesz space. □

Example A.5.8. If X is a metric space, then $bc\mathsf{X} = (bc\mathsf{X}, \leqslant)$ is a sublattice of the Riesz space $\mathbb{R}^{\mathsf{X}}$ (Exercise A.5.2) and a linear subspace. Hence $bc\mathsf{X}$ is a Riesz space.

Exercise A.5.14. Provide a counterexample to the claim that every linear subspace L of a Riesz space $(E, \leqslant)$ is a Riesz space under $\leqslant$.

For element u of any Riesz space $(E, \leqslant)$ we use the notation

$$|u| := u \vee (-u), \quad u^+ := u \vee 0 \quad \text{and} \quad u^- := (-u) \vee 0.$$

These points in E are called the **absolute value**, **positive part**, and **negative part** of u respectively. One easily shows that $|-u| = |u|$. Also,

Lemma A.5.11. *For any $u, v \in E$ we have*

(i) $u = u^+ - u^-$ *and* $|u| = u^+ + u^-$,

(ii) $|u| = 0$ *if and only if* $u = 0$,

(iii) $|u| + |v| = |u + v| \vee |u - v|$, *and*

(iv) *if* $u \in E_+$, *then* $|v| \leqslant u \iff v \in [-u, u]$.

Proof. For (i) we use Exercise A.5.5 to obtain $u^+ - u = (u \vee 0) - u = 0 \vee (-u) = u^-$, which gives the first equality. For the second we observe that $u \vee (-u) = u \vee (-u) + u - u = (2u \vee 0) - u = 2u^+ - u = 2u^+ - u^+ + u^- = u^+ + u^-$. For (iii) we refer to Theorem 5.3 of Zaanen (2012). Regarding (iv), we have

$$|v| \leqslant u \iff v \vee (-v) \leqslant u \iff v \leqslant u \text{ and } -v \leqslant u \iff -u \leqslant v \leqslant u. \qquad \square$$

Notice that (iii) implies the triangle inequality $|u + v| \leqslant |u| + |v|$.

Exercise A.5.15. Prove that if K is a positive linear operator on Riesz space E, then $|Ku| \leqslant K|u|$ for all $u \in E$.

Exercise A.5.16. Let S be an order preserving self-map on a subset V of a Riesz space E. Suppose there exists a order continuous linear operator $K \colon E \to E$ such that $|Sv - Sw| \leqslant K|v - w|$ for all $v, w \in V$. Prove that S is order continuous on V.

A.5.2.4 Riesz Spaces of Measurable Functions

Let $(\mathsf{X}, \mathscr{A}, \mu)$ be a σ-finite measure space. Fix $p \in [1, \infty)$. As usual, we let

- $m\mathsf{X}$ be the real-valued Borel measurable functions on $(\mathsf{X}, \mathscr{A})$,
- $bm\mathsf{X}$ be the bounded functions in $m\mathsf{X}$, and
- $\mathcal{L}_p := \mathcal{L}_p(\mathsf{X}, \mathscr{A}, \mu)$ be all $f \in m\mathsf{X}$ with $\int |f|^p \, \mathrm{d}\mu < \infty$.

The vector spaces $m\mathsf{X}$, $bm\mathsf{X}$ and $\mathcal{L}_p$ are all Riesz spaces when paired with the pointwise partial order $\leqslant$. Both $bm\mathsf{X}$ and $\mathcal{L}_p$ are subsets of $m\mathsf{X}$ and, in addition, $bm\mathsf{X} \subset \mathcal{L}_p$ whenever μ is finite.

Exercise A.5.17. Show that $m\mathsf{X}$ is a Riesz space. Next, show that $bm\mathsf{X}$ and $\mathcal{L}_p$ are Riesz spaces using Lemma A.5.10.

Exercise A.5.18. Let V be a linear subspace of $m\mathsf{X}$ and let K be the **integral operator** defined by

$$(Kv)(x) = \int v(x')N(x, \mathrm{d}x') \qquad (v \in V, \; x \in \mathsf{X}),$$

where N is a transition kernel (see §A.5.4.1) on $(\mathsf{X}, \mathscr{A})$ and we assume that $KV \subset V$. Show that K is a positive linear operator on V. In addition, prove the following: If V is closed under pointwise suprema (see §A.1.2.6), then K is order continuous on V.

A.5.2.5 Almost Everywhere Pointwise Order

Let $(\mathsf{X}, \mathscr{A}, \mu)$ be as in the previous section and fix $p \in [1, \infty)$. Let $L_p := L_p(\mathsf{X}, \mathscr{A}, \mu)$ be the Banach space of equivalence classes defined in §A.4.2.5. Let $\leqslant$ be defined by $f \leqslant g$ if and only if $\{x \in \mathsf{X} : f(x) > g(x)\}$ has μ-measure zero.

EXERCISE A.5.19. The relation $\leqslant$ introduced above can be stated more formally as follows: given equivalence classes f, g in L_p we write $f \leqslant g$ if, for any functions $f_0 \in f$ and $g_0 \in g$, the set $\{x \in \mathsf{X} : f_0(x) > g_0(x)\}$ has μ-measure zero. Using this definition, show that L_p is partially ordered under $\leqslant$.

The space $(L_p, \leqslant)$ just described is a Riesz space. For example, if $f, g \in L_p$, then $|f \vee g| \leqslant |f| + |g|$, and $\int |f| \, \mathrm{d}\mu$ and $\int |g| \, \mathrm{d}\mu$ are both finite. Hence $f \vee g \in L_p$.

A.5.2.6 Archimedean Riesz Space

A Riesz space E is called **Archimedean**, if, for each $u \in E_+$, we have $u/n \downarrow 0$. All Riesz spaces encountered in this book are Archimedean.

Example A.5.9. Let L_p be as defined in §A.5.2.5 and fix $u \in L_p$ with $u \geqslant 0$. Clearly 0 is a lower bound of $(u_n) := (u/n)_{n \in \mathbb{N}}$. If w is another lower bound, then

$$\int_B w \, \mathrm{d}\mu \leqslant \frac{1}{n} \int_B u \, \mathrm{d}\mu \quad \text{for all } n \in \mathbb{N} \text{ and } B \in \mathscr{A},$$

so $w \leqslant 0$. Hence 0 is a greatest upper bound and $u_n \downarrow 0$. In particular, L_p is Archimedean.

A.5.2.7 Dedekind completeness of Riesz Space

Since each Riesz space is a partially ordered space, the notions of Dedekind and σ-Dedekind completeness apply directly. Moreover, when testing these forms of completeness, one-sided conditions suffice. The following one-sided condition is particularly simple.

Lemma A.5.12. *A Riesz space E is σ-Dedekind complete if and only if every bounded increasing sequence in E has a supremum.*

For a proof of Lemma A.5.12, see Theorem 12.1 of Zaanen (2012).

We will make repeated use of the following fact.

Lemma A.5.13. *The Riesz space $L_p(\mathsf{X}, \mathscr{A}, \mu)$ from §A.5.2.5 is Dedekind complete.*

A proof of Lemma A.5.13 can be found in Example 12.5 of Zaanen (2012).

Several interesting function spaces are naturally ordered by the pointwise partial order. Next we study the completeness properties of such Riesz spaces. We will make use of the following lemma, in the statement of which, for $(v_n) \subset \mathbb{R}^{\mathsf{X}}$, the symbol $\sup_n v_n$ indicates the pointwise supremum.

Lemma A.5.14. *Let* X *be any nonempty set and let* $\leqslant$ *be the pointwise partial order on* $\mathbb{R}^{\mathsf{X}}$. *Let* $E = (E, \leqslant)$ *be a Riesz space contained in* $\mathbb{R}^{\mathsf{X}}$. *If* $\sup_n v_n$ *is in* E *whenever* $(v_n) \subset E$ *is increasing and bounded above, then* E *is* σ*-Dedekind complete.*

Proof. Let E be as above and let (v_n) be a sequence in V that is increasing and bounded above. By assumption, $s := \sup_n v$ exists in E. By Exercise A.1.11 we have $\bigvee_n v_n = s$. In view of Lemma A.5.12, the space E is σ-Dedekind complete. □

Now let $(\mathsf{X}, \mathscr{A}, \mu)$ be a σ-finite measure space and let $m\mathsf{X}$, $bm\mathsf{X}$ and $\mathcal{L}_p := \mathcal{L}_p(\mathsf{X}, \mathscr{A}, \mu)$ be the Riesz spaces discussed in §A.5.2.4. As above, let $\leqslant$ be the pointwise partial order.

Corollary A.5.15. *The spaces* $m\mathsf{X}$, $bm\mathsf{X}$ *and* $\mathcal{L}_p$ *are all* σ*-Dedekind complete under* $\leqslant$.

Proof. Consider the poset $m\mathsf{X}$. Let (v_n) be increasing and bounded above in $m\mathsf{X}$. In view of Lemma A.5.14 we need only show that $s := \sup_n v_n$ is in $m\mathsf{X}$. This follows from existence of suprema in $\mathbb{R}$ when subsets are bounded above (so s is real-valued) and Lemma A.3.1, which implies measurability.

Next consider the poset $bm\mathsf{X}$. Let $(v_n) \subset bm\mathsf{X}$ be increasing and bounded above by $w \in bm\mathsf{X}$. By the same argument as the last paragraph, we have $s := \sup_n v_n \in m\mathsf{X}$. Moreover, $v_1 \leqslant s \leqslant w$ with $v_1, w \in bm\mathsf{X}$. Hence $s \in bm\mathsf{X}$. The claim now follows from Lemma A.5.14.

Finally, suppose $V = \mathcal{L}_p$ for some $p \geqslant 1$. Let $(v_n) \subset \mathcal{L}_p$ be increasing and bounded above by $w \in \mathcal{L}_p$. As before, the pointwise supremum s, is measurable, so we need only show that $\int |s|^p \, \mathrm{d}\mu < \infty$. By assumption, $v_1 \leqslant s \leqslant w$ with $v_1, w \in \mathcal{L}_p$. But then $s \in \mathcal{L}_p$, as a result of $|s| \leqslant |v_1| + |w|$ and the fact that $\mathcal{L}_p$ is closed under addition (by Minkowski's inequality). □

A.5.3 Topology and Order

In some applications it will be helpful to draw on results that use topological or metric structure. In this section we note some elementary facts about topological, metric and normed spaces where order is also present.

A.5.3.1 Partially Ordered Space

A partial order $\preceq$ on topological space V is called **closed** if, given any two nets $(u_\alpha)_{\alpha\in\Lambda}$ and $(v_\alpha)_{\alpha\in\Lambda}$ contained in V,

$$u_\alpha \to u, \;\; v_\alpha \to v \;\text{ and }\; u_\alpha \preceq v_\alpha \text{ for all } \alpha \in \Lambda \quad \implies \quad u \preceq v. \tag{A.21}$$

A **partially ordered space**, also called a **pospace**, is a Hausdorff topological space endowed with a closed partial order. (We make the Hausdorff assumption so that sequences have unique limits.)

Example A.5.10. One canonical pospace is $\mathbb{R}^n$ paired with the the product topology and the pointwise partial order $\leqslant$. The pointwise order $\leqslant$ is closed because convergence in the product topology implies pointwise convergence (§A.2.1.6), and limits in $\mathbb{R}$ preserve the usual real-valued order $\leqslant$.

The next lemma connects topological and order convergence in partially ordered space $V = (V, \preceq)$. In the statement, $(v_\alpha)_{\alpha\in\Lambda}$ is a net in V.

Lemma A.5.16. *If $v \in V$ with $v_\alpha \to v$ and $v_\alpha \preceq v$ for all $\alpha \in \Lambda$, then $\bigvee_\alpha v_\alpha = v$.*

Proof. Let (v_α) and v be as stated. By assumption, v is an upper bound of (v_α). If w is any other upper bound, then $v_\alpha \preceq w$ for all α. Since $\preceq$ is closed and $v_\alpha \to v$, this implies $v \preceq w$. Hence v is the least upper bound of (v_α). □

The next lemma shows how global stability (see §A.2.2.1) interacts with order stability in the setting of partially ordered space.

Lemma A.5.17. *Let V be a partially ordered space and let S be an order preserving self-map on V. If S is globally stable on V, then S is strongly order stable on V.*

Proof. Let S, V have the stated properties and let $\bar v$ be the unique fixed point of S in V. If $v \in V$ and $v \preceq S\,v$, then, iterating on this inequality and using the fact that S is order preserving, we have $v \preceq S^n v$ for all $n \in \mathbb{N}$. Since the partial order is closed and S is globally stable, taking the limit gives $v \preceq \bar v$. Using this inequality and $v \preceq S^n v$, we have $v \preceq S^n v \preceq S^n \bar v = \bar v$ for all n. Since $S^n v \to \bar v$, Lemma A.5.16 implies that $S^n v \uparrow \bar v$. Hence strong upward stability holds. The proof of strong downward stability is similar. □

The following result can be used to compare fixed points of operators. In the statement, $V = (V, \preceq)$ is a pospace and $\mathcal{S}(V)$ is all self-maps on V, ordered pointwise (i.e., for $S, T \in \mathcal{S}(V)$, we have $S \preceq T$ if and only if $Sv \preceq Tv$ for all $v \in V$).

Proposition A.5.18. *Fix S, T in $\mathcal{S}(V)$. If $S \preceq T$ and, in addition, T is order preserving and globally stable on V, then its unique fixed point dominates any fixed point of S.*

Exercise A.5.20. Prove Proposition A.5.18.

A.5.3.2 Sup-Nonexpansive Metrics

Let $V = (V, \preceq)$ be an partially ordered set and let d be a metric on V. We call d **sup-nonexpansive** if

$$d\left(\vee_\alpha v_\alpha, \vee_\alpha w_\alpha\right) \leqslant \sup_\alpha d(v_\alpha, w_\alpha) \tag{A.22}$$

for any subsets (v_α) and (w_α) of V such that the suprema $\vee_\alpha v_\alpha$ and $\vee_\alpha w_\alpha$ exist.

Example A.5.11. If $V = b\mathsf{X}$ for some set X and d is the supremum metric, then d is sup-nonexpansive. Indeed, given subsets (v_α) and (w_α) of V such that $\vee_\alpha v_\alpha$ and $\vee_\alpha w_\alpha$ exist, as well as fixed $x \in \mathsf{X}$, we have

$$\left|\left(\vee_\alpha v_\alpha\right)(x) - \left(\vee_\alpha w_\alpha\right)(x)\right| = \left|\sup_\alpha v_\alpha(x) - \sup_\alpha w_\alpha(x)\right| \leqslant \sup_\alpha \left|v_\alpha(x) - w_\alpha(x)\right|,$$

where the last inequality is by Lemma A.1.2. From this we obtain

$$\left|\left(\vee_\alpha v_\alpha\right)(x) - \left(\vee_\alpha w_\alpha\right)(x)\right| \leqslant \sup_\alpha d(v_\alpha, w_\alpha) \quad \text{for all } x \in \mathsf{X}.$$

Taking the supremum over x completes the proof.

Sup-nonexpansive metrics will be useful for us because contraction properties are passed from collections of mappings to their upper envelopes. The next lemma explains. In the statement, $\mathbb{T} := \{T_\sigma : \sigma \in \Sigma\}$ is a collection of self-maps on V, d is a metric on V, and $Tv := \vee_\sigma T_\sigma v$ at each $v \in V$.

Lemma A.5.19. *If Tv is well-defined at all $v \in V$, d is sup-nonexpansive, and each $T_\sigma \in \mathbb{T}$ is a contraction of modulus β on V, then T is a contraction of modulus β on V.*

Proof. For any $v, w \in V$, we have

$$d(Tv, Tw) = d\left(\vee_\sigma T_\sigma v, \vee_\sigma T_\sigma w\right) \leqslant \vee_\sigma d(T_\sigma v, T_\sigma w) \leqslant \beta d(v, w)$$

In particular, T is a contraction of modulus β on V. □

A.5.3.3 Banach Lattices

Let $E = (E, \leqslant)$ be a Riesz space and let $\| \cdot \|$ be a complete norm on E, so that $(E, \| \cdot \|)$ is a Banach space. If the norm is compatible with the order structure on E, in the sense that $\|u\| \leqslant \|v\|$ whenever $|u| \leqslant |v|$, then $\|\cdot\|$ is called a **lattice norm** and $E := (E, \leqslant, \|\cdot\|)$ is called a **Banach lattice**.

Example A.5.12. We can easily verify that if X is any metric space, $\| \cdot \|$ is the supremum norm on $bc\mathsf{X}$, and $\leqslant$ is the pointwise partial order, then $bc\mathsf{X} = (bc\mathsf{X}, \leqslant, \| \cdot \|)$ is a Banach lattice. We already showed that $(bc\mathsf{X}, \leqslant)$ is a Riesz space (Example A.5.8). In addition, $(bc\mathsf{X}, \| \cdot \|)$ is a Banach space (Example A.4.16). Finally, the norm is clearly a lattice norm. Hence $bc\mathsf{X}$ is a Banach lattice.

Example A.5.13. Let $(L_p, \leqslant)$ be the Riesz space defined in §A.5.2.5, where $f \leqslant g$ if $\mu\{f > g\} = 0$. Paired with this partial order, the Banach space L_p becomes a Banach lattice.

If (v_n) is a sequence in E then convergence is as defined for sequences in normed linear space (see §A.4.2.1): $v_n \to v$ means that $\|v_n - v\| \to 0$ as $n \to \infty$. This should not be confused with $v_n \uparrow v$ and $v_n \downarrow v$, which are defined in terms of suprema and infima (see §A.5.1.3). Some relationships between the different forms of convergence are discussed in the next theorem.

Theorem A.5.20. *If $E = (E, \leqslant, \| \cdot \|)$ is a Banach lattice, then*

(i) *$\leqslant$ is a closed partial order on E under the norm topology,*

(ii) *if $(v_n) \subset V$ is increasing and $v_n \to v$, then $v_n \uparrow v$, and*

(iii) *if $(v_n) \subset V$ is decreasing and $v_n \to v$, then $v_n \downarrow v$.*

Proof. A proof of part (i) can be found in 18.4 of Zaanen (2012). Regarding (ii), suppose first that v_n is increasing and $v_n \to v$. Fix $m \in \mathbb{N}$. By part (i) and $v_m \leqslant v_n$ for all $n \geqslant m$ we have $v_m \leqslant v$ for all m. Hence v is an upper bound of (v_n). If w is another upper bound of (v_n), then $v_n \leqslant w$ for all n and hence, using part (i) once more, $v \leqslant w$. Hence v is a least upper bound of (v_n). The proof of (iii) is similar. □

A.5.3.4 Weighted Sup-Norm Function Spaces

In this section we introduce a class of Banach lattices that are useful for handing unbounded dynamic programming problems. To this end, let X be a topological space.

A **weight function** on X is a mapping $\ell \in m\mathsf{X}$ with $\ell(x) \geqslant 1$ for all $x \in \mathsf{X}$. Given a weight function ℓ and $v \in \mathbb{R}^{\mathsf{X}}$ we introduce the **ℓ-weighted supremum norm**

$$\|v\| := \sup_{x \in \mathsf{X}} \frac{|v(x)|}{\ell(x)}.$$

In this setting, we let

- $b_\ell \mathsf{X}$ be all $v \in \mathbb{R}^{\mathsf{X}}$ such that $\|v\| < \infty$,
- $b_\ell m\mathsf{X}$ be all Borel measurable functions in $b_\ell \mathsf{X}$, and
- $b_\ell c\mathsf{X}$ be all continuous functions in $b_\ell \mathsf{X}$.

Elements of $b_\ell \mathsf{X}$ are called **ℓ-bounded** functions.

Exercise A.5.21. Show that $b_\ell \mathsf{X}$, $b_\ell m\mathsf{X}$ and $b_\ell c\mathsf{X}$ are all linear subspaces of $\mathbb{R}^{\mathsf{X}}$ under the usual pointwise notions of addition and scalar multiplication of functions.

Exercise A.5.22. Show that $bm\mathsf{X} \subset b_\ell m\mathsf{X}$ and $bc\mathsf{X} \subset b_\ell c\mathsf{X}$.

Exercise A.5.23. Show that $\|\cdot\|$ is a norm on $b_\ell \mathsf{X}$.

Exercise A.5.24. Prove that convergence in $b_\ell \mathsf{X}$ implies pointwise convergence; that is, if (w_n) is a sequence in $b_\ell \mathsf{X}$ and $\|w_n - w\| \to 0$ for some $w \in b_\ell \mathsf{X}$, then $w_n(x) \to w(x)$ for every $x \in \mathsf{X}$.

The next theorem gives conditions under which the spaces discussed above are Banach lattices. Proofs can be found in §12.2.1 of Stachurski (2022).

Theorem A.5.21. *Both $b_\ell \mathsf{X}$ and $b_\ell m\mathsf{X}$ are Banach lattices under the norm $\|\cdot\|$ and the usual pointwise order. If ℓ is a continuous function, then $b_\ell c\mathsf{X}$ is also a Banach lattice.*

A.5.3.5 Positive Operators on Banach Lattices

If E is a Banach lattice, then, as in §A.4.3, we take $\mathcal{B}(E)$ to be the norm bounded (and hence norm continuous) linear self-maps on E. Let $\mathcal{B}_+(E)$ be the positive linear self-maps on E.

Theorem A.5.22. *If E is a Banach lattice, then $\mathcal{B}_+(E) \subset \mathcal{B}(E)$.*

Proof. See Theorem 15.1 of Zaanen (2012). □

EXERCISE A.5.25. Given $A \in \mathcal{B}_+(E)$, prove that

$$\|A\| = \sup\{\|Au\| : u \in E_+ \text{ and } \|u\| = 1\}.$$

Continuing in the setting of Exercise A.5.25, the pointwise partial order on $\mathcal{B}(E)$ is defined by $A \leqslant B$ whenever $Av \leqslant Bv$ for all $v \in E$. The set $\mathcal{B}_+(E)$ coincides with the positive cone of $\mathcal{B}(E)$. On this positive cone, the spectral radius is order preserving:

Theorem A.5.23. *If $A, B \in \mathcal{B}_+(E)$ and $A \leqslant B$, then $\|A\| \leqslant \|B\|$ and $\rho(A) \leqslant \rho(B)$.*

Proof. Fix $0 \leqslant A \leqslant B$. If $u \in E_+$ and $\|u\| = 1$, then $0 \leqslant Au \leqslant Bu$, so $\|Au\| \leqslant \|Bu\|$. The bound $\|A\| \leqslant \|B\|$ now follows from Exercise A.5.25.

Regarding the second claim, we use $0 \leqslant A \leqslant B$ and induction we obtain $0 \leqslant A^k \leqslant B^k$ for all $k \in \mathbb{N}$. Hence, by the first claim, $\|A^k\| \leqslant \|B^k\|$ for all k. The inequality $\rho(A) \leqslant \rho(B)$ now follows from Gelfand's formula. □

A Banach lattice E is said to have a **σ-order continuous norm** if

$$(v_n) \subset E \text{ and } v_n \downarrow 0 \quad \implies \quad \|v_n\| \to 0.$$

Example A.5.14. The Banach lattice $L_p = (L_p(\mathsf{X}, \mathscr{A}, \mu), \leqslant)$ discussed in Example A.5.13 has order continuous norm. See, for example, Zaanen (2012), §17.

Theorem A.5.24. *Let E be Banach lattice. If E has σ-order continuous norm, then every positive linear operator from E to itself is order continuous.*

Proof. Let E be as stated and let A be a positive linear self-map on E. Fix $(v_n) \subset E$ with $v_n \downarrow 0$. Since A is a bounded linear operator (Theorem A.5.22) and E has σ-order continuous norm, we have $\|Av_n\| \leqslant \|A\|\|v_n\| \to 0$ in $\mathbb{R}$. Also, (Av_n) is decreasing because A is positive and (v_n) is decreasing. Applying Theorem A.5.20 yields $Av_n \downarrow 0$. By Exercise A.5.11, this convergence is sufficient for order continuity of A. □

In the next example, L_p is the Banach lattice discussed in Example A.5.13.

Corollary A.5.25. *If $p \in [1, \infty)$, then every positive linear operator from L_p to itself is order continuous.*

Proof. This follows from Theorem A.5.24 and Example A.5.14. □

A.5.4 Markov Models

Many dynamic programs have some form of Markov structure (or can be coerced into a Markov framework by suitably changing the state space). Here we review key ideas related to Markov processes and state some useful results.

In all of this section (§A.5.4), X is a metric space with Borel sets $\mathscr{B}$. The symbol $\mathscr{D}(\mathsf{X})$ is the set of all distributions (Borel probability measures) on X. If X is finite, then the metric on X is the discrete metric (under which all real-valued functions on X are continuous and $\mathscr{B}$ is the set of all subsets of X).

A.5.4.1 Stochastic Kernels

Let U be a second metric space. A **transition kernel** from U to X is a function N from $\mathsf{U} \times \mathscr{B}$ to $\mathbb{R}_+$ with the property that $u \mapsto N(u, B)$ is Borel measurable for each $B \in \mathscr{B}$ and $B \mapsto N(u, B)$ is a measure on $(\mathsf{X}, \mathscr{B})$ for all $u \in \mathsf{U}$. A **stochastic kernel** from U to X is a transition kernel P from U to X satisfying $P(u, \mathsf{X}) = 1$ for all $u \in \mathsf{U}$. Informally, the stochastic kernel P takes a point $u \in \mathsf{U}$ and randomly "transitions" to a new point in X via the distribution $P(u, \mathrm{d}x)$.

A common setting is where $\mathsf{U} = \mathsf{X}$. In this case we say that N is a **transition kernel on** X, while P is a **stochastic kernel on** X.

Example A.5.15. If X is finite and $p \colon \mathsf{X} \times \mathsf{X} \to [0, 1]$ obeys $\sum_{x' \in \mathsf{X}} p(x, x') = 1$ for all X, then P defined by

$$P(x, B) = \sum_{x' \in B} p(x, x') \qquad (x \in \mathsf{X},\ B \subset \mathsf{X})$$

is a stochastic kernel on X.

Example A.5.16. If μ is a σ-finite mesure on $(\mathsf{X}, \mathscr{B})$ and $p \colon \mathsf{X} \times \mathsf{X} \to \mathbb{R}_+$ is Borel measurable with $\int p(x, x') \, \mathrm{d}x' = 1$ for all X, then P defined by

$$P(x, B) = \int_B p(x, x') \, \mathrm{d}x' \qquad (x \in \mathsf{X},\ B \subset \mathsf{X})$$

is a stochastic kernel on X.

An X-valued stochastic process $(X_t)_{t=0}^{\infty}$ on $(\Omega, \mathscr{F}, \mathbb{P})$ is called **(P, ψ)-Markov** if $X_0 \stackrel{d}{=} \psi$ and $\mathbb{P}\{X_{t+1} \in B \mid X_t\} = P(X_t, \mathrm{d}x')$ with probability one for all $t \geqslant 0$. If $\psi = \delta_x$ for some $x \in \mathsf{X}$, then we say $(X_t)_{t \geqslant 0}$ is **(P, x)-Markov**. We also say that $(X_t)_{t \geqslant 0}$ is **P-Markov** if $(X_t)_{t \geqslant 0}$ is (P, ψ)-Markov for some $\psi \in \mathscr{D}(\mathsf{X})$.

Example A.5.17 (Stochastic recursive sequence)**.** Suppose $(X_t)_{t\geqslant 0}$ is defined by

$$X_{t+1} = F(X_t, W_{t+1}), \quad (W_t)_{t\geqslant 1} \overset{\text{IID}}{\sim} \varphi, \quad X_0 \sim \psi \tag{A.23}$$

where $(W_t)_{t\geqslant 1}$ and IID random elements taking values in metric space Z, $F\colon \mathsf{X}\times\mathsf{Z} \to \mathsf{X}$ is Borel measurable, and X_0 and $(W_t)_{t\geqslant 0}$ are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and are jointly independent. As a stochastic process, (X_t) is P-Markov when

$$P(x,B) := \mathbb{P}\{F(x, W_{t+1}) \in B\} = \int \mathbb{1}_B[F(x,z)]\varphi(dz) \qquad (x \in \mathsf{X},\ B \in \mathscr{B}). \tag{A.24}$$

A.5.4.2 Markov Operators

As before, let $m\mathsf{X}$ be all the measurable functions on X and let P be a stochastic kernel on X. Given $h \in m\mathsf{X}$ we set

$$(Ph)(x) := \int h(x')P(x, \mathrm{d}x') \qquad (x \in \mathsf{X}). \tag{A.25}$$

whenever the integral is well-defined. We call P the **Markov operator** generated by the stochastic kernel P. We use the same symbol because stochastic kernels and Markov operators can be placed in one-to-one correspondence via

$$P(x,B) = (P\mathbb{1}_B)(x) \qquad (x \in \mathsf{X},\ B \in \mathscr{B}). \tag{A.26}$$

(The stochastic kernel is on the left and the Markov operator is on the right.)

(We already studied a version of P in Exercise A.4.8 on page 192.) Intuitively, $(Ph)(x)$ represents the expectation of $h(X_{t+1})$ given $X_t = x$. We extend this interpretation below.

Example A.5.18. The Markov operator generated by the kernel P from (A.23) takes the form

$$(Ph)(x) = \int h[F(x,z)]\varphi(dz) \qquad (x \in \mathsf{X},\ h \in bm\mathsf{X}). \tag{A.27}$$

This expression is intuitive because $(Ph)(x)$ represents the expectation of $h(X_{t+1})$ given $X_t = x$, and $h(X_{t+1}) = h[F(X_t, W_{t+1})]$. Hence $(Ph)(x) = \mathbb{E}h[F(x, W_{t+1})]$, which is the right hand side of (A.27).

Lemma A.5.26. *Fix $(V, \leqslant) \subset (m\mathbb{R}^{\mathsf{X}}, \leqslant)$ and let P be a stochastic kernel on X. If $PV \subset V$, then P is order continuous on V.*

Proof. Let V and P be as stated. Let (v_n) be an increasing sequence in V with $v_n \uparrow v \in V$. By Lemma A.1.4, v_n converges pointwise to v. Fixing $x \in \mathsf{X}$ and applying the monotone convergence theorem (Theorem A.3.7), we have

$$\lim_n \int v_n(x')P(x, \mathrm{d}x') = \int v(x')P(x, \mathrm{d}x').$$

Since this holds for all x, another application of Lemma A.1.4 gives $Pv_n \uparrow Pv$. □

Given a stochastic kernel P on X, a distribution $\varphi \in \mathscr{D}$ is called **stationary** for P if

$$\varphi(B) = \int P(x, B)\varphi(\mathrm{d}x) \quad \text{for all } B \in \mathscr{B}.$$

Add at least one example.

A.5.4.3 General Properties

The following lemma lists useful properties of the Markov operator (A.25) when considered as a linear operator on $bm\mathsf{X}$, the set of bounded Borel measurable functions on X. Proofs can be found in Meyn and Tweedie (2009).

Add details on location of proofs.

Lemma A.5.27. *If P is a stochastic kernel on X and $h \mapsto Ph$ is the Markov operator defined by (A.25), then the following statements hold:*

(i) *P maps $bm\mathsf{X}$ to itself.*

(ii) *$P\mathbb{1}_{\mathsf{X}} = \mathbb{1}_{\mathsf{X}}$ pointwise on X.*

(iii) *P is order preserving on $bm\mathsf{X}$: if $h, g \in bm\mathsf{X}$ and $h \leqslant g$, then $Ph \leqslant Pg$*

(iv) *P is linear on $bm\mathsf{X}$.*

(v) *$\|P^t\| = 1$ for all $t \geqslant 0$, where $\| \cdot \|$ is the operator norm on $\mathcal{B}(bm\mathsf{X})$.*

(vi) *$\rho(P) = 1$, where ρ is spectral radius on $\mathcal{B}(bm\mathsf{X})$.*

(Note that (vi) follows from (v) and Gelfand's formula for the spectral radius.)

Here is a fundamental result linking the stochastic kernel P, Markov operator P, and any P-Markov process $(X_t)_{t \geqslant 0}$. For a proof see Meyn and Tweedie (2009).

Theorem A.5.28. *Let P be a stochastic kernel on X and let $(X_t)_{t \geqslant 0}$ be P-Markov. The corresponding Markov operator P on $bm\mathsf{X}$ obeys*

$$(P^t h)(x) = \mathbb{E}[h(X_t) \,|\, X_0 = x] \quad \textit{for all } t \geqslant 0 \textit{ and } h \in bm\mathsf{X}. \tag{A.28}$$

A.5.4.4 Markov Operators on Integrable Functions

Sometimes we wish to consider Markov operator a linear operators over a space of integrable functions. To this end, let P be a stochastic kernel on X and let φ be stationary for P. As before, we use the same symbol P for the Markov operator defined in (A.25). The space $L_1(\varphi) := L_1(\mathsf{X}, \mathscr{B}, \varphi)$ is the Banach lattice discussed in Example **??**.

Lemma A.5.29. *The following statements hold:*

(i) *P is an element of $\mathcal{B}(L_1(\varphi))$; that is, a bounded linear self-map from $L_1(\varphi)$ to itself.*

(ii) *P is order preserving on $L_1(\varphi)$*

(iii) *$\|P^t\| = 1$ for all $t \geqslant 0$, where $\| \cdot \|$ is the operator norm on $\mathcal{B}(L_1(\varphi))$.*

(iv) *$\rho(P) = 1$, where ρ is spectral radius on $\mathcal{B}(L_1(\varphi))$.*

Add proofs or references.

A.5.4.5 Feller Properties

Let P be a stochastic kernel from a metric space U to metric space X. We consider the mapping

$$(Ph)(u) := \int h(x')P(u, \mathrm{d}x'), \qquad (h \in bm\mathsf{X}).$$

For dynamic programming we will need conditions under which Ph is continuous. Here we investigate this topic.

To begin, we say that P is

- **weak Feller** if Ph is continuous on U whenever $h \in bc\mathsf{X}$ and
- **strong Feller** if Ph is continuous on U whenever $h \in bm\mathsf{X}$.

The weak Feller property is sometimes abbreviated to **Feller**. The Feller property is relatively easy to obtain. The next example shows a typical scenario.

Example A.5.19. In studying Markov decision processes we often work with dynamics of the form

$$X_{t+1} = F(X_t, A_t, W_{t+1}) \quad \text{with} \quad (W_t)_{t \geqslant 1} \overset{\text{IID}}{\sim} \varphi \in \mathscr{D}(\mathsf{W}),$$

where (X_t, A_t) takes values in $\mathsf{G} \subset \mathsf{X} \times \mathsf{A}$ and (X_t) takes values in X. Here X, A and W are metric spaces and $F \colon \mathsf{G} \times \mathsf{W} \to \mathsf{X}$ is Borel measurable. The corresponding stochastic kernel from G to X is given by

$$(Ph)(x, a) := \int h(x')P(x, a, \mathrm{d}x') = \int h(F(x, a, w))\varphi(\mathrm{d}w).$$

If $(x, a) \mapsto F(x, a, w)$ is continuous for all $w \in \mathsf{W}$, then P is weak Feller. Indeed, taking $(x_n, a_n) \to (x, a)$ in G and assuming $h \in bc\mathsf{X}$, the dominated convergence theorem yields

$$(Ph)(x_n, a_n) = \int h(F(x_n, a_n, w))\varphi(\mathrm{d}w) \to \int h(F(x, a, w))\varphi(\mathrm{d}w) = (Ph)(x, a)$$

as $n \to \infty$. In particular, Ph is continuous. Since Ph is clearly bounded, the kernel P is weak Feller. More generally, P is weak Feller whenever $(x, a) \mapsto F(x, a, w)$ is continuous for φ-almost all $w \in \mathsf{W}$.

The strong Feller property requires more conditions, since we need to map a potentially discountinuous function h into a continuous function Ph. For this we rely on smoothing properties of the integral. To obtain these properties we introduce a "dominating" measure μ on $(\mathsf{X}, \mathscr{B})$, which we assume to be σ-finite. A map p from $\mathsf{U} \times \mathsf{X}$ to $\mathbb{R}$ is called a **density kernel** from U to X with dominated measure μ if p is Borel measurable,

$$p \geqslant 0 \quad \text{and,} \quad \int p(u, x')\mu(\mathrm{d}x') = 1 \quad \text{for all } u \in \mathsf{U}.$$

We say that stochastic kernel P from U to X has density kernel p with dominating measure μ if p is a density kernel on X and

$$P(u, B) = \int_B p(u, x')\mu(\mathrm{d}x') \quad \text{for all } (u, B) \in \mathsf{U} \times \mathscr{B}.$$

If the dominating measure μ is not identified in the discussion below then we will be referring to Lebesgue measure, and we write $\mathrm{d}x$ instead of $\mu(\mathrm{d}x)$. The following lemma shows how a continuous density kernel can transform discountinuous functions into continuous ones under integration.

Lemma A.5.30. *If stochastic kernel P from U to X has density kernel p with dominating measure μ and, in addition, the mapping $u \mapsto p(u, x')$ is continuous on U for all μ-almost all $x' \in \mathsf{X}$, then P is strong Feller.*

Proof. Fix $h \in bm\mathsf{X}$ and $u_n \to u$ in U. With M such that $|h| \leqslant M$, we have

$$|(Ph)(u_n) - (Ph)(u)| \leqslant M \int |p(u_n, x') - p(u, x')|\mu(\mathrm{d}x').$$

By the continuity condition on p, we have $p(u_n, x') \to p(u, x')$ for all μ-almost all $x' \in \mathsf{X}$. The claim now follows from Scheffé's lemma (p. A.4.7). □

Example A.5.20. Suppose that $\mathsf{U} = \mathbb{R}^k$ and $\mathsf{X} = \mathbb{R}^\ell$. Let g be a continuous map from U to X. We consider the stochastic kernel from U to X given by $(Ph)(u) = h[g(u) + w]\varphi(w)\,\mathrm{d}w$, where the density φ is continuous on X. In this setting, P is strong Feller. Indeed, the change of variable $x' = g(u) + w$ yields

$$(Ph)(u) = \int h[g(u) + w]\varphi(w)\,\mathrm{d}w = \int h(x')\varphi(x' - g(u))\,\mathrm{d}x'$$

for all $h \in bm\mathsf{X}$. Specializing h to indicator functions shows that $p(x, x') = \varphi(x' - g(x))$ is a density kernel for P. The strong Feller property now follows from continuity of φ and g, combined with lemma A.5.30.

A.5.5 Orders over Distributions

Distributions are objects that decision makers naturally have preferences over. For example, speculators care about probability distributions over returns of prospective investments, often preferring distributions that offer high average returns with low risk. A planner might have preferences over the cross-sectional distributions of consumption and wealth. In this section, we discuss common methods for ordering distributions and their relationships with each other.

A.5.5.1 Stochastic Dominance

Let X be a metric space and let $\mathscr{D}(\mathsf{X})$ be the set of all distributions (i.e., Borel probability measures) on X. Let $ib\mathsf{X}$ be the increasing bounded real-valued functions on X. For μ and ν in $\mathscr{D}(\mathsf{X})$, we say that

- ν **first order stochastically dominates** μ and write $\mu \preceq_{\mathrm{F}} \nu$ if

$$\int u(x)\mu(\mathrm{d}x) \leqslant \int u(x)\nu(\mathrm{d}x) \quad \text{for every } u \text{ in } ib\mathsf{X} \text{ and}$$

- ν **second order stochastically dominates** μ and write $\mu \preceq_{\mathrm{S}} \nu$ if

$$\int u(x)\mu(\mathrm{d}x) \leqslant \int u(x)\nu(\mathrm{d}x) \quad \text{for every concave } u \text{ in } ib\mathsf{X}.$$

If we refer to stochastic dominance without explicitly stating the order, then the understanding is that we mean *first* order stochastic dominance.

Example A.5.21. If x and y are points in $\mathbb{R}$ with $x \leqslant y$, then $\delta_x \preceq_{\mathrm{F}} \delta_y$, since, for any increasing function u we have $u(x) \leqslant u(y)$.

Exercise A.5.26. Let Y be a random variable on $\mathbb{R}$ with distribution μ. Let m be a nonnegative constant and let ν be the distribution of $Y+m$. Show that μ is stochastically dominated by ν.

Suppose now that X is a Borel subset of $\mathbb{R}$ and fix $F, G \in \mathscr{D}(\mathsf{X})$. We understand F and G as cumulative distribution functions. When testing first order stochastic dominance, one can show (add ref) that it is sufficient to restrict attention to increasing functions $u \in b\mathsf{X}$ that take the form $u(x) = \mathbb{1}\{a < x\}$ for some $a \in \mathsf{X}$. Recalling the interpretation of the integral given in (A.7), this leads to the statement that $F \preceq_{\mathrm{F}} G$ if and only if $1 - F(a) \leqslant 1 - G(a)$ for all $a \in \mathsf{X}$, or

$$F \preceq_{\mathrm{F}} G \iff G(x) \leqslant F(x) \quad \text{for all } x \in \mathsf{X} \tag{A.29}$$

Exercise A.5.27. The relation $\preceq_{\mathrm{F}}$ yields a partial order on $\mathscr{D}(\mathsf{X})$. Prove this in the one-dimensional setting from the previous paragraph, where $\mathsf{X} \subset \mathbb{R}$.

Exercise A.5.28. Let $\mathsf{X} \subset \mathbb{R}$ and consider the set $\mathscr{D}(\mathsf{X})$ paired with the metric

$$d_\infty(F, G) = \sup_{x \in \mathsf{X}} |F(x) - G(x)|$$

In this context, the metric d_∞ is usually called the **Kolmogorov distance**. Show that $\preceq_{\mathrm{F}}$ is a closed partial order on $(\mathscr{D}(\mathsf{X}), d_\infty)$.

A.5.5.2 Montone Likelihood Ratios

Here is a property that implies first order stochastic dominance: Consider a pair of distributions (F, G) with positive densities f and g on an interval I contained in $\mathbb{R}$. We say that (f, g) has a **monotone likelihood ratio** if f/g is increasing on I; that is, if

$$x, x' \in I \text{ and } x \leqslant x' \implies \frac{f(x)}{g(x)} \leqslant \frac{f(x')}{g(x')} \tag{A.30}$$

Example A.5.22. The exponential density is $p(x, \lambda) = \lambda e^{-\lambda x}$ on $\mathbb{R}_+$, where λ is a positive constant. Taking the ratio $r(x) = p(x, \lambda_1)/p(x, \lambda_2)$ of exponential densities with $\lambda_1 \leqslant \lambda_2$, we have

$$r(x) = \frac{\lambda_1}{\lambda_2} \exp((\lambda_2 - \lambda_1)x) \qquad (x \in \mathbb{R}_+).$$

Since r is increasing in x, the monotone likelihood ratio property holds.

Proposition A.5.31. *If (f, g) has the monotone likelihood ratio property on I, then $G \preceq_F F$.*

Proof. Let $a := \inf I$ and $b := \sup I$. (These values can be infinite.) Writing the monotone likelihood ratio property as

$$x \leqslant x' \implies f(x)g(x') \leqslant f(x')g(x) \tag{A.31}$$

and integrating with respect to x from a to x' gives $F(x')g(x') \leqslant f(x')G(x')$. Also, integrating (A.31) with respect to x' from x to b gives $f(x)[1-G(x)] \leqslant [1-F(x)]g(x)$. Setting $x = x' = y$ in the last two inequalities yields

$$\frac{1-G(y)}{1-F(y)} \leqslant \frac{g(y)}{f(y)} \leqslant \frac{G(y)}{F(y)}.$$

This implies $F(y) \leqslant G(y)$ for arbitrary y, so $G \preceq_{\mathrm{F}} F$. □

A.5.5.3 Mean-Preserving Spreads

We will be concerned with analyzing how behavior changes when decisions become "riskier" in some sense. To analyze such senarios, we introduce the notion of a mean-preserving spread. In particular, for a given distribution φ, we say that ψ is a **mean-preserving spread** of φ if there exists a pair of random variables (Y, Z) such that

$$\mathbb{E}[Z \,|\, Y] = 0, \quad Y \stackrel{d}{=} \varphi \quad \text{and} \quad Y + Z \stackrel{d}{=} \psi$$

Thus, ψ is a mean-preserving spread of φ if it adds noise without changing the mean.

Exercise A.5.29. Let $\varphi = N(0,1)$ and let $\psi = N(0,2)$. Show that ψ is a mean-preserving spread of φ.

Exercise A.5.30. Prove that if φ is a mean-preserving spread of ψ, then ψ second order stochastically dominates φ. [Hint: Use Jensen's inequality.]

A.6 Chapter Notes

References, background, historical notes.

Appendix B

Solutions

Solution to Exercise 1.1.1. Let Assumption 1.1.1 hold and fix $v \in V$. Let σ be any policy in Σ. Since v and u are bounded, the function

$$(T_\sigma v)(w) = u(\sigma(w)) + \beta \int v(R(w - \sigma(w)) + y)\varphi(\mathrm{d}y)$$

is also bounded. Measurability follows from measurability of v, continuity of u and measurability of σ.

Solution to Exercise 1.1.2. Lemma 1.1.2 tells us that T maps V into $bc\mathbb{R}_+$, which is a subset of $bm\mathbb{R}_+$. In particular, T is a self-map on V. For the contraction property, we apply the sup inequality from Lemma A.1.2 and the triangle inequality for integrals to obtain

$$\begin{aligned} |(Tv)(w) - (Tv')(w)| &\leqslant \max_{0 \leqslant c \leqslant w} \beta \int |v(R(w-c)+y) - v'(R(w-c)+y)| \, \varphi(\mathrm{d}y) \\ &\leqslant \beta \|v - v'\|. \end{aligned}$$

Taking the supremum gives $\|Tv - Tv'\| \leqslant \beta \|v - v'\|$.

Solution to Exercise 1.2.1. Fix $T_\sigma \in \mathbb{T}_{\mathrm{OS}}$, which we can write as $T_\sigma v = r_\sigma + \beta P_\sigma v$ (see (1.4) on page 4). Recalling that the utility function is bounded (by Assumption 1.1.1), the constant $M := \sup u$ is finite. With $\bar{v} := M/(1-\beta)$, we have $T_\sigma \bar{v} \leqslant M + \beta P_\sigma \bar{v} = M + \beta M/(1-\beta) = \bar{v}$. Hence $(V, \mathbb{T}_{\mathrm{OS}})$ is bounded above. Order continuity follows from Exercise A.5.18.

Solution to Exercise 1.2.2. Since $(V, \mathbb{T})$ is bounded above we can take a $u \in V$ with $T_\sigma u \preceq u$ for all σ. By downward stability, $v_\sigma \preceq u$ for all σ. Hence u is an upper bound of V_Σ.

Solution to Exercise 1.2.5. Let the stated conditions hold. Suppose first that σ is optimal, so that $v_\sigma = v_\top$. Since $Tv_\top = v_\top$ we have $Tv_\sigma = v_\sigma$. For the reverse implication, suppose that $Tv_\sigma = v_\sigma$ holds. As $v_\top$ is the unique fixed point of T in V we have $v_\sigma = v_\top$. Since $v_\top$ is the greatest element of V_Σ, this proves that σ is an optimal policy.

Solution to Exercise 1.2.6. Convergence of OPI implies convergence of VFI because OPI reduces to VFI when $m = 1$ (since $Wv = T_\sigma v$ when σ is v-greedy, so $W = T$).

Solution to Exercise 1.3.1. Fix $v \in V_U \cap V_0$. Let $(v_n) := (T^n v)$, as in the proof of Theorem 1.3.2. The sequence (v_n) is increasing, with $v_n \preceq v_\top$ for all n (see Lemma 1.2.11). By $v_n \to v_\top$ and Lemma A.5.16 we get $T^n v \uparrow v_\top$.

Solution to Exercise 1.4.1. Let $(V, \mathbb{T})^\partial$ be the dual of $(V, \mathbb{T})$. Each $T_\sigma \in \mathbb{T}$ is a self-map on the poset V^∂. Moreover, for any $T_\sigma \in \mathbb{T}$, we have

$$v \preceq^\partial w \implies w \preceq v \implies T_\sigma w \preceq T_\sigma v \implies T_\sigma v \preceq^\partial T_\sigma w.$$

Hence T_σ is order preserving on V^∂.

Solution to Exercise 1.4.2. Regarding (i), fix $v \in V$. Policy σ is v-min-greedy for $(V, \mathbb{T})$ if and only if $T_\sigma v \preceq T_\tau v$ for all $\tau \in \Sigma$, which is equivalent to $T_\tau v \preceq^\partial T_\sigma v$ for all $\tau \in \Sigma$. Hence σ is v-min-greedy for $(V, \mathbb{T})$ if and only if σ is v-max-greedy for $(V, \mathbb{T})^\partial$.

Claim (ii) follows from (i). Claim (iii) is immediate from Exercise A.1.15 on page 158. Claims (iv)–(v) are also straightforward, and details are left to the reader.

Solution to Exercise 2.1.1. Each T_σ has a unique fixed point in $\mathbb{R}^\mathsf{X}$ given by $v_\sigma := (I - \beta P_\sigma)^{-1} r_\sigma$. T_σ is upward stable on $\mathbb{R}^\mathsf{X}$ because, given $v \in \mathbb{R}^\mathsf{X}$ with $v \leqslant T_\sigma v$, we have $(I - \beta P_\sigma) v \leqslant r_\sigma$ and, since $(I - \beta P_\sigma)^{-1} = \sum_{t=1}^\infty (\beta P_\sigma)^t$ is nonnegative, $v \leqslant (I - \beta P_\sigma)^{-1} r_\sigma = v_\sigma$. Reversing the inequalities shows that downward stability holds.

Solution to Exercise 2.1.3. Since $\mathbb{R}^\mathsf{X}$ is endowed with the pointwise partial order, for given $v \in \mathbb{R}^\mathsf{X}$ and $x \in \mathsf{X}$, the ADP Bellman operator $Tv = \bigvee_\sigma T_\sigma v$ reduces to

$$(T\, v)(x) = \sup_{\sigma \in \Sigma} (T_\sigma\, v)(x) = \sup_{\sigma \in \Sigma} \left\{ r(x, \sigma(x)) + \beta \sum_{x'} v(x') P(x, \sigma(x), x') \right\}.$$

By the definition of Σ, we can also write this as

$$(T\,v)(x) = \max_{a \in \Gamma(x)} \left\{ r(x,a) + \beta \sum_{x'} v(x') P(x,a,x') \right\}, \tag{2.6}$$

which is identical to (2.5).

Solution to Exercise 2.1.4. Regarding the first question, fixing $\sigma \in \Sigma$ we have

$$v \in \hat{V} \implies |T_\sigma v| \leqslant M + \beta P |v| \leqslant \frac{M}{1-\beta} \implies Tv \in \hat{V}.$$

Hence T_σ is a self-map on $\hat{V}$. The proof that $v_\sigma \in \hat{V}$ is left to the reader.

Solution to Exercise 2.1.5. For arbitrary $v, w \in \mathbb{R}^{\mathsf{X}}$ and $x \in \mathsf{X}$,

$$\begin{aligned}
|(Tv)(x) - (Tw)(x)| &\leqslant \beta \max_{a \in A} \left| \sum_{x'} v(x') P(x,a,x') - \sum_{x'} w(x') P(x,a,x') \right| \\
&\leqslant \beta \max_{a \in A} \sum_{x'} |v(x') - w(x')| \, P(x,a,x') \\
&\leqslant \beta \|v - w\|,
\end{aligned}$$

where $\|\cdot\|$ is the supremum norm, and hence $\|Tv - Tw\| \leqslant \beta \|v - w\|$. In the first step, the $\max$ operation is taken outside the absolute value via the inequality in Lemma A.1.2.

Solution to Exercise 2.1.6. We have discussed the fact that that the ADP $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\mathrm{MDP}})$ is regular and globally stable. The ADP is also bounded above, since with integer M such that $|r| \leqslant M$ and $v := M\,(1-\beta)$ we have $T_\sigma v \leqslant v$ for all $\sigma \in \Sigma$. Since $\mathbb{R}^{\mathsf{X}} = \mathbb{R}^{\mathsf{X}}$ is Dedekind complete (Example A.5.2), the conditions of Theorem 1.3.4 are verified.

Solution to Exercise 2.1.7. In the solution to Exercise 2.1.6 we discussed the fact that $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\mathrm{MDP}})$ is regular and bounded above, and that $\mathbb{R}^{\mathsf{X}} = \mathbb{R}^{\mathsf{X}}$ is Dedekind complete. Moreover, each policy operator has the form $T_\sigma\, v = r_\sigma + K_\sigma\, v$ where $K_\sigma := \beta P_\sigma$. Since $\rho(\beta P_\sigma) = \beta < 1$, the conditions in part (b) of Theorem 1.3.9 are verified.

Solution to Exercise 2.1.8. We showed that $(\mathbb{R}^{\mathsf{X}}, \mathbb{T}_{\mathrm{MDP}})$ is regular and hence $(\hat{V}, \mathbb{T}_{\mathrm{MDP}})$ is regular. Moreover, since every $T_\sigma \in \mathbb{T}_{\mathrm{MDP}}$ is globally stable on $\mathbb{R}^{\mathsf{X}}$ and hence $\hat{V}$, the ADP $(\hat{V}, \mathbb{T}_{\mathrm{MDP}})$ globally stable and, therefore, strongly order stable (Lemma A.5.17). Given that $\hat{V}$ is chain complete (Exercise A.5.1), Theorem 1.2.13 implies that the fundamental optimality properties hold and VFI, OPI, and HPI all converge.

Solution to Exercise 2.1.11. Fix $v \in bm\mathsf{X}$ and $\sigma \in \Sigma$. The composition, arithmetic and integral operations in the definition of T_σ all preserve measurability, so the Borel measurability restrictions on the primitives imply that $T_\sigma\, v$ is Borel measurable. Regarding boundedness, taking M such that $|r| \leqslant M$, we have

$$|(T_\sigma v)(x)| \leqslant |r(x, \sigma(x)| + \beta \left| \int v(x')\Pi(x, \sigma(x), \mathrm{d}x') \right| \leqslant M + \beta \|v\|.$$

Hence $T_\sigma\, v$ is also bounded.

Solution to Exercise 2.1.13. Fix $v \in bc\mathsf{X}$. The proof that Tv is bounded is similar to the proof that T_σ is bounded, which is in the solution to Exercise 2.1.11. Continuity of Tv follows from Theorem A.3.3 on page 179.

Solution to Exercise 2.1.15. Fix $\sigma \in \Sigma$ and $v \in bm\mathsf{X}$. We need to show that $T_\sigma\, v$ is bounded and measurable. Measurability follows easily from $v \in bm\mathsf{X}$ combined with measurability of σ and the primitives r and p. Boundedness holds because, with $\| \cdot \|$ as the supremum norm,

$$|(T_\sigma\, v)(x)| \leqslant |r(x, \sigma(x'))| + \beta \int |v(x')| p(x, \sigma(x), x') \mu(\mathrm{d}x'), \leqslant \|r\| + \beta \|v\|.$$

Finally, each T_σ is order preserving on $bm\mathsf{X}$. We conclude that $(bm\mathsf{X}, \mathbb{T}_{\mathrm{SF}})$ is an ADP.

Solution to Exercise 2.2.5. We set $\mathsf{X} = \mathsf{S} \times \mathsf{E}$ and $\mathsf{A} = \{0, 1\}$. The feasible correspondence Γ is defined by $\Gamma(x) = \mathsf{A}$ for all x. The set A is given the discrete topology. The reward function is set to

$$r(x, a) = r(s, p, a) = a(ps - m(s)) - (1 - a)c,$$

while the stochastic kernel P from G to X is described by

$$(Ph)(x, a) = (Ph)(s, p, a) = \int h(f(s, a), p')\varphi(\mathrm{d}p').$$

Since f is continuous on $\mathsf{S} \times \mathsf{A}$, the kernel P is weak Feller. With a small amount of effort, it can be shown that the Bellman equation for this weak Feller MDP agrees with (2.24).

Solution to Exercise 2.2.7. Fix $(x, a) \in \mathsf{G}$ and $v, w \in V$. Let (2.31) hold. By the monotonicty of B we have

$$B(x,a,v) = B(x,a,w+v-w) \leqslant B\left(x,a,w+\frac{|v-w|}{\ell}\ell\right) \leqslant B\left(x,a,w+\|v-w\|\,\ell\right).$$

Applying (2.31) and rearranging gives $(B(x,a,v)-B(x,a,w))/\ell \leqslant \lambda\|v-w\|$. Reversing the roles of v and w yields (U3).

Solution to Exercise 2.2.9. Since $u'(c) \to 0$ as $c \to \infty$ and u is concave, there exists a $c_0 \in \mathbb{R}_+$ such that $u'(c) \leqslant 1$ when $c \geqslant c_0$. By the mean value theorem, for all $c \geqslant c_0$,

$$u(c) \leqslant u(c_0) + c - c_0 = u(c_0) + \frac{c-c_0}{c}c.$$

Let $\gamma := u(c_0)$ and $\delta := \frac{c-c_0}{c} \leqslant 1$. Then $u(c) \leqslant \gamma + \delta c$ for all $c \geqslant c_0$.

Now consider $c < c_0$. Since u is increasing, $u(c) \leqslant u(c_0) := \gamma$. Hence, for any non-negative $\delta \leqslant 1$, $u(c) \leqslant \gamma + \delta c$. Combining the two cases gives the statement in the exercise.

Solution to Exercise 2.3.1. Fix $\sigma \in \Sigma$ and $q, f \in \mathbb{R}^{\mathsf{G}}$. For each $(x,a) \in \mathsf{G}$, we have

$$\begin{aligned}
|(S_\sigma q)(x,a) - (S_\sigma f)(x,a)| &\leqslant \beta \sum_{x'} |q(x',\sigma(x')) - f(x',\sigma(x'))| P(x,a,x') \\
&\leqslant \beta \sum_{x'} \|q-f\|_\infty P(x,a,x') = \beta\|q-f\|_\infty.
\end{aligned}$$

Taking the supremum over (x,a) yields $\|S_\sigma q - S_\sigma f\|_\infty \leqslant \beta\|q-f\|_\infty$.

Solution to Exercise 2.3.2. Fix $\sigma \in \Sigma$. Clearly, S_σ is a self-map on $\mathbb{R}^{\mathsf{G}}$. For any $q \leqslant f \in \mathbb{R}^{\mathsf{G}}$, we have $q(x,a) \leqslant f(x,a)$ for each $(x,a) \in \mathsf{G}$, and therefore $(S_\sigma q)(x,a) \leqslant (S_\sigma f)(x,a)$ for each $(x,a) \in \mathsf{G}$. This implies $S_\sigma q \leqslant S_\sigma f$, so S_σ is order preserving.

Solution to Exercise 2.3.3. If $\sigma(x) \in \operatorname{argmax}_{a\in\Gamma(x)} q(x,a)$ for all $x \in \mathsf{X}$, then $q(x,\sigma(x)) = \max_{a\in\Gamma(x)} q(x,a)$ for all $x \in \mathsf{X}$. Hence, for any $\tau \in \Sigma$, we have $(S_\tau q)(x,a) \leqslant (S_\sigma q)(x,a)$ for each $(x,a) \in \mathsf{G}$, which implies that $S_\tau q \leqslant S_\sigma q$. Hence, σ is q-greedy policy.

Solution to Exercise 2.3.4. Fix $q \in \mathbb{R}^{\mathsf{G}}$ and choose σ such that $\sigma(x) \in \operatorname{argmax}_{a\in\Gamma(x)} q(x,a)$ for all $x \in \mathsf{X}$. By Exercise 2.3.4, $\sigma \in \Sigma$ is q-greedy. The representation (2.37) now follows from $Sq = S_\sigma q$ (see Lemma 1.2.1).

Solution to Exercise 2.3.5. Since A is finite and Γ is nonempty, the characterization of greedy policies in Exercise 2.3.3 implies that $(\mathbb{R}^{\mathsf{G}}, \mathbb{S})$ is regular. By Exercise 2.3.1, each $S_\sigma \in \mathbb{S}$ is a contraction of modulus β on $\mathbb{R}^{\mathsf{G}}$. The claim now follows from Theorem 1.3.5.

Solution to Exercise 2.3.6. It is clear from the definition of the policy operators in (2.39) that, given $f \in \mathbb{R}^{\mathsf{G}}$, a policy $\sigma \in \Sigma$ is f-greedy whenever $\sigma(x) \in \operatorname{argmax}_{a \in \Gamma(x)} f(x, a)$ for all $x \in \mathsf{X}$. Since A is finite and the correspondence Γ is nonempty, at least one such policy exists.

Solution to Exercise 2.3.7. Fix $f \in \mathbb{R}^{\mathsf{G}}$ and let σ be such that $\sigma(x) \in \operatorname{argmax}_{a \in \Gamma(x)} f(x, a)$ for all $x \in \mathsf{X}$. This policy is f-greedy and, therefore, $Tf = T_\sigma f$ (see Lemma 1.2.1). As a result, $Tf = f$ holds if and only if $T_\sigma f = f$. Clearly, for this choice of σ, we have $T_\sigma f = f$ if and only if f solves (2.38).

Solution to Exercise 2.3.8. Fix $f \in \mathbb{R}^{\mathsf{G}}$. By Exercise 2.3.7, we can pick up a f-greedy policy σ satisfying $\sigma(x) \in \operatorname{argmax}_{a \in \Gamma(x)} f(x, a)$ for all $x \in \mathsf{X}$. It follows from $f = Tf = T_\sigma f$ (see Lemma 1.2.1).

Solution to Exercise 2.3.12. Let F be as stated and fix $t \in \mathbb{N}$. Let $E = A + BF$. Observe that $\|x_t\| \leqslant \|E^t\| \|x_0\|$, where the first norm on the right hand side is the operator norm on $\mathbb{R}^{k \times k}$. By Gelfand's formula for the spectral radius (page 200), there exists an $\varepsilon \in (0, 1)$ such that $\|E^t\| \leqslant \varepsilon^t$ for all sufficiently large t. Hence $\|x_t\| \leqslant \varepsilon^t \|x_0\|$ for all such t. The claim follows.

Solution to Exercise 2.3.14. Fix $N \in \mathcal{N}$ and $x \in \mathbb{R}^k$. Let $F = \mathbf{F}N$. Since Fx is the maximizer of (2.46) when $v(x) = x^\top N x$, we have

$$x^\top F^\top R F x + x^\top Q x + x^\top (A + BF)^\top N (A + BF) x \geqslant u^\top R u + x^\top Q x + v(Ax + Bu) \tag{2.56}$$

for any $u \in \mathbb{R}^m$. Setting G to be any control matrix and letting $u = Gx$ allows us to write (2.56) as $x^\top T_G N x \leqslant x^\top T_F N x$. Hence $T_G N \preccurlyeq T_F N$.

Now let $v(x) = x^\top N x$ and let F be any control matrix. If $T_G N \preccurlyeq T_F N$ holds for any control matrix G, then it holds when when G is N-maximal. Using this choice of G and fixed $x \in \mathbb{R}^k$, we get $x^\top T_G N x \leqslant x^\top T_F N x$ and hence

$$u_F^\top R u_F + x^\top Q x + v(Ax + Bu_F) \geqslant \max_{u \in \mathbb{R}^m} \left\{ u^\top R u + x^\top Q x + v(Ax + Bu) \right\}$$

where $u_F = Fx$. Lemma 2.3.5 now implies that $F = -(B^\top N B + R)^{-1} B^\top N A$. Hence F is N-maximal.

Solution to Exercise 3.1.3. We address the first claim. Let F be an order isomorphism F from V to $\hat{V}$ and observe that, by reflexivity, $Fv = Fv'$ implies $Fv \trianglelefteq Fv'$ and $Fv' \trianglelefteq Fv$. Since F is order isomorphic, this yields $v \preceq v'$ and $v' \preceq v$. Using antisymmetry, we get $v = v'$. Hence F is one-to-one as well as onto, and therefore bijective.

Solution to Exercise 3.1.5. Let $\{p_\alpha\}_{\alpha \in \Lambda}$ be a subset of V, let V, $\hat{V}$, and F be as stated, and let $\bar{p} := \bigvee_\alpha p_\alpha$. We need to show that $\bar{q} := F\bar{p}$ is the supremum of $\{Fp_\alpha\}_{\alpha \in \Lambda}$. First, $p_\alpha \preceq \bar{p}$ for all α, so $Fp_\alpha \preceq F\bar{p} = \bar{q}$ for all α. In particular, $\bar{q}$ is an upper bound of $\{Fp_\alpha\}_{\alpha \in \Lambda}$. Moreover, if u is any upper bound of $\{Fp_\alpha\}_{\alpha \in \Lambda}$, then $Fp_\alpha \preceq u$ and hence $p_\alpha \preceq F^{-1}u$ for all α, so $\bar{p} \preceq F^{-1}u$. But then $\bar{q} = F\bar{p} \preceq u$. Hence $\bar{q}$ is the supremum of $\{Fp_\alpha\}_{\alpha \in \Lambda}$, as was to be shown.

Solution to Exercise 3.1.8. Let v be the unique fixed point of S in V. Then, using $F \circ S = \hat{S} \circ F$, we have $\hat{S}Fv = FSv = Fv$, so Fv is a fixed point of $\hat{S}$ in $\hat{V}$. Moreover, if w is another fixed point of $\hat{S}$ in $\hat{V}$, then, by rearranging $F \circ S = \hat{S} \circ F$, we obtain $SF^{-1}w = F^{-1}\hat{S}w = F^{-1}w$, so $F^{-1}w$ is a fixed point of S. Since v is the only fixed point of S, we then have $v = F^{-1}w$, or $w = Fv$. In particular, Fv is the only fixed point of $\hat{S}$.

Solution to Exercise 3.1.9. Let D be the set of all dynamical systems (V, S) where V is partially ordered. For (V, S) and $(\hat{V}, \hat{S})$ in D, we write $(V, S) \sim (\hat{V}, \hat{S})$ when (V, S) and $(\hat{V}, \hat{S})$ are order conjugate. We claim that $\sim$ is reflexive, symmetric and transitive. Reflexivity is obvious: every (V, S) in D is order conjugate to itself under the identity map I. Symmetry is also straightforward: If (V, S) and $(\hat{V}, \hat{S})$ are order conjugate under F, then (V, S) and $(\hat{V}, \hat{S})$ are order conjugate under F^{-1}. Finally, if $(V, S) \sim (V', S')$ under F and $(V', S') \sim (V'', S'')$ under G, then $G \circ F$ is an order isomorphism from V to V'' and

$$G \circ F \circ S = G \circ S' \circ F = S'' \circ G \circ F \text{ on } V.$$

Hence $(V, S) \sim (V'', S'')$ and $\sim$ is also transitive.

Solution to Exercise 3.1.10. Let (V, S) and $(\hat{V}, \hat{S})$ be order conjugate under F, with respective fixed points v and $\hat{v} = Fv$. Let S be order stable on V and let $\hat{w}$ be an element of $\hat{V}$ satisfying $\hat{S}\hat{w} \preceq \hat{w}$. Then $F^{-1}\hat{S}\hat{w} \preceq F^{-1}\hat{w}$ and hence $SF^{-1}\hat{w} \preceq F^{-1}\hat{w}$. But then $v \preceq F^{-1}\hat{w}$, by downward stability of S. Applying F gives $\hat{v} \preceq \hat{w}$. Hence $\hat{S}$ is downward stable on $\hat{V}$. Similarly, if $\hat{w}$ is an element of $\hat{V}$ satisfying $\hat{w} \preceq \hat{S}\hat{w}$, then $F^{-1}\hat{w} \preceq F^{-1}\hat{S}\hat{w} = SF^{-1}\hat{w}$. By upward stability of S on V, we have $F^{-1}\hat{w} \preceq v$. Applying F gives $\hat{w} \preceq \hat{v}$, so $\hat{S}$ is upward stable on $\hat{V}$. Together, these results show that $\hat{S}$ is order stable on $\hat{V}$.

Solution to Exercise 3.1.13. Suppose first that $\theta > 0$, so that $v_1 = m_1$ and $v_2 = m_2$. Then

$$\hat{T}_\sigma v_1 = \left\{ r_\sigma^\alpha + \beta \frac{\min r^\alpha - \varepsilon}{1-\beta} \right\}^\theta \geqslant \left\{ \min r^\alpha + \beta \frac{\min r^\alpha - \varepsilon}{1-\beta} \right\}^\theta = \left\{ \frac{\min r^\alpha - \beta\varepsilon}{1-\beta} \right\}^\theta > m_1 = v_1.$$

In addition,

$$\hat{T}_\sigma v_2 = \left\{ r_\sigma^\alpha + \beta \frac{\max r^\alpha + \varepsilon}{1-\beta} \right\}^\theta \leqslant \left\{ \max r + \beta \frac{\max r^\alpha + \varepsilon}{1-\beta} \right\}^\theta = \left\{ \frac{\max r^\alpha + \beta\varepsilon}{1-\beta} \right\}^\theta < m_2 = v_2.$$

If $\theta < 0$, then $v_1 = m_2$ and $v_2 = m_1$, so

$$\hat{T}_\sigma v_1 = \left\{ r_\sigma^\alpha + \beta \frac{\max r^\alpha + \varepsilon}{1-\beta} \right\}^\theta \geqslant \left\{ \max r^\alpha + \beta \frac{\max r^\alpha + \varepsilon}{1-\beta} \right\}^\theta = \left\{ \frac{\max r^\alpha + \beta\varepsilon}{1-\beta} \right\}^\theta > m_2 = v_1.$$

In addition,

$$\hat{T}_\sigma v_2 = \left\{ r_\sigma^\alpha + \beta \frac{\min r^\alpha - \varepsilon}{1-\beta} \right\}^\theta \leqslant \left\{ \min r^\alpha + \beta \frac{\min r^\alpha - \varepsilon}{1-\beta} \right\}^\theta = \left\{ \frac{\min r^\alpha - \beta\varepsilon}{1-\beta} \right\}^\theta < m_1 = v_2.$$

Solution to Exercise 3.1.14. For $c, t > 0$, let $f(t) := (c + \beta t^{1/\theta})^\theta$. Simple calculations show that $f' > 0$, that $f'' < 0$ when $\theta < 0$ or $1 \leqslant \theta$, and that $f'' > 0$ when $0 < \theta \leqslant 1$. Hence f is concave when $\theta < 0$ or $1 \leqslant \theta$, and convex when $0 < \theta \leqslant 1$. The claim in the exercise follows easily from these facts and the definition of $\hat{T}_\sigma$ in (3.14).

Solution to Exercise 3.1.15. Fix $\sigma \in \Sigma$ and $v \in V$. On the one hand,

$$F T_\sigma v = (T_\sigma v)^\gamma = \left\{ r_\sigma + \beta (P_\sigma v^\gamma)^{\alpha/\gamma} \right\}^{\gamma/\alpha} = \left\{ r_\sigma + \beta (P_\sigma v^\gamma)^{1/\theta} \right\}^\theta .$$

On the other,

$$\hat{T}_\sigma F v = \hat{T}_\sigma v^\gamma = \left\{ r_\sigma + \beta (P_\sigma v^\gamma)^{1/\theta} \right\}^\theta .$$

Hence $F \circ T_\sigma = \hat{T}_\sigma \circ F$ on I, as claimed.

Solution to Exercise 4.1.1. Fix $\sigma \in \Sigma$. Straightforward arguments show that T_σ is a contraction of modulus β on $L_1(\varphi)$. Since $L_1(\varphi)$ is complete and $\beta \in (0,1)$, it follows that T_σ has a unique fixed point. Hence $(L_1(\varphi), \mathbb{T})$ is well-posed.

Solution to Exercise 4.1.2. This follows immediately from (4.6), which shows $T_\sigma v$ at a v-greedy policy σ, and Lemma 1.2.1.

Solution to Exercise 4.1.4. We set $\Gamma(w) = \{0, 1\}$ for every $w \in \mathsf{W}$, while $V = bm\mathsf{W}$ and

$$B(w, a, v) = a\frac{w}{1-\beta} + (1-a)\left[c + \beta \int v(w')\varphi(\mathrm{d}w')\right] \tag{4.8}$$

The action a takes values in $\{0, 1\}$, where $a = 1$ means accept and $a = 0$ means reject. The state space is W. The tuple (Γ, V, B) is an RDP. The monotonicity in (2.19) is easily verified. For 2.20 we note that a feasible policy is a measurable map $\sigma\colon \mathsf{W} \to \{0, 1\}$ and, given any such σ, the function

$$m(w) := \sigma(w)\frac{w}{1-\beta} + (1-\sigma(w))\left[c + \beta \int v(w')P(w, \mathrm{d}w')\right]$$

is again in $bm\mathsf{X}$. (The function m is bounded because W is bounded.)

Solution to Exercise 4.1.9. We apply Proposition A.5.18 on page 215. Fixing $w \in \mathbb{R}_+$, it suffices to show that the value $f(w) := c(1-\beta) + \beta \int \max\{w', w\}\, \varphi(\mathrm{d}w')$ shifts up when β increases. This is true when $c \leqslant \bar{w}$, because $f(w)$ is the weighted average of two terms and the second term is larger than the first:

$$\int \max\{w', w\}\, \varphi(\mathrm{d}w') \geqslant \int w'\varphi(\mathrm{d}w') = \bar{w} \geqslant c.$$

Increasing β puts more weight on the larger term, so $f(w)$ increases with β.

Solution to Exercise 4.1.11. Fix $v \in L_1(\varphi)$ and $\sigma \in \Sigma$. For the self-map property, we need to show that $\sigma e + (1-\sigma)(c + \beta P v)$ is again in $L_1(\varphi)$. Borel measurability is obvious from Borel measurability elements of Σ and assumptions on the primitives. Regarding φ-integrability, it suffices to show that the individual terms in the sum are integrable. That σe is integrable follows from Assumption 4.1.2. Also, $(1-\sigma)c$ is integrable because φ is a probability measure. Finally, $0 \leqslant (1-\sigma)\beta P v \leqslant P v$ and P maps $L_1(\varphi)$ to itself (see Lemma A.5.29).

The order preserving property of T_σ follows from the fact that P is a positive linear operator.

Solution to Exercise 4.1.13. To see that T_σ is order continuous, observe that T_σ can be expressed as $T_\sigma v = a + K_\sigma v$, where $a \in L_1(\varphi)$ and $K_\sigma = \beta(1-\sigma)P \in \mathcal{B}(L_1(\varphi))$. By Corollary A.5.25, the operator K_σ is order continuous. It follows that T_σ is itself order continuous (see, e.g., Exercise A.5.6 on page 207).

Solution to Exercise 4.1.14. Since $\beta(1-\sigma)P \leqslant \beta P$ in the pointwise order on $\mathcal{B}(L_1(\varphi))$, an application of Theorem A.5.23 on page 218 yields $\rho(\beta(1-\sigma)P) \leqslant \rho(\beta P) = \beta\rho(P) =$

$\beta < 1$. Hence, by the Neumann series lemma, and in particular Corollary A.4.11, T_σ has unique fixed point in $L_1(\varphi)$ given by v_σ in (4.22).

Solution to Exercise 4.1.15. Fix $\sigma \in \Sigma$. Using the power series representation from the Neumann series lemma, we have

$$v_\sigma = \sum_{t \geqslant 0} [\beta(1-\sigma)P]^t (\sigma e + (1-\sigma)c) \leqslant \sum_{t \geqslant 0} (\beta P)^t (e + c) = \bar{v}.$$

Next we show that $v \in V$ implies $Tv \in V$. To this end, fix $v \in V$. Evidently $0 \leqslant T_\sigma v$. Moreover, since $v \leqslant \bar{v}$,

$$T_\sigma v \leqslant e + c + \beta P v \leqslant e + c + \beta P(I - \beta P)^{-1}(e + c).$$

Using the power series representation, the right hand side can be expressed as

$$e + c + \beta P \sum_{t \geqslant 0} (\beta P)^t (e + c) = e + c + \sum_{t \geqslant 1} (\beta P)^t (e + c) = \sum_{t \geqslant 0} (\beta P)^t (e + c) = \bar{v}.$$

We have confirmed that $0 \leqslant T_\sigma v \leqslant \bar{v}$ when $v \in V$, so each T_σ is a self-map on V.

Solution to Exercise 4.1.16. Since V is an order interval in the Dedekind complete Banach lattice $L_1(\varphi)$, V is σ-chain complete (see Lemma A.5.4). Moreover, $(V, \mathbb{T})$ is regular, well-posed and order continuous (see Exercises 4.1.12–4.1.14). The claims (i)–(iii) now follow from Theorem 1.2.14.

Solution to Exercise 4.1.17. If W is finite, then the set of policies Σ, which is the set of maps from W into $\{0, 1\}$, is also finite. In addition, $(V, \mathbb{T})$ is order stable (see, e.g., Lemma A.5.3 on page 203). Hence Theorem 1.2.12 applies.

Solution to Exercise 4.1.21. Fix $h \in L_1(\varphi)$. By Jensen's inequality and (4.15), we have

$$\begin{aligned} |(\hat{T}g)(z) - (\hat{T}h)(z)| &\leqslant \beta\, \mathbb{E}_z \left| \max\left\{ \frac{w'}{1-\beta}, g(z') \right\} - \max\left\{ \frac{w'}{1-\beta}, h(z') \right\} \right| \\ &\leqslant \beta\, \mathbb{E}_z \left| g(z') - h(z') \right|. \end{aligned}$$

Let Z be a draw from φ. Taking the expectation of the last inequality with $z = Z$ and using the fact that φ is stationary gives

$$\mathbb{E}|(\hat{T}g)(Z) - (\hat{T}h)(Z)| \leqslant \beta \mathbb{E}\mathbb{E}_Z \left| g(z') - h(z') \right| = \beta \mathbb{E} \left| g(z') - h(z') \right|.$$

This proves that $\|\hat{T}g - \hat{T}h\| \leqslant \beta \|g - h\|$, where $\|\cdot\|$ is the norm on $L_1(\varphi)$.

Solution to Exercise 4.1.24. Fix $v \in L_1(\varphi)$ and consider the policy $\sigma = \mathbb{1}\{q \geqslant v\}$. Under this policy we have

$$\sigma q + (1-\sigma)v = q \vee v \geqslant \tau q + (1-\tau)v \quad \text{for all } \tau \in \Sigma.$$

Since K is a positive operator, this yields $T_\sigma v \geqslant T_\tau v$ for all $\tau \in \Sigma$. Hence σ is v-greedy and $(L_1(\varphi), \mathbb{T})$ is regular.

Solution to Exercise 4.1.25. Fix $v \in L_1(\varphi)$ and recall from Lemma 1.2.1 that, when σ is v-greedy, we have $Tv = T_\sigma v$. We saw in Exercise 4.1.24 that $\sigma = \mathbb{1}\{q \geqslant v\}$ is v-greedy. Moreover, for this choice of σ, we have $T_\sigma v = \pi + K(\sigma q + (1-\sigma)v) = \pi + K(q \vee v)$.

Solution to Exercise 4.1.26. Observe that $K(1-\sigma) \leqslant K$ and $\rho(K) < 1$ together imply that $\rho(K(1-\sigma)) < 1$ (see Theorem A.5.23 on page 218). The Neumann series lemma now implies that v_σ in (4.33) is well-defined and is the unique fixed point of T_σ in $L_1(\varphi)$.

Solution to Exercise 4.1.27. For fixed $v, h \in V$ we have $T_\sigma(v+h) = \pi + K[\sigma q + (1-\sigma)(v+h)] \leqslant T_\sigma v + Kh$. Since K is a positive linear operator and $\rho(K) < 1$, the ADP $(V, \mathbb{T})$ is eventually Blackwell contracting and all the conclusions of Theorem 1.3.8 hold.

Solution to Exercise 4.2.2. Fix $\sigma \in \Sigma$ and take $(v_n) \subset V$ and $v \in V$ with $v_n \uparrow v$. We claim that $T_\sigma v_n \uparrow T_\sigma v$. Since V is closed under pointwise suprema, we know (by Lemma A.1.4) that v_n converges up to v pointwise, and, moreover, that our proof will be complete if we can show that $T_\sigma v_n$ converges up to $T_\sigma v$ pointwise. For this it is enough to show that, fixing $w \in \mathsf{W}$, the integral $\int \beta[v_n(w')]P(w, \mathrm{d}w')$ converges up to $\int \beta[v(w')]P(w, \mathrm{d}w')$. This convergence holds by monotonicty of the integral, monotone pointwise convergence of v_n to v, continuity of β and the dominated convergence theorem.

Solution to Exercise 4.2.6. Let $\|\cdot\|$ be the supremum norm on $bm(0,1)$. First we show that $\hat{T}$ is a self-mapping on $bm(0,1)$. To this end, pick any $\omega \in bm(0,1)$ and consider the function $\hat{T}\omega$ defined by (4.38). Evidently $\hat{T}\omega$ is Borel measurable. To see that this function is bounded, observe that, by the triangle inequality and the fact that φ_π is a density,

$$(\hat{T}\omega)(\pi) \leqslant (1-\beta)c + \beta \max\{M, \|\omega\|\} \tag{4.39}$$

The right hand side does not depend on π so $\hat{T}\omega$ is bounded as claimed.

Next let's establish the contraction property. Fix $\omega, \psi \in bm(0,1)$ and $\pi \in (0,1)$. Using the triangle inequality for integrals and the bound (4.15) on page 119 yields

$$|(\hat{T}\omega)(\pi) - (\hat{T}\psi)(\pi)| \leqslant \beta \int |\omega[\kappa(w',\pi)] - \psi[\kappa(w',\pi)]| \; \varphi_\pi(w') \; \mathrm{d}w' \leqslant \beta\|\omega - \psi\|.$$

Taking the supremum over π gives $\|\hat{T}\omega - \hat{T}\psi\| \leqslant \beta\|\omega - \psi\|$.

Solution to Exercise 4.2.7. Let ω be bounded and continuous on $(0,1)$. To show that $\hat{T}\omega$ is continuous, we need to prove that

$$\int \max\{w', \omega[\kappa(w',\pi_n)]\} \; \varphi_{\pi_n}(w') \; \mathrm{d}w' \to \int \max\{w', \omega[\kappa(w',\pi)]\} \; \varphi_\pi(w') \; \mathrm{d}w'$$

when (π_n) is a sequence converging to $\pi \in (0,1)$, then For fixed w', both $\kappa(w',\pi)$ and $\varphi_\pi(w')$ are continuous in π, so, by the dominated convergence theorem (page 186), it suffices to show that

$$H_n(w') := \max\{w', \omega[\kappa(w',\pi_n)]\} \; \varphi_{\pi_n}(w')$$

satisfies $\sup_n |H_n(w')| \leqslant H(w')$ for some $H\colon [0,M] \to \mathbb{R}$ with $\int H(w')\,\mathrm{d}w' < \infty$. Such an H does indeed exist: one suitable choice is

$$H(w') := \max\{M, \|\omega\|_\infty\} \; (f(w') + g(w')).$$

Solution to Exercise 4.2.8. For fixed α and any increasing bounded function u, we have

$$\int u \,\mathrm{d}H_\alpha = \alpha \int u \,\mathrm{d}G + \alpha \left(\int u \,\mathrm{d}F - \int u \,\mathrm{d}G \right)$$

By the fact that $G \preceq_{\mathrm{F}} F$ and u is increasing, this expression is increasing in α. Hence $\alpha \leqslant \beta$ implies $H_\alpha \preceq_{\mathrm{F}} H_\beta$ as claimed.

Solution to Exercise 4.2.10. Let $K := \beta P$ and $J := \gamma P$. Fix $f, g \in bm\mathrm{W}$ and $\sigma \in \Sigma$. Pointwise on W, we have

$$|T_\sigma f - T_\sigma g| = |\sigma J(f-g) + (1-\sigma)K(f-g)| \leqslant |J(f-g)| \vee |K(f-g)|.$$

But $|J(f-g)| \leqslant \gamma P|f-g|$ and hence

$$|J(f-g)| \leqslant \gamma\|f-g\|$$

A similar argument shows that $|K(f-g)| \leqslant \beta \|f-g\|$. Hence

$$|T_\sigma f - T_\sigma g| \leqslant (\beta \|f-g\|) \vee (\beta \|f-g\|) = (\beta \vee \gamma)\|f-g\|.$$

Since $\gamma < 1$, the operator T_σ is a contraction of modulus $\beta \vee \gamma$.

Solution to Exercise 4.2.12. It follows from (4.50) that if $c \leqslant \hat{c}$ and Q and $\hat{Q}$ are the corresponding continuation value operators, then $\hat{Q}h \leqslant Qh$ for any $h \in \mathbb{R}^{\mathsf{S}}$. Since Q is order preserving on $bc\mathsf{S}$ and also globally stable, Proposition A.5.18 on page 215 implies that, pointwise on S, the fixed point h of Q is larger than the fixed point $\hat{h}$ of $\hat{Q}$.

Now consider the decision problems of the two firms described in the exercise. The first one has lower maintenance cost and stops when $p_t s_t - m \geqslant h(s_t)$. The second one has higher maintenance costs and stops when $p_t s_t - m \geqslant \hat{h}(s_t)$. Since $\hat{h}(s_t) \leqslant h(s_t)$, we know that if the first firm decides to harvest then the second firm does too. In particular, the firm with higher maintenance costs harvest no later than the firm with low maintenance costs.

Solution to Exercise 4.2.13. The Bellman operator is

$$Tv(s,p) = \max_{a \in \{0,1\}} \left\{ a(ps - m) - (1-a)c + \beta \int v(q(s)(1-a), p')\varphi(\mathrm{d}p') \right\}$$

Given $v, w \in \mathbb{R}^{\mathsf{X}}$, we have, by [change this to cite §A.5.3.2] followed by the triangle inequality for integrals,

$$\begin{aligned}
|Tv(s,p) - Tw(s,p)| &\leqslant \beta \max_{a \in \{0,1\}} \left| \int [v(q(s)(1-a), p') - w(q(s)(1-a), p')]\varphi(p')\,\mathrm{d}p' \right| \\
&\leqslant \beta \max_{a \in \{0,1\}} \int |v(q(s)(1-a), p') - w(q(s)(1-a), p')|\varphi(p')\,\mathrm{d}p'
\end{aligned}$$

Hence

$$|Tv(s,p) - Tw(s,p)| \leqslant \beta \|v - w\|_\infty$$

Taking the supremum over all (s,p) in X leads to

$$\|Tv - Tw\|_\infty \leqslant \beta \|v - w\|_\infty$$

Since v, w were arbitrary elements of $\mathbb{R}^{\mathsf{X}}$, the contraction claim is established.

Solution to Exercise A.1.4. Point (i) is obvious: If $x = y$ then $x + y = x \vee y + x \wedge y$ certainly holds. If $x < y$, then $x \vee y + x \wedge y = y + x = x + y$. Regarding (ii), observe that, for any real values x, y, we have $x - y \leqslant x \vee y - x \wedge y$ and $y - x \leqslant x \vee y - x \wedge y$. Hence (ii) holds.

Solution to Exercise A.1.8. If A is order bounded, with $A \subset [u, v] \subset \mathbb{R}^n$, then, given $a \in A$, we have $|a_i| \leqslant |u_i| \vee |v_i| \leqslant \|u\|_\infty \vee \|u\|_\infty =: M$ for all i, and hence $\|a\|_\infty \leqslant M$. Hence A is bounded with respect to the norm $\| \cdot \|_\infty$, and therefore with respect to any norm on $\mathbb{R}^n$ by equivalence of norms (see §A.4.2.1). Conversely, if A is bounded, with $\|a\|_\infty \leqslant M$ for all $a \in A$, then $-M\mathbb{1} \leqslant a \leqslant M\mathbb{1}$ for all $a \in A$. Hence A is order bounded.

Solution to Exercise A.1.11. We prove (i). If $s \in V$, then $s \in U(G)$. Also, if $w \in U(G)$, then $g \leqslant w$ for all $g \in G$ and hence $s \leqslant w$. This proves that $s = \bigvee G$.

Solution to Exercise A.1.12. Let C be as stated and let G be all $f_n \in C$ with $f_n(x) = x^{1/n}$ for all x. Clearly $s(x) = \mathbb{1}\{x > 0\}$, which is not in C. At the same time, if $g \in C$ dominates all f_n, then g is continuous and equals 1 for all $x \in (0, 1]$, so $g \equiv 1$. Thus, $U(G)$, the set of upper bounds of G in C, is equal to $\{1\}$. In particular, $\bigvee G = 1$, so $\bigvee G$ and s are distinct.

Solution to Exercise A.1.13. Suppose first that $b := \bigvee \emptyset$ exists in V. If $v \in V$, then v is an upper bound of $\emptyset$, since the statement $u \preceq v$ for all $u \in \emptyset$ is vacuously true. Hence $b \preceq v$. This proves that b is the least element of V.

Next, suppose that b is the least element of V. Then b is a lower bound of $\emptyset$, since the statement $b \preceq v$ for all $v \in \emptyset$ is vacuously true. As the least element of V, b is also a least upper bound of $\emptyset$, so $\bigvee \emptyset = b$.

Solution to Exercise A.1.14. Let V, I, and D be as stated and let $s = \bigvee D$ in $(V, \preceq)$. Let U_V be the set of upper bounds of D in V and let U_I be the set of upper bounds of D in I (i.e., in the poset $(I, \preceq)$). Since $a \preceq d \preceq b$ for all $d \in D$ and s is the least element of U_V, it must be that $a \preceq s \preceq b$. Since $s \in U_V$ we have $d \preceq s$ for all $d \in D$ and hence $s \in U_I$. Moreover, for any $t \in U_I$, we have $t \in V$ and $d \preceq t$ for all $d \in D$, so $t \in U_V$. But then $s \preceq t$, so s is a least element of U_I. In particular, s is the supremum of A in $(I, \preceq)$.

Solution to Exercise A.1.16. Here's a proof of part (i). Assume that $v_n \uparrow v$ and $v_n \preceq u_n \preceq v$ for all n. Evidently v is an upper bound of (u_n). If w is any other bound of (u_n), then w is also an upper bound of (v_n). But v is the least upper bound of (v_n), so $v \preceq w$. This proves that v is the least upper bound of (u_n).

Solution to Exercise A.1.17. Let (f_n) be a sequence in $ib\mathsf{X}$ such that $f_n \to f$ for some $f \in b\mathsf{X}$. The function f is increasing because, for $x, x' \in \mathsf{X}$ with $x \preceq x'$, we have $f_n(x) \leqslant f_n(x')$ for all n, and hence, taking the limit, $f(x) \leqslant f(x')$.

Solution to Exercise A.1.18. Part (i) follows easily from a simple induction argument. Regarding part (ii), let $S, T \in \mathcal{S}$ obey $S \preceq T$. We claim that $S^k \preceq T^k$ holds for all $k \in \mathbb{N}$. Clearly it holds for $k = 1$. If it also holds at $k - 1$, then, for any $u \in V$, we have $S^k u = SS^{k-1}u \leqslant ST^{k-1}u \leqslant TT^{k-1}u = T^k u$, where we used the induction hypothesis, the order preserving property of S and the assumption that $S \preceq T$.

Solution to Exercise A.1.19. By the Neumann series lemma, S has a unique fixed point in $\mathbb{R}^k$ given by $\bar{v} := (I - A)^{-1} r$. Fix $v \in \mathbb{R}^k$ with $v \leqslant Sv$. Since S is order preserving, the sequence $(S^n v)$ is increasing. The n-th element of this sequence is $S^n v = A^n v + \sum_{i=0}^{n-1} A^i r$. Since $\rho(A) < 1$, the pointwise limit of this sequence is $\sum_{i=0}^{\infty} A^i r = (I - A)^{-1} r = \bar{v}$. By Lemma A.1.4, $\bar{v}$ is also the supremum of $(S^n v)$. This proves strong upward stability. Reversing the inequalities shows that strong downward stability also holds.

Solution to Exercise A.1.20. Under the stated conditions, $v \preceq Sv$ implies $v \preceq \bar{v}$. Since S^m is also order preserving, applying S^m to both sides of this inequality yields $S^m v \preceq \bar{v}$.

Solution to Exercise A.1.22. The proofs are straightforward. For example, to see that d satisfies the triangle inequality, pick any $u, v, w \in V$. We claim that $d(u, v) \leqslant d(u, w) + d(w, v)$. If $u = v$, this bound is trivial, so suppose they are distinct. We then need to show that $1 \leqslant d(u, w) + d(w, v)$. Suppose to the contrary that $d(u, w) + d(w, v) = 0$. It follows that $u = w$ and $v = w$. But then $u = v$, which is a contradiction.

Solution to Exercise A.1.34. Let K be a nonempty subset of V. Under the discrete metric, any sequence in V taking only distinct values has no convergent subsequence. Hence, if K is infinite, then K is not compact (or even precompact). Conversely, if K is finite, then every sequence has a constant subsequence. So K is compact.

Solution to Exercise A.2.2. Let V, T, u^* and u be as stated and let $u_m := T^m u$ for each $m \in \mathbb{N}$. By assumption, $u_m \to u^*$. Moreover, the sequence (Tu_m) is just u_m with the first element omitted, so $Tu_m \to u^*$ also holds. At the same time, T is continuous at u^*, so $u_m \to u^*$ implies $Tu_m \to u^*$. We have, therefore, exhibited a sequence that converges to both u^* and Tu^*. Since V is Hausdorff, and hence limits are unique, it follows that $Tu^* = u^*$.

Solution to Exercise A.2.6. It suffices to show that S has only one fixed point in V. To this end, let $\bar{u}$ be a fixed point of S in V. Since $\bar{u} = S^k \bar{u} \to u^*$, there exists a sequence in V converging to both $\bar{u}$ and u^*. Since V is Hausdorff, $\bar{u} = u^*$. Hence u^* is the only fixed point of S in V.

Solution to Exercise A.2.9. By Theorem A.2.10, it suffices to show that L^k is a contraction on $(\mathbb{M}^{n\times n}, \|\cdot\|)$ for some $k \in \mathbb{N}$. Iterating with L from arbitrary $\Sigma \in \mathbb{M}^{n\times n}$, we obtain

$$L^k(\Sigma) = A^k\Sigma(A^k)^\top + A^{k-1}M(A^{k-1})^\top + \cdots + M$$

Hence, for any Σ, T in $\mathbb{M}^{n\times n}$, we have

$$\begin{aligned}\|L^k(\Sigma) - L^k(T)\| &= \left\|A^k\Sigma(A^k)^\top - A^kT(A^k)^\top\right\| \\ &= \left\|A^k(\Sigma - T)(A^k)^\top\right\| \\ &\leqslant \|A^k\| \cdot \|\Sigma - T\| \cdot \|(A^k)^\top\|\end{aligned}$$

Transposes don't change norms, so $\|(A^k)^\top\| = \|A^k\|$ and hence $\|L^k(\Sigma) - L^k(T)\| \leqslant \|A^k\|^2\|\Sigma - T\|$. Since $r(A) < 1$, we can find a $k \in \mathbb{N}$ and a constant $\lambda < 1$, both independent of Σ and T, such that $\|L^k(\Sigma) - L^k(T)\| \leqslant \lambda\|\Sigma - T\|$. Then L^k is a contraction on $\mathbb{M}^{n\times n}$, as was to be shown.

Solution to Exercise A.2.10. The result is immediate if $L\sum_{i\geqslant 0} B_i = \sum_{i\geqslant 0} LB_i$ where $B_i := A^iM(A^\top)^i$. This is true by continuity of L (see, e.g., Exercise A.2.8).

Solution to Exercise A.2.11. We treat only the second question, since the first uses a similar (and easier) argument. Fix nonzero $b \in \mathbb{R}^n$. Because $\Sigma^* := \sum_{i\geqslant 0} A^iM(A^\top)^i$, it suffices to show that, for any i, the term

$$f(b) := b^\top A^iM(A^\top)^ib = b^\top A^iM(A^i)\top b = (A^ib)^\top M(A^ib)^\top$$

is positive. Since A is nonsingular, A^i is nonsingular for all i. Given this fact and the assumption that b is nonzero, the product A^ib is nonzero. Because M is positive definite, we see that $f(b) > 0$, as required.

Solution to Exercise A.4.6. It suffices to provide a basis of $\mathbb{M}^{n\times k}$ with nk elements. For this basis, consider $B := \{B_{ij}\}_{(i,k)\in[n]\times[k]}$, where B_{ij} is the $n \times k$ matrix with 1 at row i, column j, and zeros elsewhere. For a any collection of scalars $\{\alpha_{ij}\}$, the sum $S := \sum_i\sum_j \alpha_{ij}B_{ij}$ is the matrix with element α_{ij} at each i, j and hence $S = 0$ implies $\alpha_{ij} = 0$ for all i, j. Hence B is linearly independent. Clearly $\operatorname{span} B = \mathbb{M}^{n\times k}$. Hence B is a basis, as claimed (and $|B| = nk$).

Solution to Exercise A.4.14. To prove that $(\mathbb{R}^n, \|\cdot\|_1)$ is complete, let (x_k) be Cauchy in $(\mathbb{R}^n, \|\cdot\|_1)$, where $\|\cdot\|_1$ is the ℓ_1 norm. Let $x_k^i \in \mathbb{R}$ be the i-th component of x_k. We have

$$|x_j^i - x_k^i| \leqslant \sum_{i=1}^{n} |x_j^i - x_k^i| = \|x_j - x_k\|,$$

from which it easily follows that each component is Cauchy in $\mathbb{R}$. From this and the completeness of $\mathbb{R}$, stated in Theorem A.1.3 on page 152, all of these component sequences converge. From Theorem **??**, componentwise convergence in $\mathbb{R}^n$ implies norm convergence. Hence every Cauchy sequence in $(\mathbb{R}^n, \|\cdot\|_1)$ converges, as was to be shown.

To complete the proof of Theorem A.4.4, we take $(E, \|\cdot\|)$ to be a normed vector space with $\dim E = n$. By Theorem A.4.3, there exists a norm $\|\cdot\|_0$ on E such that $(E, \|\cdot\|_0)$ is isometrically isomorphic to $(\mathbb{R}^n, \|\cdot\|_1)$. Hence, by Exercise A.4.12, the space $(E, \|\cdot\|_0)$ is complete. Exercise A.4.9 now tells us that $(E, \|\cdot\|)$ is also complete.

Solution to Exercise A.4.17. Because $\mathbb{1}$ is an eigenvalue (since $K\mathbb{1} = \beta\mathbb{1}$), the definition of the spectral radius (see (A.20)) implies that $\rho(K) \geqslant \beta$. At the same time, for any $v \in L_1(\varphi)$ with $\|v\| = 1$ we have $\|Kv\| = \beta$, so $\|K\| = \beta$. Hence $\rho(K) \leqslant \|K\| = \beta$. We conclude that $\rho(K) = \beta$.

Solution to Exercise A.5.1. The set $\mathbb{R}^n$ with the pointwise partial order is a lattice but not chain complete.

Solution to Exercise A.5.5. Regarding (i), suppose $u \leqslant 0$ and $v \leqslant 0$. Adding v to both sides of $u \leqslant 0$ gives $u + v \leqslant v \leqslant 0$. Regarding (ii), suppose $u \leqslant v$. Adding $-v$ to both sides gives $u - v \leqslant 0$. Adding $-u$ to both sides completes the proof. Regarding (iii), set $s := (u \vee v) + w$ and $t := (u+w) \vee (v+w)$. On one hand, $s \geqslant u+w$ and $s \geqslant v+w$, so $s \geqslant t$. On the other hand, $t \geqslant u + w$ and $t \geqslant v + w$, so $t - w \geqslant u \vee v$. Hence $t \geqslant u \vee v + w = s$. Regarding (iv), set $s := \alpha(u \vee v)$ and suppose $\alpha > 0$. (The case $\alpha = 0$ is trivial.) We have $u \leqslant s/\alpha$ and $v \leqslant s/\alpha$, so $\alpha u \leqslant s$ and $\alpha v \leqslant s$. Moreover, if $(\alpha u) \vee (\alpha v) \leqslant s'$, then $u \leqslant s'/\alpha$ and $v \leqslant s'/\alpha$, so $s = \alpha(u \vee v) \leqslant s'$. In other words, s is the least upper bound of $\{\alpha u, \alpha v\}$.

Solution to Exercise A.5.6. Suppose $u_n \uparrow 0$ and $v_n \uparrow 0$. Let U be the set of upper bounds of $(u_n + v_n)$. Since $u_n \leqslant 0$ and $v_n \leqslant 0$ for all n we see that $0 \in U$. Fixing any $w \in U$, monotonicity of the sequences gives $u_n + v_m \leqslant w$ for all n, m, from which we obtain $u_n \leqslant w - v_m$ for all n, m and hence $0 \leqslant w - v_m$ (because 0 is the supremum of (u_n)). Rearranging gives $v_m \leqslant w$ for all m and hence $0 \leqslant w$. This proves that 0 is a least element of U, so 0 is the supremum of $(u_n + v_n)$.

Regarding the second claim, suppose $u_n \uparrow u$ and fix $b \in E$. Let U be the set of upper bounds of $(u_n + b)$. Since $u_n \leqslant u$ for all n we see that $u + b \in U$. If $w \in U$, then $u_n \leqslant w - b$

for all n, so $u \leqslant w - b$, or $u + b \leqslant w$. This proves that $u + b$ is a least element of U, so $u + b$ is the supremum of $(u_n + b)$.

Solution to Exercise A.5.11. We prove that (iii) implies (i) and leave other details to the reader. Let A be a positive linear operator mapping $E \to E$. Fix $(v_n) \subset E$ with $v_n \uparrow v \in E$. Using results from Lemma A.5.9 we have $v - v_n \downarrow 0$ and hence, by (iii), $A(v - v_n) \downarrow 0$. Using linearity and results from Lemma A.5.9 again, we obtain $Av_n \uparrow Av$. Hence (i) holds.

Solution to Exercise A.5.14. The set of differentiable functions in $bc(0,1)$, the bounded continuous functions on $(0,1)$, is a linear subspace of $bc(0,1)$ but not a lattice – and hence not a Riesz space.

Solution to Exercise A.5.15. Let K be as described and fix $u \in E$. We use linearity, positivity, and Lemma A.5.11 to obtain

$$|Ku| = |K(u^+ - u^-)| = |Ku^+ - Ku^-| \leqslant |Ku^+| + |Ku^-| = Ku^+ + Ku^- = K(u^+ + u^-) = K|u|.$$

Solution to Exercise A.5.16. Let E, V and S be as stated. Fix $(v_n) \subset V$ with $v_n \uparrow v \in V$. We claim that $S\,v_n \uparrow S\,v$. We have $0 \leqslant S\,v - S\,v_n \leqslant K|v - v_n| = K(v - v_n) \downarrow 0$, where $\downarrow 0$ is by order continuity of K (see Exercise A.5.11). Hence $S\,v - S\,v_n \downarrow 0$. By Lemma A.5.9, this gives $S\,v_n \uparrow S\,v$.

Solution to Exercise A.5.18. Let V, K and N be as stated. Positivity and linearity follow from basic properties of the integral (see §A.3.2.3). Regarding order continuity, fix $v_n \uparrow v$. We claim that $Kv_n \uparrow Kv$. Since V is closed under pointwise suprema, Lemma A.1.4 implies that v_n increases to v pointwise. Moreover, K is positive and therefore order preserving, so Kv_n is increasing and bounded above by Kv. As V is closed under pointwise suprema, it follows that Kv_n increases pointwise to the pointwise supremum $\sup_n Kv_n$. By the monotone convergence theorem, the limit is Kv. Applying Lemma A.1.4 again, we get $Kv_n \uparrow Kv$.

Solution to Exercise A.5.19. Assume the conditions of the proposition and let u_T be the unique fixed point of T. Let u_S be any fixed point of S. Since $S \preceq T$, we have $u_S = Su_S \preceq Tu_S$. Applying T to both sides of this inequality and using the order preserving property of T and transitivity of $\preceq$ gives $u_S \preceq T^2 u_S$. Continuing in this fashion yields $u_S \preceq T^k u_S$ for all $k \in \mathbb{N}$. Taking the limit in k and using the fact that $\preceq$ is closed under limits gives $u_S \preceq u_T$.

Solution to Exercise A.5.20. Assume the conditions of the proposition and let u_T be the unique fixed point of T in V. Let u_S be any fixed point of S. Since $S \preceq T$, we have $u_S = Su_S \preceq Tu_S$. Applying T to both sides of this inequality and using the order-preserving property of T and transitivity of $\preceq$ gives $u_S \preceq T^2 u_S$. Continuing in this fashion yields $u_S \preceq T^k u_S$ for all $k \in \mathbb{N}$. Taking the limit in k and using the fact that $\preceq$ is closed gives $u_S \preceq u_T$.

Solution to Exercise A.5.22. If $v \in bm\mathsf{X}$, then $|v| \leqslant M$ for some $M \in \mathbb{N}$. But then $|v|/\ell \leqslant M$, since $\ell \leqslant 1$. Hence $v \in\subset b_\ell m\mathsf{X}$. The proof of the second case is similar.

Solution to Exercise A.5.23. The only nontrivial part of the proof is the triangle inequality. This is still quite straightforward: If $u, v \in b_\ell\mathsf{X}$, then, using the triangle inequality in $\mathbb{R}$,

$$\left|\frac{u}{\ell} + \frac{v}{\ell}\right| \leqslant \left|\frac{u}{\ell}\right| + \left|\frac{v}{\ell}\right| \leqslant \|u\| + \|v\|.$$

Taking the supremum on the left-hand side completes the proof.

Solution to Exercise A.5.24. Pick any $x \in \mathsf{X}$. We have $|w_n(x)/\ell(x) - w(x)/\ell(x)| \leqslant \|w_n - w\| \to 0$ and hence $|w_n(x) - w(x)| \leqslant \|w_n - w\|\ell(x) \to 0$.

Solution to Exercise A.5.25. Let s be the supremum of $\|Au\|$ over all $u \in E_+$ with $\|u\| = 1$. Clearly $s \leqslant \|A\|$. To see that the reverse inequality holds, fix $u \in E$ with $\|u\| = 1$ and let $v = |u| \in E_+$. We claim that $\|Au\| \leqslant \|Av\|$, which suffices for $\|A\| \leqslant s$. To verify this claim, we use $|Au| \leqslant A|u| = Av = |Av|$ and the lattice norm property of $\|\cdot\|$ to obtain $\|Au\| \leqslant \|Av\|$.

Solution to Exercise A.5.27. The claim is that $\preceq_{\mathrm{F}}$ yields a partial order on $\mathscr{D}(\mathsf{X})$. Reflexivity and transitivity are immediate from the definition. Asymmetry follows from the characterization in (A.29).

Solution to Exercise A.5.29. If we take Y and Z to be independent and standard normal, then $\mathbb{E}[Z \mid Y] = \mathbb{E}Z = 0$, $Y \stackrel{d}{=} \varphi$ and, since sums of independent Gaussians and Gaussian, $Y + Z \stackrel{d}{=} N(0, 2) = \psi$. Hence ψ is a mean-preserving spread of φ, as claimed.

Solution to Exercise A.5.30. Let φ be a mean-preserving spread of ψ. Then there exists a random pair (Y, Z) such that

$$Y \stackrel{d}{=} \psi, \quad Y + Z \stackrel{d}{=} \varphi \quad \text{and} \quad \mathbb{E}[Z \mid Y] = 0.$$

Fixing arbitrary concave $u \in ib\mathbb{R}$ and applying Jensen's inequality,

$$\mathbb{E}\, u(Y+Z) = \mathbb{E}\,\mathbb{E}[u(Y+Z)\,|\,Y] \leqslant \mathbb{E}\, u(\mathbb{E}[Y+Z\,|\,Y]) = \mathbb{E}\, u(Y).$$

Therefore $\int u(x)\varphi(\mathrm{d}x) = \mathbb{E}\, u(Y+Z) \leqslant \mathbb{E} u(Y) = \int u(x)\psi(\mathrm{d}x)$.

Bibliography

Alessandria, G., Arkolakis, C., and Ruhl, K. J. (2021). Firm dynamics and trade. *Annual Review of Economics*, 13:253–280.

Aliprantis, C. D. and Border, Kim, C. (2006). *Infinite dimensional analysis: a hitchhiker's guide*. Springer-Verlag, New York, 3 edition.

Aliprantis, C. D. and Burkinshaw, O. (1998). *Principles of real analysis*. Academic Press, 3 edition.

Aliprantis, C. D. and Burkinshaw, O. (2006). *Positive operators*, volume 119. Springer Science & Business Media.

Bartle, R. G. and Sherbert, D. R. (2011). *Introduction to real analysis*. Hoboken, NJ: Wiley, 4 edition.

Bátkai, A., Fijavž, M. K., and Rhandi, A. (2017). Positive operator semigroups. *Operator Theory: advances and applications*, 257.

Bäuerle, N. and Jaśkiewicz, A. (2018). Stochastic optimal growth model with risk sensitive preferences. *Journal of Economic Theory*, 173:181–200.

Bäuerle, N., Jaśkiewicz, A., and Nowak, A. S. (2021). Stochastic dynamic programming with non-linear discounting. *Applied Mathematics & Optimization*, 84(3):2819–2848.

Bäuerle, N. and Rieder, U. (2011). *Markov decision processes with applications to finance*. Springer Science & Business Media.

Bellman, R. (1957). *Dynamic programming*. American Association for the Advancement of Science.

Benhabib, J., Bisin, A., and Luo, M. (2015). Wealth distribution and social mobility in the us: A quantitative approach. Technical report, National Bureau of Economic Research.

Bertsekas, D. (2012). *Dynamic programming and optimal control*, volume 1. Athena Scientific.

Bertsekas, D. (2021). *Rollout, policy iteration, and distributed reinforcement learning*. Athena Scientific.

Bertsekas, D. P. (1977). Monotone mappings with application in dynamic programming. *SIAM Journal on Control and Optimization*, 15(3):438–464.

Bertsekas, D. P. (2017). Regular policies in abstract dynamic programming. *SIAM Journal on Optimization*, 27(3):1694–1727.

Bertsekas, D. P. (2022). *Abstract dynamic programming*. Athena Scientific, 3 edition.

Borovička, J. and Stachurski, J. (2020). Necessary and sufficient conditions for existence and uniqueness of recursive utilities. *The Journal of Finance*.

Çınlar, E. (2011). *Probability and stochastics*, volume 261. Springer Science & Business Media.

Davey, B. A. and Priestley, H. A. (2002). *Introduction to lattices and order*. Cambridge University Press.

de Castro, L. and Galvao, A. F. (2019). Dynamic quantile models of rational behavior. *Econometrica*, 87(6):1893–1939.

de Castro, L., Galvao, A. F., and Nunes, D. (2022). Dynamic economics with quantile preferences. *Available at SSRN 4108230*.

De Nardi, M., Fella, G., and Paz-Pardo, G. (2020). Nonlinear household earnings dynamics, self-insurance, and welfare. *Journal of the European Economic Association*, 18(2):890–926.

Denardo, E. V. (1967). Contraction mappings in the theory underlying dynamic programming. *Siam Review*, 9(2):165–177.

Dudley, R. M. (2002). *Real analysis and probability*, volume 74. Cambridge University Press.

Engel, K.-J. and Nagel, R. (2006). *A short course on operator semigroups*. Springer Science & Business Media.

Fei, Y., Yang, Z., Chen, Y., and Wang, Z. (2021). Exponential Bellman equation and improved regret bounds for risk-sensitive reinforcement learning. *Advances in Neural Information Processing Systems*, 34:20436–20446.

Goebel, K. and Kirk, W. A. (1990). *Topics in metric fixed point theory*. Number 28. Cambridge university press.

Green, L., Myerson, J., and McFadden, E. (1997). Rate of temporal discounting decreases with amount of reward. *Memory & cognition*, 25:715–723.

Gu, Y., Subramaniyan, A., Dunn, T., Khadem, A., Chen, K.-Y., Paul, S., Vasimuddin, M., Misra, S., Blaauw, D., and Narayanasamy, S. (2023). Gendp: A framework of dynamic programming acceleration for genome sequencing analysis. In *Proceedings of the 50th Annual International Symposium on Computer Architecture*, pages 1–15.

Guo, X. and Hernández-Lerma, O. (2009). *Continuous-time Markov decision processes*. Springer.

Hansen, L. P. and Scheinkman, J. A. (2012). Recursive utility in a Markov environment with stochastic growth. *Proceedings of the National Academy of Sciences*, 109(30):11967–11972.

Hernández-Lerma, O. and Lasserre, J. B. (2012). *Discrete-time Markov control processes: basic optimality criteria*, volume 30. Springer Science & Business Media.

Hopenhayn, H. A. and Prescott, E. C. (1992). Stochastic monotonicity and stationary distributions for dynamic economies. *Econometrica*, pages 1387–1406.

Jahn, J. (2020). *Introduction to the theory of nonlinear optimization*. Springer Nature, 4 edition.

Jänich, K. (1994). Linear algebra. undergraduate texts in mathematics. *Springer-Verlag, New York*, 7:8.

Jaśkiewicz, A., Matkowski, J., and Nowak, A. S. (2014). On variable discounting in dynamic programming: applications to resource extraction and other economic models. *Annals of Operations Research*, 220:263–278.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica*, pages 649–670.

Kechris, A. (2012). *Classical descriptive set theory*, volume 156. Springer Science & Business Media.

Kikuchi, T., Nishimura, K., Stachurski, J., and Zhang, J. (2021). Coase meets Bellman: Dynamic programming for production networks. *Journal of Economic Theory*, 196:105287.

Kochenderfer, M. J., Wheeler, T. A., and Wray, K. H. (2022). *Algorithms for decision making*. The MIT Press.

Kreyszig, E. (1978). *Introductory functional analysis with applications*, volume 1. wiley New York.

Kristensen, D., Mogensen, P. K., Moon, J. M., and Schjerning, B. (2021). Solving dynamic discrete choice models using smoothing and sieve methods. *Journal of Econometrics*, 223(2):328–360.

Li, X. and Xie, L. (2021). Online abstract dynamic programming with contractive models.

Ljungqvist, L. and Sargent, T. J. (2018). *Recursive macroeconomic theory*. MIT press, 4 edition.

Lu, J., Luo, Y., Saito, K., and Xin, Y. (2024). Did harold zuercher have time-separable preferences?

McCall, J. J. (1970). Economics of Information and Job Search. *The Quarterly Journal of Economics*, 84(1):113–126.

Meyer-Nieberg, P. (2012). *Banach lattices*. Springer Science & Business Media.

Meyn, S. P. and Tweedie, R. L. (2009). *Markov chains and stochastic stability*. Cambridge University Press.

Puterman, M. L. (2005). *Markov decision processes: discrete stochastic dynamic programming*. Wiley Interscience.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of Harold Zurcher. *Econometrica*, pages 999–1033.

Rust, J. (1994). Structural estimation of Markov decision processes. *Handbook of Econometrics*, 4:3081–3143.

Rust, J. (1996). Numerical dynamic programming in economics. *Handbook of computational economics*, 1:619–729.

Sargent, T. J. and Stachurski, J. (2025a). Dynamic programming: Finite states.

Sargent, T. J. and Stachurski, J. (2025b). Dynamic programs on partially ordered sets. *SIAM Journal on Optimization and Control*, page in press.

Schorfheide, F., Song, D., and Yaron, A. (2018). Identifying long-run risks: A Bayesian mixed-frequency approach. *Econometrica*, 86(2):617–654.

Stachurski, J. (2022). *Economic dynamics: theory and computation*. MIT Press, 2 edition.

Stachurski, J. and Toda, A. A. (2019). An impossibility theorem for wealth in heterogeneous-agent models with limited heterogeneity. *Journal of Economic Theory*, 182:1–24.

Sterk, V., Sedláček, P., and Pugsley, B. (2021). The nature of firm growth. *American Economic Review*, 111(2):547–579.

Stokey, N. L. and Lucas, R. E. (1989). *Recursive methods in dynamic economics*. Harvard University Press.

Szepesvari, C. (1998). Non-markovian policies in sequential decision problems. *Acta Cybernetica*, 13(3):305–318.

Tsitsiklis, J. N. (1994). Asynchronous stochastic approximation and q-learning. *Machine learning*, 16:185–202.

Verdu, S. and Poor, H. V. (1987). Abstract dynamic programming models under commutativity conditions. *SIAM Journal on Control and Optimization*, 25(4):990–1006.

Watkins, C. J. C. H. (1989). Learning from delayed rewards. Technical report, PhD Thesis, King's College, Cambridge United Kingdom.

Zaanen, A. C. (2012). *Introduction to operator theory in Riesz spaces*. Springer.

Zhang, Z. (2012). *Variational, topological, and partial order methods with their applications*, volume 29. Springer.