Markov Decision Processes in Economic Dynamics: A Concise Treatment via the McCall Job Search Problem

Abstract

This note develops the theory of infinite horizon discounted Markov decision processes for students who already know discrete time Markov chains. The presentation is motivated and illustrated by the McCall job search model. The core results on existence and uniqueness of the value function, the principle of optimality, and the policy improvement theorem are stated and proved under standard conditions. The McCall problem is solved by direct application of these theorems. Further classic examples and exercises are provided with references to the literature.

1 Introduction and Motivation

Many dynamic economic problems reduce to recursive choice under Markovian uncertainty. The McCall job search model offers a minimal and transparent instance. An unemployed worker observes a sequence of wage offers over time, chooses when to accept, and discounts the future. The search environment is i.i.d. across time, hence Markov relative to the current information set. The optimal strategy has a reservation wage characterization that follows from the general theorems of Markov decision processes. The goal is to formulate the abstract model and give proofs of the main results that deliver the McCall solution as a corollary.

2 The MDP Framework

Let $(\mathsf{X}, \mathcal{X})$ be a measurable state space and $(\mathsf{A}, \mathcal{A})$ a measurable action space. For each $x \in \mathsf{X}$ a nonempty feasible correspondence $\Gamma(x) \subset \mathsf{A}$ is given. A one period payoff function $r : \mathsf{X} \times \mathsf{A} \to \mathbb{R}$ is measurable and bounded. The dynamics are governed by a stochastic kernel $P(\cdot \mid x, a)$ on $(\mathsf{X}, \mathcal{X})$ that yields the distribution of next period state X' given current (x, a). The discount factor $\beta \in (0, 1)$ is fixed.

A policy is a measurable map $\pi : \mathsf{X} \to \mathsf{A}$ with $\pi(x) \in \Gamma(x)$ for all x. Let Π denote the set of stationary policies. Starting from $x_0 = x$ and a policy π, the induced controlled process $\{X_t\}_{t \geq 0}$ is a time homogeneous Markov chain with transition kernel $P_\pi(B \mid x) := P(B \mid x, \pi(x))$ and realized actions $A_t = \pi(X_t)$. The value of π from state x is

$$V_\pi(x) := \mathbb{E}_x^\pi \left[\sum_{t=0}^{\infty} \beta^t r(X_t, \pi(X_t)) \right],$$

well defined when r is bounded. The optimal value function is $V^*(x) := \sup_{\pi \in \Pi} V_\pi(x)$.

Definition 2.1 (Bellman Operator)**.** Define T on the Banach space $(\mathcal{B}, \|\cdot\|_\infty)$ of bounded measurable functions $v : \mathsf{X} \to \mathbb{R}$ by

$$(Tv)(x) := \sup_{a \in \Gamma(x)} \left\{ r(x,a) + \beta \int v(x')\, P(dx' \mid x, a) \right\}.$$

Theorem 2.2 (Contraction and Fixed Point)**.** *Assume r is bounded and $\beta \in (0,1)$. Then T is a β-contraction on $(\mathcal{B}, \|\cdot\|_\infty)$. Hence there exists a unique $V^* \in \mathcal{B}$ with $V^* = TV^*$, and for any $v_0 \in \mathcal{B}$ the sequence $v_{n+1} := Tv_n$ converges to V^* in sup norm at geometric rate β.*

Proof. Let $v, w \in \mathcal{B}$ and $x \in \mathsf{X}$. For any $a \in \Gamma(x)$ one has

$$r(x,a) + \beta \int v\, dP(\cdot \mid x,a) \le r(x,a) + \beta \int w\, dP(\cdot \mid x,a) + \beta \|v - w\|_\infty.$$

Taking suprema over a on the left and right sides yields $(Tv)(x) \le (Tw)(x) + \beta\|v - w\|_\infty$. The symmetric inequality gives $\|(Tv) - (Tw)\|_\infty \le \beta \|v - w\|_\infty$. Banach's fixed point theorem implies existence and uniqueness of a fixed point and global convergence of value iteration. □

Theorem 2.3 (Principle of Optimality)**.** *Let V^* be the unique fixed point of T. For each x there exists an action $a^*(x) \in \Gamma(x)$ that attains the supremum in the Bellman equation*

$$V^*(x) = \sup_{a \in \Gamma(x)} \left\{ r(x,a) + \beta \int V^*(x')\, P(dx' \mid x, a) \right\}.$$

Any stationary policy π^* that selects such maximizers is optimal. Moreover $V_{\pi^*} = V^*$.

Proof. Fix x. Let a_n be a sequence in $\Gamma(x)$ that ε_n-maximizes the right-hand side with $\varepsilon_n \downarrow 0$. Boundedness and the measurability structure allow measurable selection of maximizers under mild regularity, for example compactness of each $\Gamma(x)$ and upper semicontinuity of $a \mapsto r(x,a) + \beta \int V^*(x') P(dx' \mid x, a)$. Under these conditions choose a measurable maximizer $a^*(x)$. For any such selector π^* define $Q(x,a) := r(x,a) + \beta \int V^*(x') P(dx' \mid x,a)$. Then $V^*(x) = Q(x, \pi^*(x))$. The tower property and induction yield

$$V^*(x) = \mathbb{E}_x^{\pi^*} \left[\sum_{t=0}^{n} \beta^t r(X_t, \pi^*(X_t)) + \beta^{n+1} V^*(X_{n+1}) \right].$$

Letting $n \to \infty$ and using boundedness shows $V^*(x) = V_{\pi^*}(x)$. Optimality follows since V^* dominates V_π for all π by the definition of T and monotonicity. □

Theorem 2.4 (Policy Improvement)**.** *Let π be any stationary policy and define the Q-function $Q_\pi(x,a) := r(x,a) + \beta \int V_\pi(x') P(dx' \mid x, a)$. Define a new policy $\pi'(x) \in \arg\max_{a \in \Gamma(x)} Q_\pi(x,a)$. Then $V_{\pi'} \ge V_\pi$ pointwise. Moreover equality holds if and only if π is optimal.*

Proof. By definition $T_\pi v := r(\cdot, \pi(\cdot)) + \beta P_\pi v$ is a β-contraction on $\mathcal{B}$ with fixed point V_π, and $Tv \geq T_\pi v$ for all v. Since π' maximizes Q_π, one has $T_{\pi'} V_\pi = TV_\pi \geq T_\pi V_\pi = V_\pi$. Monotonicity and contraction of $T_{\pi'}$ yield $V_{\pi'} = \lim_{n\to\infty} T_{\pi'}^n V_\pi \geq V_\pi$. If $V_{\pi'} = V_\pi$, then $TV_\pi = T_{\pi'} V_\pi = V_\pi$, hence $V_\pi = V^*$ and π is optimal. □

Remark 2.5. Compactness and continuity assumptions are standard to guarantee measurable selectors. In finite state and action models these assumptions hold trivially. In Borel models one can use measurable selection theorems as in Bertsekas and Shreve.

3 The McCall Job Search Problem

Time is discrete. An unemployed worker receives an i.i.d. sequence of wage offers $\{W_t\}$ drawn from a known distribution F on $[0, \infty)$. Each period the worker either accepts and receives the offered wage forever, or rejects and remains unemployed. The flow value of unemployment is $b \in \mathbb{R}$ and the discount factor is $\beta \in (0, 1)$. Acceptance yields a permanent job with constant wage and no separation. The objective is to maximize expected discounted income.

Define the state space as $\mathsf{X} = [0, \infty) \cup \{\varnothing\}$, where $\varnothing$ denotes the unemployed state in which an offer w is observed as an ancillary signal. It is convenient to embed the signal in the state and write the relevant Bellman equations directly. Let V_U denote the value of being unemployed before seeing the current offer, $V_U(w)$ the value at the decision point after observing w, and $V_E(w)$ the value of employment at wage w. The one period payoffs are b in unemployment and w in employment. The transition is degenerate in employment and i.i.d. in unemployment.

The recursive system is

$$V_E(w) = \frac{w}{1-\beta}, \qquad V_U = \beta \int V_U(w)\, F(dw),$$

$$V_U(w) = \max\left\{\frac{w}{1-\beta},\ b + \beta V_U\right\}.$$

The first identity follows from the geometric sum under constant wage. The second identity is the law of iterated expectations for the pre-offer value. The third identity is the Bellman equation at the post-offer decision.

Proposition 3.1 (Reservation Wage)**.** *There exists a reservation wage $\overline{w}$ such that it is optimal to accept if and only if $w \geq \overline{w}$. It is characterized by*

$$\frac{\overline{w}}{1-\beta} = b + \beta V_U,$$

and $\overline{w}$ solves the scalar fixed point equation

$$\frac{\overline{w}}{1-\beta} = b + \beta \int \max\left\{\frac{w}{1-\beta},\ \frac{\overline{w}}{1-\beta}\right\} F(dw).$$

Proof. Define $A(w) := \frac{w}{1-\beta}$ and $R := b + \beta V_U$. The function A is increasing in w while R is a constant. The decision $\max\{A(w), R\}$ has a monotone structure. Since A is continuous and strictly increasing, there exists a unique threshold $\overline{w}$ at which $A(\overline{w}) = R$. Accept if and only if $A(w) \geq R$, which is equivalent to $w \geq \overline{w}$. Substituting $R = \frac{\overline{w}}{1-\beta}$ into the pre-offer value identity gives

$$V_U = \beta \int \max\left\{\frac{w}{1-\beta}, \frac{\overline{w}}{1-\beta}\right\} F(dw).$$

Combining with $R = b + \beta V_U$ yields the stated scalar fixed point equation. □

Theorem 3.2 (Existence and Uniqueness of the Reservation Wage)**.** *Assume* $\mathbb{E}[W] < \infty$*. There exists a unique* $\overline{w} \in [0, \infty)$ *satisfying the fixed point equation in Proposition 3.1. The optimal stationary policy that accepts if and only if* $w \geq \overline{w}$ *attains the optimal value function* V^**.*

Proof. Define $\Phi(\omega) := b + \beta \int \max\left\{\frac{w}{1-\beta}, \frac{\omega}{1-\beta}\right\} F(dw)$ on $[0, \infty)$. The function Φ is increasing and β-Lipschitz in the sup norm metric inherited from the scalar space since

$$|\Phi(\omega) - \Phi(\omega')| \leq \beta \int \left|\max\{A(w), A(\omega)\} - \max\{A(w), A(\omega')\}\right| F(dw) \leq \beta \frac{|\omega - \omega'|}{1-\beta}.$$

Hence the map $\omega \mapsto (1-\beta)\Phi(\omega)$ is a contraction on $[0, \infty)$ with modulus β. It has a unique fixed point $\overline{w}$. By Theorem 2.3 the policy that selects a maximizer of the Bellman operator is optimal. The reservation policy is such a maximizer by Proposition 3.1, hence it is optimal and yields V^*. □

Remark 3.3. Comparative statics are immediate. The reservation wage is increasing in b and in first order stochastic dominance shifts of F. Differentiability requires additional conditions. Under a log-concavity assumption on $1 - F$ the acceptance probability is increasing in $\overline{w}$, which simplifies welfare and duration calculations.

4 Main Theorems for Finite and Borel MDPs

For finite state and action sets the theory is purely algebraic.

Theorem 4.1 (Finite MDPs)**.** *If* X *and* A *are finite, then the Bellman operator* T *is a contraction on* $\mathbb{R}^{|\mathsf{X}|}$ *with modulus* β*, the fixed point* V^* *is unique, and there exists an optimal stationary deterministic policy. Value iteration and policy iteration converge globally.*

Proof. Boundedness is trivial and the contraction property follows from Theorem 2.2. Existence of a stationary deterministic optimal policy follows because the maximization defining T is over a finite set at each state. Policy iteration is a monotone sequence of policies with strictly improving values until optimality, which occurs in finitely many steps because the policy set is finite. □

In general Borel models compactness and continuity ensure existence of maximizers and measurability of selectors.

Theorem 4.2 (Borel MDPs with Discounting)**.** *Suppose* X *and* A *are Borel spaces, the feasible correspondence* Γ *has nonempty compact values and measurable graph, the function* $(x, a) \mapsto r(x, a)$ *is bounded and continuous, and* $(x, a) \mapsto \int v(x')P(dx' \mid x, a)$ *is continuous for each bounded continuous* v*. Then* T *maps bounded continuous functions into themselves, admits a unique fixed point* V^**, and there exists a stationary measurable optimal policy.*

Proof. Contraction is as before. Continuity and compactness yield upper semicontinuity of the objective and existence of maximizers. The measurable maximum theorem gives a measurable selector $x \mapsto a^*(x)$. The rest follows from Theorem 2.3. □

5 Applying the General Theory to McCall

The McCall problem is an optimal stopping problem. Embed it in the MDP framework by taking $\mathsf{X} = [0, \infty)$ as the current offer, $\mathsf{A} = \{\text{accept}, \text{ reject}\}$, $r(w, \text{accept}) = w$, $r(w, \text{reject}) = b$, $P(\cdot \mid w, \text{accept})$ degenerate at an absorbing employment state with value $w/(1 - \beta)$ which can be folded into current payoff, and $P(\cdot \mid w, \text{reject}) = F$. Boundedness holds if $\mathbb{E}[W] < \infty$ and b is finite. Theorem 4.2 gives existence of V^* and an optimal stationary policy. The policy improvement theorem ensures that the threshold policy dominates any other stationary rule, and the reservation wage computed above is the unique fixed point that supports optimality.

6 An Extension after McCalls: Skill Depreciation, Turbulence, and State-Dependent Reservation Wages

This section develops the Ljungqvist–Sargent extension of the McCall job search model, as presented in Sargent's lecture notes and subsequent joint work. The modification introduces a skill state that can depreciate during unemployment. A turbulence parameter governs the likelihood of losing high skills when jobless, thereby lowering future earning potential. This mechanism generates state-dependent reservation wages and helps explain persistently high unemployment following adverse turbulence shocks. For formal treatments see Ljungqvist and Sargent (1998, 2010).

6.1 Environment

Time is discrete. A worker is risk neutral with discount factor $\beta \in (0, 1)$. In each period an unemployed worker draws a wage offer $W \sim F$ independently across time and independent of the skill state defined below. If the worker accepts an offer w, she receives the constant wage w forever. There is no separation once employed in the baseline computation; separations and on-the-job wage resets are discussed in a remark.

Skills take two levels $\mathsf{K} = \{H, L\}$ with $H > L > 0$. The skill state evolves only while unemployed according to a Markov transition that depends on a turbulence parameter $\tau \in [0, 1]$:

$$\Pr\{K' = L \mid K = H, \text{ unemployed}\} = \tau, \qquad \Pr\{K' = L \mid K = L, \text{ unemployed}\} = 1.$$

Thus H skills erode to L at rate τ when unemployed, while L persists. When employed, the wage is fixed at the accepted w and the skill process is irrelevant for payoffs in the baseline. The unemployed flow payoff is $b \in \mathbb{R}$. Write $U(k)$ for the pre-offer value of unemployment at skill $k \in \{H, L\}$ and $E(w)$ for the value of accepting wage w.

6.2 Bellman Problem and Principle of Optimality

At an unemployment decision node with current skill k and observed offer w, the post-offer value is

$$U^{\mathrm{obs}}(k, w) = \max\left\{E(w),\ b + \beta\,\mathbb{E}[U(K') \mid k, \text{unemp.}]\right\}.$$

With a constant wage forever once employed,

$$E(w) = \frac{w}{1-\beta}.$$

The pre-offer unemployment value averages over the offer distribution,

$$U(k) = \int U^{\mathrm{obs}}(k, w)\, F(dw).$$

Define the continuation value in unemployment

$$R(k) := b + \beta\,\mathbb{E}[U(K') \mid k, \text{unemp.}] = \begin{cases} b + \beta\big[(1-\tau)U(H) + \tau U(L)\big], & k = H, \\ b + \beta\, U(L), & k = L. \end{cases}$$

The Bellman operator on $\mathbb{R}^2$ given by $U \mapsto \tilde{U}$ with

$$\tilde{U}(k) = \int \max\left\{\frac{w}{1-\beta},\ R(k)\right\} F(dw)$$

is a β-contraction in the sup norm by the standard argument, so a unique fixed point $U^* = (U^*(H), U^*(L))$ exists and iteration converges to it. The principle of optimality implies that any measurable selector that attains the pointwise maximum is optimal. These statements are direct applications of Theorems 2.2 and 2.3 from the main text.

6.3 State-Dependent Reservation Wages and Closed-Form Equations

Monotonicity of $w \mapsto w/(1-\beta)$ implies a threshold rule at each k.

Proposition 6.1 (Two reservation wages)**.** *For $k \in \{H, L\}$ there exists a reservation wage $\overline{w}_k$ such that accepting w is optimal if and only if $w \geq \overline{w}_k$. They satisfy*

$$\frac{\overline{w}_k}{1-\beta} = R(k),$$

and the fixed point equations

$$U(k) = \int_0^{\overline{w}_k} R(k)\, F(dw) + \int_{\overline{w}_k}^{\infty} \frac{w}{1-\beta}\, F(dw), \tag{1}$$

$$R(k) = b + \beta\, \mathbb{E}[U(K') \mid k, \mathit{unemp.}], \tag{2}$$

with $R(k) = (1-\beta)^{-1}\overline{w}_k$.

Proof. The function $w \mapsto \max\{w/(1-\beta), R(k)\}$ is increasing with a unique crossing at $w = (1-\beta)R(k)$. Splitting the integral at that crossing gives (1). The definition of $R(k)$ is (2). The fixed point relation between $\overline{w}_k$ and $R(k)$ follows from $E(\overline{w}_k) = R(k)$. □

It is convenient to express $U(k)$ directly in terms of $\overline{w}_k$. Using $R(k) = (1-\beta)^{-1}\overline{w}_k$ in (1) yields

$$U(k) = \frac{1}{1-\beta}\left[\overline{w}_k\, F(\overline{w}_k) + \int_{\overline{w}_k}^{\infty} w\, F(dw)\right]. \tag{3}$$

Substituting (3) into (2) and multiplying both sides by $(1-\beta)$ gives a pair of scalar equations for $(\overline{w}_H, \overline{w}_L)$:

$$\overline{w}_H = (1-\beta)b + \beta\Big[(1-\tau)\,\Phi(\overline{w}_H) + \tau\,\Phi(\overline{w}_L)\Big], \tag{4}$$

$$\overline{w}_L = (1-\beta)b + \beta\,\Phi(\overline{w}_L), \tag{5}$$

where the auxiliary map Φ collects the bracket in (3):

$$\Phi(c) := c\,F(c) + \int_c^{\infty} w\, F(dw).$$

Equations (4)–(5) pin down the two reservation wages. Note that (5) decouples and can be solved first. Then (4) follows as a convex combination in $\overline{w}_H$ and $\overline{w}_L$.

Theorem 6.2 (Existence, uniqueness, and comparative statics)**.** *Suppose* $\mathbb{E}[W] < \infty$*. Then* (4)–(5) *have a unique solution* $(\overline{w}_H, \overline{w}_L) \in [0,\infty)^2$*. Moreover* $\overline{w}_L$ *is increasing in* b *and* β*, and* $\overline{w}_H$ *is increasing in* b*,* β*, and* $\overline{w}_L$*, and decreasing in* τ*. If* F *shifts in the sense of first-order stochastic dominance, both reservation wages increase.*

Proof. For any $c \geq 0$, Φ is increasing and 1–Lipschitz: $0 \leq \Phi'(c) = F(c) \leq 1$ almost everywhere. Define $T_L(c) = (1-\beta)b + \beta\,\Phi(c)$ on $[0,\infty)$. Then $|T_L(c) - T_L(c')| \leq \beta|c-c'|$, so T_L is a contraction with modulus β. It has a unique fixed point $\overline{w}_L$. Define $T_H(c) = (1-\beta)b + \beta[(1-\tau)\,\Phi(c) + \tau\,\Phi(\overline{w}_L)]$. Since $|T_H(c) - T_H(c')| \leq \beta(1-\tau)|c-c'|$, T_H is a contraction with modulus $\beta(1-\tau) < 1$. This yields a unique fixed point $\overline{w}_H$. Monotone comparative statics follow from monotonicity of T_L and T_H in their arguments and parameters. In particular T_H is decreasing in τ, hence the fixed point $\overline{w}_H$ is decreasing in τ. First-order stochastic dominance raises Φ pointwise, hence increases the fixed points. □

Computation that is often omitted. Given an explicit F, equations (4)–(5) simplify to closed form. Two examples:

Example 1: Exponential offers. If $W \sim \text{Exp}(\lambda)$, then $F(c) = 1 - e^{-\lambda c}$ and $\int_c^\infty w\, F(dw) = \int_c^\infty w\lambda e^{-\lambda w} dw = c\, e^{-\lambda c} + \lambda^{-1} e^{-\lambda c}$. Hence $\Phi(c) = c(1 - e^{-\lambda c}) + ce^{-\lambda c} + \lambda^{-1}e^{-\lambda c} = c + \lambda^{-1}e^{-\lambda c}$. Therefore

$$\overline{w}_L = (1-\beta)b + \beta\left(\overline{w}_L + \lambda^{-1}e^{-\lambda \overline{w}_L}\right) \quad \Rightarrow \quad (1-\beta)\overline{w}_L = (1-\beta)b + \beta\,\lambda^{-1}e^{-\lambda\overline{w}_L}.$$

This scalar equation can be solved by a single variable root finder. Then

$$(1-\beta)\overline{w}_H = (1-\beta)b + \beta\left[(1-\tau)\lambda^{-1}e^{-\lambda\overline{w}_H} + \tau\,\lambda^{-1}e^{-\lambda\overline{w}_L}\right].$$

Example 2: Bounded uniform offers. If $W \sim U[0,\omega]$, then for $0 \le c \le \omega$,

$$\Phi(c) = c\frac{c}{\omega} + \int_c^\omega \frac{w}{\omega}dw = \frac{c^2}{\omega} + \frac{\omega^2 - c^2}{2\omega} = \frac{c^2+\omega^2}{2\omega}.$$

Hence

$$\overline{w}_L = (1-\beta)b + \beta\,\frac{\overline{w}_L^2 + \omega^2}{2\omega}, \qquad \overline{w}_H = (1-\beta)b + \beta\left[(1-\tau)\frac{\overline{w}_H^2+\omega^2}{2\omega} + \tau\,\frac{\overline{w}_L^2+\omega^2}{2\omega}\right],$$

two scalar quadratic fixed point equations with unique positive solutions.

6.4 Welfare and Duration

Let $p_k := 1 - F(\overline{w}_k)$ be the acceptance probability at skill k. The expected unemployment duration starting at k before skill erosion occurs is geometrically distributed with parameter p_k in the first period. With erosion, the total expected duration from H solves

$$\text{Dur}(H) = p_H \cdot 1 + (1-p_H)\left[1 + (1-\tau)\,\text{Dur}(H) + \tau\,\text{Dur}(L)\right], \qquad \text{Dur}(L) = \frac{1}{p_L}.$$

Solving gives

$$\text{Dur}(H) = \frac{1 + (1-p_H)\tau\,\text{Dur}(L)}{p_H + (1-p_H)\tau}.$$

Higher τ raises $\text{Dur}(H)$ by lowering $\overline{w}_H$ less than it lowers the chance of remaining in H, which is the Ljungqvist–Sargent turbulence mechanism.

Remark 6.3 (Separations and wage resets)**.** Ljungqvist and Sargent also consider job separations and on-the-job wage resets driven by a Markov process. Then the employment value solves $V_E = \mathbb{E}\big[\sum_{t\ge0}\beta^t W_t\big]$ with a linear recursion, and the accept/reject comparison becomes $V_E(w)$ versus $R(k)$. Under a finite-state wage process or i.i.d. resets, the same contraction arguments deliver existence and a state-dependent reservation wage vector. See the references below for those variants.

6.5 What the extension teaches

The two-threshold system (4)–(5) shows how unemployment insurance and turbulence interact. A rise in τ directly depresses the continuation value from waiting while high skilled and thus lowers $\overline{w}_H$, while $\overline{w}_L$ remains pinned by (5). For calibrated F this mechanism reproduces long spells after turbulence shocks without invoking changes in preferences or discounting.

6.6 Exercises

Prove Theorem 6.2 for a general Borel skill space with an unemployment transition kernel $G_U(\cdot \mid k)$ that is stochastically decreasing in a scalar turbulence parameter. Show that the vector reservation wage map remains a contraction and that thresholds are decreasing in turbulence in the monotone likelihood ratio order when F has an increasing failure rate.

Specialize Example 1 to calibrate (β, b, λ) monthly and compute $(\overline{w}_H, \overline{w}_L)$ over a grid of τ. Plot acceptance probabilities and expected durations. Explain the nonlinearity in the response of durations as τ increases from 0 to 0.3.

Augment the baseline with a Poisson separation probability δ for employed workers. Show that the accept value becomes $E(w) = \frac{w}{1-\beta(1-\delta)}$ and re-derive the system (4)–(5) with $(1-\beta)$ replaced by $(1-\beta(1-\delta))$. Compare the implied reservation wages across δ.

7 Further Classic Examples

Example 7.1 (Scarf's Lost Sales Inventory Control)**.** A single item is reviewed each period. The state is the inventory position. The action is the order quantity. Demand is i.i.d. with known distribution. Holding and shortage costs are convex. The optimal policy is a base stock rule characterized by a critical level. See Scarf (1960).

Example 7.2 (Rust's Bus Engine Replacement)**.** A decision maker chooses each period whether to replace or continue operating a bus engine. The state is accumulated mileage, controlled by a Markov transition with stochastic use. Replacement resets mileage and incurs a cost. Continuation yields stochastic maintenance costs. The optimal policy is of threshold type. See Rust (1987).

Example 7.3 (Investment with Convex Adjustment Costs)**.** The state is the capital stock and the action is investment. The payoff is profit net of a convex adjustment cost. Productivity follows a Markov process. Under standard conditions the value function is concave and the optimal policy is increasing in productivity. See Stokey and Lucas with Prescott (1989), chapter 9, and Bertsekas and Shreve (1978), chapter 9.

8 Exercises

1. Prove monotonicity and discounting of the Bellman operator. Show that if $v \leq w$ then $Tv \leq Tw$, and that $T(v + c) = Tv + \beta c$ for any constant c. Deduce an a priori sup norm

bound on V^* in terms of $\|r\|_\infty$.

2. In the McCall model assume W has a continuous density with compact support. Show that the reservation wage is continuous in b and in β. Derive the elasticity of the acceptance probability with respect to b.

3. For a finite MDP implement the policy improvement algorithm and prove that it terminates in finitely many steps. Explain why termination need not be monotone in the number of states across instances.

4. Consider Scarf's inventory model with linear holding and penalty costs and i.i.d. demand. Derive the base stock structure using the K-convexity method or show that value iteration preserves convexity and prove that a threshold policy is optimal.

5. In Rust's engine replacement model assume mileage evolves as a birth process with reflecting boundary at a maximum. Prove that the continuation value is increasing in the state and that the optimal policy is a one sided threshold. State the conditions under which the threshold is unique.

6. Let productivity follow a finite state Markov chain in the investment model. Prove that value iteration preserves concavity and that the optimal investment policy is increasing in productivity if the profit function satisfies single crossing.

References

[1] Bellman, R. (1957). *Dynamic Programming.* Princeton University Press.

[2] Bertsekas, D. and Shreve, S. (1978). *Stochastic Optimal Control: The Discrete-Time Case.* Academic Press.

[3] Lippman, S. and McCall, J. (1976). The economics of job search: a survey. *Economic Inquiry* 14, 155–189.

[4] McCall, J. (1970). Economics of information and job search. *Quarterly Journal of Economics* 84, 113–126.

[5] Puterman, M. (1994). *Markov Decision Processes.* Wiley.

[6] Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica* 55, 999–1033.

[7] Scarf, H. (1960). The optimality of (s, S) policies in the dynamic inventory problem. In *Mathematical Methods in the Social Sciences*, Stanford University Press.

[8] Stokey, N., Lucas, R., with Prescott, E. (1989). *Recursive Methods in Economic Dynamics.* Harvard University Press.

[9] Phelps, E. S. (1968). Money-wage dynamics and labor-market equilibrium. *Journal of Political Economy*, 76(4), 678–711.

[10] Ljungqvist, L. and Sargent, T. J. (2010). European unemployment and turbulence. In *Handbook of Monetary Economics*, vol. 3, 1175–1237. North-Holland.

[11] Sargent, T. J. (various years). *European Unemployment: From a Worker's Perspective* [course packet, "phelpsall.pdf"]. Available at `tomsargent.com`.