Corrections for
ABSTRACT DYNAMIC PROGRAMMING
by Dimitri P. Bertsekas
Athena Scientific

Last Updated: 2/4/14

p. 143 (-3) Change Eq. (4.10) to

$$J_{\pi^k[x]}(x) \leq J^*(x) + \epsilon_k. \tag{4.10}$$

p. 159 (-15) Change "$J_{\mu^k} \to J^*$" to "$J_k \to J^*$"

p. 165 (-5) Change "$T_{\mu^0}^{m_0} J_0 \geq J_1$" to "$T_{\mu^0}^{m_0} J_0 = J_1$"

p. 177 (-13) Change "Prop. 3.2.4" to "Prop. 3.2.3"

p. 178 (+14) Change "(S is equal to $\Re^2$ here)" to "(S is equal to $S = \big\{J \mid J(1) > 0,\ J(2) > 0\big\}$ here)"

p. 180 (+3) Change "infinite horizon examples" to "infinite horizon models"

p. 185 (-2) Change "X_k" to

$$\hat U_k(x) = \big\{u \in U(x) \mid f(x,u,w) \in X_k,\ \forall\ w \in W(x,u)\big\}$$

p. 240 Replace the last line with "It can be seen that $U_k(x,\lambda)$ is equal to the set

$$\hat U_k(x) = \big\{u \in U(x) \mid f(x,u,w) \in X_k,\ \forall\ w \in W(x,u)\big\}$$

given in the statement of the exercise."

p. 243 Add the reference

[CaR11] Canbolat, P. G., and Rothblum, U. G., 2011. "(Approximate) Iterated Successive Approximations Algorithm for Sequential Decision Processes," Technion Report; appeared in Annals of Operations Research, Vol. 208, pp. 309-320.

pp. 171-178 Section 4.5, Affine Monotonic Models, is fine as is, but it assumes finite state and control spaces. Given here is a revised version, which applies to infinite state and control spaces as well.

4.5 AFFINE MONOTONIC MODELS

In this section, we consider the case

$$T_\mu J = A_\mu J + b_\mu, \tag{4.40}$$

where for each μ, b_μ is a given function in $R^+(X)$, the set of all nonnegative real-valued functions on X, and $A_\mu : E^+(X) \mapsto E^+(X)$ is a given mapping, where $E^+(X)$ is the set of all nonnegative extended real-valued functions on X. We assume that A_μ has the "linearity" property

$$A_\mu(J_1 + J_2) = A_\mu J_1 + A_\mu J_2, \qquad \forall\ J_1, J_2 \in E^+(X). \tag{4.41}$$

Thus if $J, J' \in E^+(X)$ with $J' \geq J$, we have $A_\mu(J' - J) \geq 0$ [since A_μ maps $E^+(X)$ to $E^+(X)$] and hence [using also Eq. (4.41)] $A_\mu J' = A_\mu J + A_\mu(J' - J) \geq A_\mu J$, so that A_μ and T_μ are monotone in the sense that

$$J, J' \in E^+(X), \quad J \leq J' \qquad \Rightarrow \qquad A_\mu J \leq A_\mu J', \quad T_\mu J \leq T_\mu J'.$$

(In the preceding equations we use our convention $\infty + \infty = \infty - \infty = r + \infty = \infty + r = \infty$ for any real number r; see Appendix A.) We refer to this model, with a function $\bar J \in R^+(X)$, as an *affine monotonic* model.

An example of this model is when X is a countable set, A_μ is defined by the transition probabilities corresponding to μ, and $\bar J(x) \equiv 0$. Then we obtain the countable-state case of the negative DP model of [Str66], which is fully covered by the theory of Section 4.3, under Assumption I.

Another special case is the multiplicative model of Example 1.2.8, where X and U are finite sets, $\bar J$ is the unit function ($\bar J = e$), and for transition probabilities $p_{xy}(u)$ and function $g(x, u, y) \geq 0$, we have

$$H(x, u, J) = \sum_{y \in X} p_{xy}(u) g(x, u, y) J(y). \tag{4.42}$$

Thus with $b_\mu = 0$ and the matrix A_μ having components

$$A_\mu(x, y) = p_{xy}\big(\mu(x)\big) g\big(x, \mu(x), y\big),$$

we obtain an affine monotonic model.

In a variant of the multiplicative model that involves a cost-free and absorbing termination state 0, similar to SSP problems, H may contain a "constant" term, i.e., have the form

$$H(x, u, J) = p_{x0}(u) g(x, u, 0) + \sum_{y \in X} p_{xy}(u) g(x, u, y) J(y), \tag{4.43}$$

in which case $b_\mu(x) = p_{x0}\big(\mu(x)\big) g\big(x, \mu(x), 0\big)$. A special case of this model is the risk-sensitive SSP problem with exponential cost function, which will be discussed later in Section 4.5.3.

In the next two subsections we will consider two alternative lines of semicontractive model analysis. The first assumes the monotone increase condition $T\bar{J} \geq \bar{J}$, and relies on Assumption I of this chapter. The second line of analysis follows the approach of Section 3.2.1 (irregular policies have infinite cost for some $x \in X$), based on Assumption 3.2.1 with an appropriate choice of a subset S of real-valued functions. Analyses based on the monotone decrease condition $T\bar{J} \leq \bar{J}$, and on the perturbation-based approach of Section 3.2.2 are also possible, but will not be pursued in detail. Of course the strong results of Chapter 2 may also apply when there is a weighted sup-norm for which A_μ is a contraction for all μ over $B(X)$, and with the same modulus.

4.5.1 Increasing Affine Monotonic Models

In this subsection we assume that the condition $T\bar{J} \geq \bar{J}$ holds and that the remaining two conditions of Assumption I are satisfied. Then the affine monotonic model admits a straightforward analysis with a choice

$$S \subset \big\{ J \in E^+(X) \mid J \geq \bar{J} \big\}, \tag{4.44}$$

based on the theory of Section 4.4.1 and the parts of Section 4.3 that relate to the monotone increase Assumption I. In particular, we have the following proposition.

Proposition 4.5.1: Consider the affine monotonic model, assuming that $T\bar{J} \geq \bar{J}$ and that the remaining conditions of Assumption I hold. Assume that there exists an optimal S-regular policy, where S satisfies Eq. (4.44). Then:

(a) The optimal cost function J^* is the unique fixed point of T within S.

(b) A policy μ is optimal if and only if $T_\mu J^* = TJ^*$.

(c) Under the compactness assumptions of Prop. 4.3.14, we have $T^k J \to J^*$ for every $J \in S$.

Proof: (a) Follows from Prop. 4.4.1(a).

(b) Follows from Prop. 4.3.9.

(c) Follows from Prop. 4.4.1(c). **Q.E.D.**

4.5.2 Nonincreasing Affine Monotonic Models

We now consider the affine monotonic model without assuming the monotone increase condition $T\bar{J} \geq \bar{J}$. We will use the approach of Section 3.2.1, assuming that $\bar{J} \in S$ and that S is equal to one of the three choices

$$S = R^+(X) = \big\{ J \mid 0 \le J(x) < \infty,\ \forall\ x \in X \big\},$$

$$S = R_p^+(X) = \big\{ J \mid 0 < J(x) < \infty,\ \forall\ x \in X \big\},$$

$$S = R_b^+(X) = \left\{ J \ \Big|\ 0 < \inf_{x \in X} J(x) \le \sup_{x \in X} J(x) < \infty \right\}.$$

Note that if X is finite, we have $R_p^+(X) = R_b^+(X)$.

We first derive an expression for the cost function of a policy and obtain conditions for S-regularity. Using the form of T_μ and the "linearity" condition (4.41), we have

$$T_\mu^k J = A_\mu^k J + \sum_{m=0}^{k-1} A_\mu^m b_\mu, \qquad \forall\ J \in S,\ k = 1, 2, \ldots.$$

By definition, μ is S-regular if $J_\mu \in S$, and $\lim_{k\to\infty} T_\mu^k J = J_\mu$ for all $J \in S$, or equivalently if for all $J \in S$ we have

$$\limsup_{k\to\infty}\ A_\mu^k J + \sum_{m=0}^{\infty} A_\mu^m b_\mu = \limsup_{k\to\infty}\ A_\mu^k \bar{J} + \sum_{m=0}^{\infty} A_\mu^m b_\mu \in S.$$

Letting $J = 2\bar{J}$ and using the fact $A_\mu^k(2\bar{J}) = 2A_\mu^k \bar{J}$ [cf. Eq. (4.41)], we see that $A_\mu^k \bar{J} \to 0$. It follows that *μ is S-regular if and only if*

$$\lim_{k\to\infty} A_\mu^k J = 0,\quad \forall\ J \in S, \qquad \text{and} \qquad \sum_{m=0}^{\infty} A_\mu^m b_\mu \in S. \tag{4.45}$$

We will now consider conditions for Assumption 3.2.1 to hold, so that the results of Prop. 3.2.1 will follow. For the choices $S = R^+(X)$ and $S = R_b^+(X)$, parts (a), (b), and (f) of this assumption are automatically satisfied [a proof, to be given later, will be required for part (f) and the case $S = R_b^+(X)$]. For the choice $S = R_p^+(X)$, part (a) of this assumption is automatically satisfied, while part (b),

$$\inf_{\mu:\, R_p^+(X)\text{-regular}} J_\mu \in R_p^+(X),$$

and part (f) will be assumed in the proposition that follows. The compactness condition of Assumption 3.2.1(d) and the technical condition of Assumption 3.2.1(e) are needed, and they will be assumed.

The critical part of Assumption 3.2.1 is (c), which requires that for each S-irregular policy μ and each $J \in S$, there is at least one state $x \in X$ such that

$$\limsup_{k\to\infty} (T_\mu^k J)(x) = \limsup_{k\to\infty} (A_\mu^k J)(x) + \sum_{m=0}^{\infty} (A_\mu^m b_\mu)(x) = \infty.$$

This part is satisfied if and only if for each S-irregular μ and $J \in S$, there is at least one $x \in X$ such that

$$\limsup_{k\to\infty} (A_\mu^k J)(x) = \infty \quad \text{or} \quad \sum_{m=0}^{\infty} (A_\mu^m b_\mu)(x) = \infty. \tag{4.46}$$

Note that this cannot be true if $S = R^+(X)$ and $b_\mu = 0$ [as in the multiplicative cost case of Eq. (4.42)], because for $J = 0$, the preceding condition is violated. On the other hand, if $S = R_p^+(X)$ or $S = R_b^+(X)$, the condition (4.46) is satisfied even if $b_\mu = 0$, provided that for each S-irregular μ and $J \in S$, there is at least one $x \in X$ with

$$\limsup_{k\to\infty} (A_\mu^k J)(x) = \infty.$$

We have the following proposition.

Proposition 4.5.2: Consider the affine monotonic model and let $S = R^+(X)$ or $S = R_p^+(X)$ or $S = R_b^+(X)$. Assume that the following hold:

(1) There exists an S-regular policy.

(2) If μ is an S-irregular policy, then for each function $J \in S$, Eq. (4.46) holds for at least one $x \in X$.

(3) The function $\hat{J}$ given by

$$\hat{J}(x) = \inf_{\mu:\, S\text{-regular}} J_\mu(x), \qquad x \in X,$$

belongs to S.

(4) The control set U is a metric space, and the set

$$\big\{u \in U(x) \mid H(x,u,J) \le \lambda\big\}$$

is compact for every $J \in S$, $x \in X$, and $\lambda \in \Re$.

(5) For each sequence $\{J_m\} \subset S$ with $J_m \uparrow J$ for some $J \in S$ we have

$$\lim_{m\to\infty} (A_\mu J_m)(x) = (A_\mu J)(x), \qquad \forall\ x \in X,\ \mu \in \mathcal{M}.$$

(6) In the case where $S = R_p^+(X)$, for each function $J \in S$, there exists a function $J' \in S$ such that $J' \le J$ and $J' \le TJ'$.

Then:

(a) The optimal cost function J^* is the unique fixed point of T within S.

(b) We have $T^k J \to J^*$ for every $J \in S$. Moreover there exists an optimal S-regular policy.

(c) A policy μ is optimal if and only if $T_\mu J^* = TJ^*$.

Proof: If $S = R^+(X)$ or $S = R_p^+(X)$, it can be verified that all the parts of Assumption 3.2.1 are satisfied, and the results follow from Prop. 3.2.1 [this includes part (f), which is satisfied by assumption in the case of $S = R_p^+(X)$; cf. condition (6)]. If $S = R_b^+(X)$, the proof is similar, but to apply Prop. 3.2.1, we need to show that Assumption 3.2.1(f) is satisfied.

To this end, we will show that for each $J \in S$, there exists a $J' \in S$ of the form $J' = \alpha \hat{J}$, where α is a scalar with $0 < \alpha < 1$, such that $J' \le J$ and $J' \le TJ'$, so again the results will follow from Prop. 3.2.1. Indeed, from Lemma 3.2.4, we have that $\hat{J}$ is a fixed point of T. For any $J \in S$, choose $J' = \alpha \hat{J}$, with $\alpha \in (0, 1)$, such that $J' \le J$, and let μ be an S-regular policy μ such that $T_\mu J' = TJ'$ [cf. Lemma 3.2.1 and condition (4)]. Then, we have $TJ' = T_\mu J' = T_\mu(\alpha \hat{J}) = \alpha A_\mu \hat{J} + b_\mu \ge \alpha(A_\mu \hat{J} + b_\mu) = \alpha T_\mu \hat{J} \ge \alpha T \hat{J} = \alpha \hat{J} = J'$. **Q.E.D.**

Note the difference between Props. 4.5.1 and 4.5.2: in the former, the uniqueness of fixed point of T is guaranteed within a smaller set of functions when $\bar{J} \in R_p^+(X)$. Similarly, the convergence of VI is guaranteed from within a smaller range of starting functions when $\bar{J} \in R_p^+(X)$.

4.5.3 Exponential Cost Stochastic Shortest Path Problems

We will now apply the analysis of the affine monotonic model to SSP problems with an exponential cost function, which is introduced to incorporate risk sensitivity in the control selection process.

Consider an SSP problem with finite state and control spaces, transition probabilities $p_{xy}(u)$, and real-valued transition costs $h(x, u, y)$. State 0 is a termination state, which is cost-free and absorbing. Instead of the standard additive cost function (cf. Example 1.2.6), we consider an exponential cost function of the form

$$J_\mu(x) = \lim_{k \to \infty} E\left\{ \exp\left(\sum_{m=0}^{k-1} h\big(x_m, \mu(x_m), x_{m+1}\big) \right) \;\Big|\; x_0 = x \right\}, \quad x \in X,$$

where $\{x_0, x_1, \ldots\}$ denotes the trajectory produced by the Markov chain under policy μ. This is an affine monotonic model with $\bar{J} = e$ and mapping

T_μ given by

$$(T_\mu J)(x) = \sum_{y \in X} p_{xy}\big(\mu(x)\big)\exp\big(h(x,\mu(x),y)\big)J(y) + p_{x0}\big(\mu(x)\big)\exp\big(h(x,\mu(x),0)\big), \qquad x \in X, \tag{4.47}$$

[cf. Eq. (4.43)]. Here A_μ and b_μ have components

$$A_\mu(x,y) = p_{xy}\big(\mu(x)\big)\exp\big(h(x,\mu(x),y)\big), \tag{4.48}$$

$$b_\mu(x) = p_{x0}\big(\mu(x)\big)\exp\big(h(x,\mu(x),0)\big). \tag{4.49}$$

Note that there is a distinction between S-irregular policies and improper policies (the ones that never terminate). In particular, there may exist improper policies, which are S-regular because they can generate some negative transition costs $h(x,u,y)$, which make A_μ contractive [cf. Eq. (4.47)]. Similarly, there may exist proper policies (i.e., terminate with probability one), which are S-irregular because for the corresponding A_μ and b_μ we have $\sum_{m=0}^{\infty}(A_\mu^m b_\mu)(x) \to \infty$ for some x.

We may consider the two cases where the condition $T\bar J \geq \bar J$ holds (cf. Section 4.5.1) and where it does not (cf. Section 4.5.2), as well as a third case where none of these conditions applies, but the perturbation-based theory of Section 3.2.2 or the contractive theory of Chapter 2 can be used. Consider first the case where $T\bar J \geq \bar J$. An example is when

$$h(x,u,y) \geq 0, \qquad \forall\ x,\, y \in X,\ u \in U(x),$$

so that from Eq. (4.47), we have $\exp\big(h(x,u,y)\big) \geq 1$, and since $\bar J = e$, it follows that $T_\mu \bar J = A_\mu \bar J + b_\mu \geq \bar J$ for all $\mu \in \mathcal{M}$. As in Section 4.5.1, by letting

$$S \subset \big\{J \in E^+(X) \mid J \geq \bar J\big\},$$

and by assuming the existence of an optimal S-regular policy, we can apply Prop. 4.5.1 to obtain the corresponding conclusions. In particular, J^* is the unique fixed point of T within S [cf. Eq. (4.44)], all optimal policies are S-regular and satisfy the optimality condition $T_\mu J^* = TJ^*$, and VI yields J^* in the limit, when initialized from within S.

On the other hand, there are interesting applications where the condition $T\bar J \geq \bar J$ does not hold. The following is an example.

Example 4.5.1 (Optimal Stopping with Risk-Sensitive Cost)

Consider an SSP problem where there are two controls at each x: *stop*, in which case we move to the termination state 0 with a cost $s(x)$, and *continue*, in which case we move to a state y, with given transition probabilities p_{xy} [at no cost if $y \neq 0$ and a cost $\bar s(x)$ if $y = 0$]. The mapping H has the form

$$H(x,u,J) = \begin{cases} \exp\big(s(x)\big) & \text{if } u = \text{stop}, \\ \sum_{y \in X} p_{xy} J(y) + p_{x0}\exp\big(\bar s(x)\big) & \text{if } u = \text{continue}, \end{cases}$$

and $\bar{J}$ is the unit function e. Here the stopping cost $s(x)$ is often naturally negative for some x (this is true for example in search problems of the type discussed in Example 3.2.1), so the condition $T\bar{J} \geq \bar{J}$ can be written as

$$\min\left\{\exp\big(s(x)\big),\ \sum_{y\in X} p_{xy} + p_{x0}\exp\big(\bar{s}(x)\big)\right\} \geq 1, \qquad \forall\ x \in X,$$

and is violated.

When the condition $T\bar{J} \geq \bar{J}$ does not hold, we may use the analysis of Section 4.5.2, under the conditions of Prop. 4.5.2, chief among which is that an S-regular policy exists, and for every S-irregular policy μ and $J \in S$, there exists $x \in X$ such that

$$\limsup_{k\to\infty}\,(A_\mu^k J)(x) = \infty \quad \text{or} \quad \sum_{m=0}^{\infty}(A_\mu^m b_\mu)(x) = \infty,$$

where A_μ and b_μ are given by Eqs. (4.48), (4.49) [cf. Eq. (4.46)], and $S = R^+(X)$ or $S = R_p^+(X)$ or $S = R_b^+(X)$.

If these conditions do not hold, we may also use the approach of Section 3.2.2, which is based on adding a perturbation δ to b_μ. We assume that the optimal cost function J_δ^* of the δ-perturbed problem is a fixed point of the mapping T_δ given by

$$(T_\delta J)(x) = \min_{u\in U(x)}\left\{\sum_{y\in X} p_{xy}(u)\exp\big(h(x,u,y)\big)J(y) + p_{x0}(u)\exp\big(h(x,u,0)\big)\right\} + \delta, \qquad x \in X,$$

and we assume existence of an optimal S-regular policy with

$$S = \big\{B(X) \mid J(x) > 0,\ \forall\ x \in X\big\},$$

where $B(X)$ is the space of bounded functions with respect to some weighted sup-norm. The remaining conditions of Assumption 3.2.2 are relatively mild and we assume that they hold. Then Prop. 3.2.3 applies and shows that J^* is equal to $\lim_{\delta\downarrow 0} J_\delta^*$ and is the unique fixed point of T within the set $\{J \in S \mid J \geq J^*\}$, and that the VI sequence $\{T^k J\}$ converges to J^* starting from a function $J \in S$ with $J \geq J^*$. Under some circumstances where there is no optimal S-regular policy, we may also be able to use Prop. 3.2.2. In particular, it may happen that for some $x \in X$, $J^*(x)$ is strictly smaller than $\lim_{\delta\downarrow 0} J_\delta^*(x)$, the optimal cost over all S-regular policies, while there may exist S-irregular policies that are optimal and attain J^*, in which case Prop. 3.2.2 applies.

The following example illustrates the possibilities, and highlights the ranges of applicability of Props. 4.5.1 and 4.5.2 (which are special cases of Props. 4.4.1 and 3.2.1, respectively), and Props. 3.2.2 and 3.2.3.

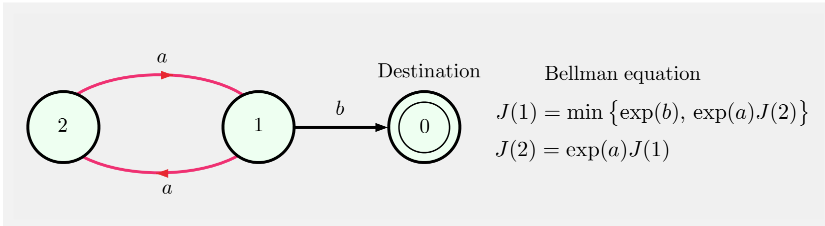

Figure 4.5.1. Shortest path problem with exponential cost function. The cost that is exponentiated is shown next to each arc.

Example 4.5.2 (Shortest Paths with Risk-Sensitive Cost)

Consider the context of the three-node shortest path problem of Section 3.1.2, but with the exponential cost function of the present subsection (see Fig. 4.5.1). Here the DP model has two states: $x = 1, 2$. There are two policies denoted μ and $\overline{\mu}$: the 1st policy is $2 \to 1 \to 0$, while the 2nd policy is $2 \to 1 \to 2$. The corresponding mappings T_μ and $T_{\overline{\mu}}$ are given by

$$(T_\mu J)(1) = \exp(b), \qquad (T_\mu J)(2) = \exp(a)J(1),$$

$$(T_{\overline{\mu}} J)(1) = \exp(a)J(2), \qquad (T_{\overline{\mu}} J)(2) = \exp(a)J(1).$$

Moreover, for $k \geq 2$, we have

$$(T_\mu^k J)(1) = \exp(b), \qquad (T_\mu^k J)(2) = \exp(a+b),$$

and

$$(T_{\overline{\mu}}^k J)(1) = \begin{cases} \big(\exp(a)\big)^k J(1) & \text{if } k \text{ is even,} \\ \big(\exp(a)\big)^k J(2) & \text{if } k \text{ is odd,} \end{cases}$$

$$(T_{\overline{\mu}}^k J)(2) = \begin{cases} \big(\exp(a)\big)^k J(1) & \text{if } k \text{ is odd,} \\ \big(\exp(a)\big)^k J(2) & \text{if } k \text{ is even.} \end{cases}$$

The cost functions of μ and $\overline{\mu}$, with $\bar{J} = e$, are

$$J_\mu(1) = \exp(b), \qquad J_\mu(2) = \exp(a+b),$$

$$J_{\overline{\mu}}(1) = J_{\overline{\mu}}(2) = \lim_{k\to\infty} \exp\left(\sum_{m=0}^{k-1} a\right) = \lim_{k\to\infty} \big(\exp(a)\big)^k.$$

Clearly the proper policy μ is S-regular, since $T_\mu^k J = J_\mu$ for all $k \geq 2$. The improper policy $\overline{\mu}$ is S-irregular when $a > 0$ since $J_{\overline{\mu}}(1) = J_{\overline{\mu}}(2) = \infty$, and when $a = 0$ (since $T_{\overline{\mu}}^k J$ depends on J), for any reasonable choice of S. However, in the case where $a < 0$ and there is a negative cycle $2 \to 1 \to 2$, $\overline{\mu}$ is optimal and $R^+(X)$-regular [but not $R_p^+(X)$-regular], since $T_{\overline{\mu}}^k J = \big(\exp(a)\big)^k J \to 0 \in R^+(X)$ for all $J \in R^+(X)$.

The major lines of analysis of semicontractive models that we have discussed are all illustrated in the five possible combinations of values of a and b given below. Each of these five combinations exhibits significantly different characteristics, and in each case the assertion about the set of fixed points of T is based on a different proposition!

(a) *Case* $a > 0$: Here the regular policy μ is optimal, and the irregular policy $\overline{\mu}$ has infinite cost for all x. It can be seen that the assumptions of Prop. 4.5.2 with $S = R_p^+(X)$ apply. Note here that $b_{\overline{\mu}} = 0$, so condition (2) of Prop. 4.5.2 is violated when $S = R^+(X)$ [the condition (4.46) is violated for $J = 0$]. Consistently with this fact, T has the additional fixed point $J = 0$ within $R^+(X)$, while value iteration starting from $J_0 = 0$ generates $T^k J_0 = 0$ for all k, and does not converge to J^*.

(b) *Case* $a = 0$ *and* $b > 0$: Here the irregular policy $\overline{\mu}$ is optimal, and the assumptions of Props. 4.5.1 and 4.5.2, with both $S = R^+(X)$ and $S = R_p^+(X)$, are violated [despite the fact that Assumption (I) holds for this case]. The assumptions of Prop. 3.2.3 are also violated because the only optimal policy is irregular. However, consistent with Prop. 3.2.2, $\lim_{\delta\downarrow 0} J_\delta^*$ is the optimal cost over the regular policies only, which is J_μ. In particular, we have

$$J_\mu(1) = \exp(b) = \lim_{\delta\downarrow 0} J_\delta^*(1) > J^*(1) = 1.$$

Here the set of fixed points of T is

$$\big\{J \mid J \le \exp(b)e,\ J(1) = J(2)\big\},$$

and contains vectors J from the range $J > J^*$ as well as from the range $J < J^*$ (however, $J^* = e$ is the "smallest" fixed point with the property $J \ge \bar{J} = e$, consistently with Prop. 4.3.3).

(c) *Case* $a = 0$ *and* $b = 0$: Here μ and $\overline{\mu}$ are both optimal, and the results of Prop. 4.5.1 apply with $S = \big\{J \mid J \ge J^* = \bar{J} = e\big\}$. However, the assumptions of Prop. 4.5.2 are violated, and indeed T has multiple fixed points within both $R_p^+(X)$ and (a fortiori) $R^+(X)$; the set of its fixed points is

$$\big\{J \mid J \le e,\ J(1) = J(2)\big\}.$$

(d) *Case* $a = 0$ *and* $b < 0$: Here the regular policy μ is optimal. However, the assumptions of Props. 4.5.1 and 4.5.2 are violated. On the other hand, Prop. 3.2.3 applies with $S = \big\{J \mid J \ge J^*\big\}$, so T has a unique fixed point within S, while value iteration converges to J^* starting from within S. Here again T has multiple fixed points within $R_p^+(X)$ and (a fortiori) $R^+(X)$; the set of its fixed points is

$$\big\{J \mid J \le \exp(b)e,\ J(1) = J(2)\big\}.$$

(e) *Case* $a < 0$: Here $\overline{\mu}$ is optimal and also $R^+(X)$-regular [but not $R_p^+(X)$-regular, since $J_{\overline{\mu}} = 0 \notin R_p^+(X)$]. However, the assumptions of Prop.

4.5.1, and Prop. 4.5.2 with both $S = R^+(X)$ and $S = R_p^+(X) = R_b^+(X)$ are violated. Still, however, our analysis applies and in a stronger form, because both T_μ and $T_{\overline{\mu}}$ are contractions. Thus we are dealing with a contractive model for which the results of Chapter 2 apply ($J^* = 0$ is the unique fixed point of T over the entire space $\Re^2$, and value iteration converges to J^* starting from any $J \in \Re^2$).

4.6 AN OVERVIEW OF SEMICONTRACTIVE MODELS AND RESULTS

Several semicontractive models and results have been discussed in this chapter and in Chapter 3, under several different assumptions, and it may be worth summarizing them. Three types of models have been considered:

(a) Models where the set S may include extended real-valued functions, an optimal S-regular policy is assumed to exist, and no other conditions are placed on S-irregular policies. These models are covered by Props. 3.1.1, 3.1.2, and 4.4.1, and they may require substantial analysis to verify the corresponding assumptions. Note here that the existence of an optimal stationary policy (regular or irregular) may not be easily verified. However, in the special case where Assumption I and the compactness assumption of Prop. 4.3.14 holds, existence of an optimal stationary policy is guaranteed, and then requiring the existence of an optimal S-regular policy may not be restrictive.

(b) Models where S consists of real-valued functions, and conditions are placed on S-irregular policies, which roughly imply that their cost is infinite from some states. There are two propositions that apply to such models: Prop. 3.1.3, which assumes also that an optimal S-regular policy exists, and Prop. 3.2.1 (and its specialized version, Prop. 4.5.2, for affine monotonic models), which indirectly guarantees existence of an optimal S-regular policy through other assumptions.

(c) Perturbation models, where S-irregular policies cannot be adequately differentiated from S-regular ones on the basis of their cost functions, but they become differentiated once a positive additive perturbation is added to their associated mapping. These models include the ones of Sections 3.2.2 and are covered by Props. 3.2.2-3.2.4.

Variants of these models may also include special structure that enhances the power of the analysis, as for example in SSP problems, linear quadratic problems, and affine monotonic and exponential cost models.

The two significant issues in the analysis of semicontractive models are how to select the set S so that an optimal S-regular policy exists, and how to verify the existence of such a policy. There seems to be no universal approach for addressing these issues, as can be evidenced by the variety of alternative sets of assumptions that we have introduced, and by the