APPENDIX A:

Notation and Mathematical Conventions

In this appendix we collect our notation, and some related mathematical facts and conventions.

A.1 SET NOTATION AND CONVENTIONS

If X is a set and x is an element of X, we write $x \in X$. A set can be specified in the form $X = \{x \mid x \text{ satisfies } P\}$, as the set of all elements satisfying property P. The union of two sets X_1 and X_2 is denoted by $X_1 \cup X_2$, and their intersection by $X_1 \cap X_2$. The empty set is denoted by $\emptyset$. The symbol $\forall$ means "for all."

The set of real numbers (also referred to as scalars) is denoted by $\Re$. The set of extended real numbers is denoted by $\Re^*$:

$$\Re^* = \Re \cup \{\infty, -\infty\}.$$

We write $-\infty < x < \infty$ for all real numbers x, and $-\infty \leq x \leq \infty$ for all extended real numbers x. We denote by $[a, b]$ the set of (possibly extended) real numbers x satisfying $a \leq x \leq b$. A rounded, instead of square, bracket denotes strict inequality in the definition. Thus $(a, b]$, $[a, b)$, and (a, b) denote the set of all x satisfying $a < x \leq b$, $a \leq x < b$, and $a < x < b$, respectively.

Generally, we adopt standard conventions regarding addition and multiplication in $\Re^*$, except that we take

$$\infty - \infty = -\infty + \infty = \infty,$$

and we take the product of 0 and ∞ or $-\infty$ to be 0. In this way the sum and product of two extended real numbers is well-defined. Division by 0 or ∞ does not appear in our analysis. In particular, we adopt the following rules in calculations involving ∞ and $-\infty$:

$$\alpha + \infty = \infty + \alpha = \infty, \qquad \forall\ \alpha \in \Re^*,$$

$$\alpha - \infty = -\infty + \alpha = -\infty, \qquad \forall\ \alpha \in [-\infty, \infty),$$

$$\alpha \cdot \infty = \infty, \qquad \alpha \cdot (-\infty) = \infty, \qquad \forall\ \alpha \in (0, \infty],$$

$$\alpha \cdot \infty = -\infty, \qquad \alpha \cdot (-\infty) = -\infty, \qquad \forall\ \alpha \in [-\infty, 0),$$

$$0 \cdot \infty = \infty \cdot 0 = 0 = 0 \cdot (-\infty) = (-\infty) \cdot 0, \qquad -(-\infty) = \infty.$$

Under these rules, the following laws of arithmetic are still valid within $\Re^*$:

$$\alpha_1 + \alpha_2 = \alpha_2 + \alpha_1, \qquad (\alpha_1 + \alpha_2) + \alpha_3 = \alpha_1 + (\alpha_2 + \alpha_3),$$

$$\alpha_1\alpha_2 = \alpha_2\alpha_1, \qquad (\alpha_1\alpha_2)\alpha_3 = \alpha_1(\alpha_2\alpha_3).$$

We also have

$$\alpha(\alpha_1 + \alpha_2) = \alpha\alpha_1 + \alpha\alpha_2$$

if either $\alpha \geq 0$ or else $(\alpha_1 + \alpha_2)$ is not of the form $\infty - \infty$.

Inf and Sup Notation

The *supremum* of a nonempty set $X \subset \Re^*$, denoted by $\sup X$, is defined as the smallest $y \in \Re^*$ such that $y \geq x$ for all $x \in X$. Similarly, the *infimum* of X, denoted by $\inf X$, is defined as the largest $y \in \Re^*$ such that $y \leq x$ for all $x \in X$. For the empty set, we use the convention

$$\sup \varnothing = -\infty, \qquad \inf \varnothing = \infty.$$

If $\sup X$ is equal to an $\overline{x} \in \Re^*$ that belongs to the set X, we say that $\overline{x}$ is the *maximum point* of X and we write $\overline{x} = \max X$. Similarly, if $\inf X$ is equal to an $\overline{x} \in \Re^*$ that belongs to the set X, we say that $\overline{x}$ is the *minimum point* of X and we write $\overline{x} = \min X$. Thus, when we write $\max X$ (or $\min X$) in place of $\sup X$ (or $\inf X$, respectively), we do so just for emphasis: we indicate that it is either evident, or it is known through earlier analysis, or it is about to be shown that the maximum (or minimum, respectively) of the set X is attained at one of its points.

A.2 FUNCTIONS

If f is a function, we use the notation $f : X \mapsto Y$ to indicate the fact that f is defined on a nonempty set X (its *domain*) and takes values in a set Y (its *range*). Thus when using the notation $f : X \mapsto Y$, we implicitly assume that X is nonempty. We will often use the *unit function* $e : X \mapsto \Re$, defined by

$$e(x) = 1, \qquad \forall\ x \in X.$$

Given a set X, we denote by $\mathcal{R}(X)$ the set of real-valued functions $J : X \mapsto \Re$, and by $\mathcal{E}(X)$ the set of all extended real-valued functions $J : X \mapsto \Re^*$. For any collection $\{J_\gamma \mid \gamma \in \Gamma\} \subset \mathcal{E}(X)$, parameterized by the elements of a set Γ, we denote by $\inf_{\gamma\in\Gamma} J_\gamma$ the function taking the value $\inf_{\gamma\in\Gamma} J_\gamma(x)$ at each $x \in X$.

For two functions $J_1, J_2 \in \mathcal{E}(X)$, we use the shorthand notation $J_1 \leq J_2$ to indicate the pointwise inequality

$$J_1(x) \leq J_2(x), \qquad \forall\ x \in X.$$

We use the shorthand notation $\inf_{i\in I} J_i$ to denote the function obtained by pointwise infimum of a collection $\{J_i \mid i \in I\} \subset \mathcal{E}(X)$, i.e.,

$$\left(\inf_{i\in I} J_i\right)(x) = \inf_{i\in I} J_i(x), \qquad \forall\ x \in X.$$

We use similar notation for sup.

Given subsets $S_1, S_2, S_3 \subset \mathcal{E}(X)$ and mappings $T_1 : S_1 \mapsto S_3$ and $T_2 : S_2 \mapsto S_1$, the *composition* of T_1 and T_2 is the mapping $T_1T_2 : S_2 \mapsto S_3$ defined by

$$(T_1T_2J)(x) = \big(T_1(T_2J)\big)(x), \qquad \forall\ J \in S_2,\ x \in X.$$

In particular, given a subset $S \subset \mathcal{E}(X)$ and mappings $T_1 : S \mapsto S$ and $T_2 : S \mapsto S$, the composition of T_1 and T_2 is the mapping $T_1T_2 : S \mapsto S$ defined by

$$(T_1T_2J)(x) = \big(T_1(T_2J)\big)(x), \qquad \forall\ J \in S,\ x \in X.$$

Similarly, given mappings $T_k : S \mapsto S$, $k = 1, \ldots, N$, their composition is the mapping $(T_1 \cdots T_N) : S \mapsto S$ defined by

$$(T_1T_2 \cdots T_NJ)(x) = \big(T_1(T_2(\cdots(T_NJ)))\big)(x), \qquad \forall\ J \in S,\ x \in X.$$

In our notation involving compositions we minimize the use of parentheses, as long as clarity is not compromised. In particular, we write T_1T_2J instead of (T_1T_2J) or $(T_1T_2)J$ or $T_1(T_2J)$, but we write $(T_1T_2J)(x)$ to indicate the value of T_1T_2J at $x \in X$.

If X and Y are nonempty sets, a mapping $T : S_1 \mapsto S_2$, where $S_1 \subset \mathcal{E}(X)$ and $S_2 \subset E(Y)$, is said to be *monotone* if for all $J, J' \in S_1$,

$$J \leq J' \qquad \Rightarrow \qquad TJ \leq TJ'.$$

Sequences of Functions

For a sequence of functions $\{J_k\} \subset \mathcal{E}(X)$ that converges pointwise, we denote by $\lim_{k\to\infty} J_k$ the pointwise limit of $\{J_k\}$. We denote by $\limsup_{k\to\infty} J_k$ (or $\liminf_{k\to\infty} J_k$) the pointwise limit superior (or inferior, respectively) of $\{J_k\}$. If $\{J_k\} \subset \mathcal{E}(X)$ converges pointwise to J, we write $J_k \to J$. Note that we reserve this notation for pointwise convergence. To denote convergence with respect to a norm $\|\cdot\|$, we write $\|J_k - J\| \to 0$.

A sequence of functions $\{J_k\} \subset \mathcal{E}(X)$ is said to be *monotonically nonincreasing* (or *monotonically nondecreasing*) if $J_{k+1} \leq J_k$ for all k (or $J_{k+1} \geq J_k$ for all k, respectively). Such a sequence always has a (pointwise) limit within $\mathcal{E}(X)$. We write $J_k \downarrow J$ (or $J_k \uparrow J$) to indicate that $\{J_k\}$ is monotonically nonincreasing (or monotonically nonincreasing, respectively) and that its limit is J.

Let $\{J_{mn}\} \subset \mathcal{E}(X)$ be a double indexed sequence, which is monotonically nonincreasing separately for each index in the sense that

$$J_{(m+1)n} \leq J_{mn}, \qquad J_{m(n+1)} \leq J_{mn}, \qquad \forall\ m, n = 0, 1, \ldots.$$

For such sequences, a useful fact is that

$$\lim_{m\to\infty} \left(\lim_{n\to\infty} J_{mn} \right) = \lim_{m\to\infty} J_{mm}.$$

There is a similar fact for monotonically nondecreasing sequences.

Expected Values

Given a random variable w defined over a probability space Ω, the expected value of w is defined by

$$E\{w\} = E\{w^+\} + E\{w^-\},$$

where w^+ and w^- are the positive and negative parts of w,

$$w^+(\omega) = \max\big\{0, w(\omega)\big\}, \qquad w^-(\omega) = \min\big\{0, w(\omega)\big\}.$$

In this way, taking also into account the rule $\infty - \infty = \infty$, the expected value $E\{w\}$ is well-defined if Ω is finite or countably infinite. In more general cases, $E\{w\}$ is similarly defined by the appropriate form of integration, and more detail will be given at specific points as needed.

APPENDIX B:

Contraction Mappings

B.1 CONTRACTION MAPPING FIXED POINT THEOREMS

The purpose of this appendix is to provide some background on contraction mappings and their properties. Let Y be a real vector space with a norm $\|\cdot\|$, i.e., a real-valued function satisfying for all $y \in Y$, $\|y\| \geq 0$, $\|y\| = 0$ if and only if $y = 0$, and

$$\|ay\| = |a|\|y\|, \quad \forall\ a \in \Re, \qquad \|y + z\| \leq \|y\| + \|z\|, \quad \forall\ y, z \in Y.$$

Let $\overline{Y}$ be a closed subset of Y. A function $F : \overline{Y} \mapsto \overline{Y}$ is said to be a *contraction mapping* if for some $\rho \in (0, 1)$, we have

$$\|Fy - Fz\| \leq \rho\|y - z\|, \qquad \forall\ y, z \in \overline{Y}.$$

The scalar ρ is called the *modulus of contraction* of F.

Example B.1 (Linear Contraction Mappings in $\Re^n$)

Consider the case of a linear mapping $F : \Re^n \mapsto \Re^n$ of the form

$$Fy = b + Ay,$$

where A is an $n \times n$ matrix and b is a vector in $\Re^n$. Let $\sigma(A)$ denote the *spectral radius* of A (the largest modulus among the moduli of the eigenvalues of A). Then it can be shown that A *is a contraction mapping with respect to some norm if and only if* $\sigma(A) < 1$.

Specifically, given $\epsilon > 0$, there exists a norm $\|\cdot\|_s$ such that

$$\|Ay\|_s \leq \big(\sigma(A) + \epsilon\big)\|y\|_s, \qquad \forall\ y \in \Re^n. \tag{B.1}$$

Thus, if $\sigma(A) < 1$ we may select $\epsilon > 0$ such that $\rho = \sigma(A) + \epsilon < 1$, and obtain the contraction relation

$$\|Fy - Fz\|_s = \big\|A(y - z)\big\|_s \leq \rho\|y - z\|_s, \qquad \forall\, y, z \in \Re^n. \tag{B.2}$$

The norm $\|\cdot\|_s$ can be taken to be a weighted Euclidean norm, i.e., it may have the form $\|y\|_s = \|My\|$, where M is a square invertible matrix, and $\|\cdot\|$ is the standard Euclidean norm, i.e., $\|x\| = \sqrt{x'x}$. †

Conversely, if Eq. (B.2) holds for some norm $\|\cdot\|_s$ and all real vectors y, z, it also holds for all complex vectors y, z with the squared norm $\|c\|_s^2$ of a complex vector c defined as the sum of the squares of the norms of the real and the imaginary components. Thus from Eq. (B.2), by taking $y - z = u$, where u is an eigenvector corresponding to an eigenvalue λ with $|\lambda| = \sigma(A)$, we have $\sigma(A)\|u\|_s = \|Au\|_s \leq \rho\|u\|_s$. Hence $\sigma(A) \leq \rho$, and it follows that if F is a contraction with respect to a given norm, we must have $\sigma(A) < 1$.

A sequence $\{y_k\} \subset Y$ is said to be a *Cauchy sequence* if $\|y_m - y_n\| \to 0$ as $m, n \to \infty$, i.e., given any $\epsilon > 0$, there exists N such that $\|y_m - y_n\| \leq \epsilon$ for all $m, n \geq N$. The space Y is said to be *complete* under the norm $\|\cdot\|$ if every Cauchy sequence $\{y_k\} \subset Y$ is convergent, in the sense that for some $y \in Y$, we have $\|y_k - y\| \to 0$. Note that a Cauchy sequence is always bounded. Also, a Cauchy sequence of real numbers is convergent, implying that the real line is a complete space and so is every real finite-dimensional vector space. On the other hand, an infinite dimensional space may not be complete under some norms, while it may be complete under other norms.

When Y is complete and $\overline{Y}$ is a closed subset of Y, an important property of a contraction mapping $F : \overline{Y} \mapsto \overline{Y}$ is that it has a unique fixed point within $\overline{Y}$, i.e., the equation

$$y = Fy$$

has a unique solution $y^* \in \overline{Y}$, called the *fixed point of* F. Furthermore, the sequence $\{y_k\}$ generated by the iteration

$$y_{k+1} = Fy_k$$

† We may show Eq. (B.1) by using the Jordan canonical form of A, which is denoted by J. In particular, if P is a nonsingular matrix such that $P^{-1}AP = J$ and D is the diagonal matrix with $1, \delta, \ldots, \delta^{n-1}$ along the diagonal, where $\delta > 0$, it is straightforward to verify that $D^{-1}P^{-1}APD = \hat{J}$, where $\hat{J}$ is the matrix that is identical to J except that each nonzero off-diagonal term is replaced by δ. Defining $\hat{P} = PD$, we have $A = \hat{P}\hat{J}\hat{P}^{-1}$. Now if $\|\cdot\|$ is the standard Euclidean norm, we note that for some $\beta > 0$, we have $\|\hat{J}z\| \leq \big(\sigma(A) + \beta\delta\big)\|z\|$ for all $z \in \Re^n$ and $\delta \in (0, 1]$. For a given $\delta \in (0, 1]$, consider the weighted Euclidean norm $\|\cdot\|_s$ defined by $\|y\|_s = \|\hat{P}^{-1}y\|$. Then we have for all $y \in \Re^n$,

$$\|Ay\|_s = \|\hat{P}^{-1}Ay\| = \|\hat{P}^{-1}\hat{P}\hat{J}\hat{P}^{-1}y\| = \|\hat{J}\hat{P}^{-1}y\| \leq \big(\sigma(A) + \beta\delta\big)\|\hat{P}^{-1}y\|,$$

so that $\|Ay\|_s \leq \big(\sigma(A) + \beta\delta\big)\|y\|_s$, for all $y \in \Re^n$. For a given $\epsilon > 0$, we choose $\delta = \epsilon/\beta$, so the preceding relation yields Eq. (B.1).

converges to y^*, starting from an arbitrary initial point y_0.

Proposition B.1: (Contraction Mapping Fixed-Point Theorem) Let Y be a complete vector space and let $\overline{Y}$ be a closed subset of Y. Then if $F : \overline{Y} \mapsto \overline{Y}$ is a contraction mapping with modulus $\rho \in (0,1)$, there exists a unique $y^* \in \overline{Y}$ such that

$$y^* = Fy^*.$$

Furthermore, the sequence $\{F^k y\}$ converges to y^* for any $y \in \overline{Y}$, and we have

$$\|F^k y - y^*\| \leq \rho^k \|y - y^*\|, \qquad k = 1, 2, \ldots.$$

Proof: Let $y \in \overline{Y}$ and consider the iteration $y_{k+1} = Fy_k$ starting with $y_0 = y$. By the contraction property of F,

$$\|y_{k+1} - y_k\| \leq \rho \|y_k - y_{k-1}\|, \qquad k = 1, 2, \ldots,$$

which implies that

$$\|y_{k+1} - y_k\| \leq \rho^k \|y_1 - y_0\|, \qquad k = 1, 2, \ldots.$$

It follows that for every $k \geq 0$ and $m \geq 1$, we have

$$\begin{aligned} \|y_{k+m} - y_k\| &\leq \sum_{i=1}^{m} \|y_{k+i} - y_{k+i-1}\| \\ &\leq \rho^k (1 + \rho + \cdots + \rho^{m-1}) \|y_1 - y_0\| \\ &\leq \frac{\rho^k}{1 - \rho} \|y_1 - y_0\|. \end{aligned}$$

Therefore, $\{y_k\}$ is a Cauchy sequence in $\overline{Y}$ and must converge to a limit $y^* \in \overline{Y}$, since Y is complete and $\overline{Y}$ is closed. We have for all $k \geq 1$,

$$\|Fy^* - y^*\| \leq \|Fy^* - y_k\| + \|y_k - y^*\| \leq \rho \|y^* - y_{k-1}\| + \|y_k - y^*\|$$

and since y_k converges to y^*, we obtain $Fy^* = y^*$. Thus, the limit y^* of y_k is a fixed point of F. It is a unique fixed point because if $\tilde{y}$ were another fixed point, we would have

$$\|y^* - \tilde{y}\| = \|Fy^* - F\tilde{y}\| \leq \rho \|y^* - \tilde{y}\|,$$

which implies that $y^* = \tilde{y}$.

To show the convergence rate bound of the last part, note that

$$\|F^k y - y^*\| = \big\|F^k y - Fy^*\big\| \le \rho\|F^{k-1}y - y^*\|.$$

Repeating this process for a total of k times, we obtain the desired result. **Q.E.D.**

The convergence rate exhibited by $F^k y$ in the preceding proposition is said to be *geometric*, and $F^k y$ is said to converge to its limit y^* *geometrically*. This is in reference to the fact that the error $\|F^k y - y^*\|$ converges to 0 faster than some geometric progression ($\rho^k\|y - y^*\|$ in this case).

In some contexts of interest to us one may encounter mappings that are not contractions, but become contractions when iterated a finite number of times. In this case, one may use a slightly different version of the contraction mapping fixed point theorem, which we now present.

We say that a function $F : Y \mapsto Y$ is an *m-stage contraction mapping* if there exists a positive integer m and some $\rho < 1$ such that

$$\big\|F^m y - F^m y'\big\| \le \rho\|y - y'\|, \qquad \forall\ y, y' \in Y,$$

where F^m denotes the composition of F with itself m times. Thus, F is an m-stage contraction if F^m is a contraction. Again, the scalar ρ is called the modulus of contraction. We have the following generalization of Prop. B.1.

Proposition B.2: (m-Stage Contraction Mapping Fixed-Point Theorem) Let Y be a complete vector space and let $\overline{Y}$ be a closed subset of Y. Then if $F : \overline{Y} \mapsto \overline{Y}$ is an m-stage contraction mapping with modulus $\rho \in (0, 1)$, there exists a unique $y^* \in \overline{Y}$ such that

$$y^* = Fy^*.$$

Furthermore, $\{F^k y\}$ converges to y^* for any $y \in \overline{Y}$.

Proof: Since F^m maps $\overline{Y}$ into $\overline{Y}$ and is a contraction mapping, by Prop. B.1, it has a unique fixed point in $\overline{Y}$, denoted y^*. Applying F to both sides of the relation $y^* = F^m y^*$, we see that Fy^* is also a fixed point of F^m, so by the uniqueness of the fixed point, we have $y^* = Fy^*$. Therefore y^* is a fixed point of F. If F had another fixed point, say $\tilde{y}$, then we would have $\tilde{y} = F^m\tilde{y}$, which by the uniqueness of the fixed point of F^m implies that $\tilde{y} = y^*$. Thus, y^* is the unique fixed point of F.

To show the convergence of $\{F^k y\}$, note that by Prop. B.1, we have for all $y \in \overline{Y}$,

$$\lim_{k\to\infty} \|F^{mk} y - y^*\| = 0.$$

Using $F^\ell y$ in place of y, we obtain

$$\lim_{k\to\infty} \|F^{mk+\ell}y - y^*\| = 0, \qquad \ell = 0, 1, \ldots, m-1,$$

which proves the desired result. **Q.E.D.**

B.2 WEIGHTED SUP-NORM CONTRACTIONS

In this section, we will focus on contraction mappings within a specialized context that is particularly important in DP. Let X be a set (typically the state space in DP), and let $v : X \mapsto \Re$ be a positive-valued function,

$$v(x) > 0, \qquad \forall\, x \in X.$$

Let $\mathcal{B}(X)$ denote the set of all functions $J : X \mapsto \Re$ such that $J(x)/v(x)$ is bounded as x ranges over X. We define a norm on $\mathcal{B}(X)$, called the *weighted sup-norm*, by

$$\|J\| = \sup_{x\in X} \frac{\big|J(x)\big|}{v(x)}. \tag{B.3}$$

It is easily verified that $\|\cdot\|$ thus defined has the required properties for being a norm. Furthermore, $\mathcal{B}(X)$ *is complete under this norm*. To see this, consider a Cauchy sequence $\{J_k\} \subset \mathcal{B}(X)$, and note that $\|J_m - J_n\| \to 0$ as $m, n \to \infty$ implies that for all $x \in X$, $\{J_k(x)\}$ is a Cauchy sequence of real numbers, so it converges to some $J^*(x)$. We will show that $J^* \in \mathcal{B}(X)$ and that $\|J_k - J^*\| \to 0$. To this end, it will be sufficient to show that given any $\epsilon > 0$, there exists an integer K such that

$$\frac{\big|J_k(x) - J^*(x)\big|}{v(x)} \le \epsilon, \qquad \forall\, x \in X,\ k \ge K.$$

This will imply that

$$\sup_{x\in X} \frac{\big|J^*(x)\big|}{v(x)} \le \epsilon + \|J_k\|, \qquad \forall\, k \ge K,$$

so that $J^* \in \mathcal{B}(X)$, and will also imply that $\|J_k - J^*\| \le \epsilon$, so that $\|J_k - J^*\| \to 0$. Assume the contrary, i.e., that there exists an $\epsilon > 0$ and a subsequence $\{x_{m_1}, x_{m_2}, \ldots\} \subset X$ such that $m_i < m_{i+1}$ and

$$\epsilon < \frac{\big|J_{m_i}(x_{m_i}) - J^*(x_{m_i})\big|}{v(x_{m_i})}, \qquad \forall\, i \ge 1.$$

The right-hand side above is less or equal to

$$\frac{\big|J_{m_i}(x_{m_i}) - J_n(x_{m_i})\big|}{v(x_{m_i})} + \frac{\big|J_n(x_{m_i}) - J^*(x_{m_i})\big|}{v(x_{m_i})}, \qquad \forall\, n \ge 1,\ i \ge 1.$$

The first term in the above sum is less than $\epsilon/2$ for i and n larger than some threshold; fixing i and letting n be sufficiently large, the second term can also be made less than $\epsilon/2$, so the sum is made less than ϵ - a contradiction. In conclusion, the space $\mathcal{B}(X)$ is complete, so the fixed point results of Props. B.1 and B.2 apply.

In our discussions, unless we specify otherwise, we will assume that $\mathcal{B}(X)$ is equipped with the weighted sup-norm above, where the weight function v will be clear from the context. There will be frequent occasions where the norm will be unweighted, i.e., $v(x) \equiv 1$ and $\|J\| = \sup_{x\in X} \big|J(x)\big|$, in which case we will explicitly state so.

Finite-Dimensional Cases

Let us now focus on the finite-dimensional case $X = \{1,\ldots,n\}$, in which case $\mathcal{R}(X)$ and $\mathcal{B}(X)$ can be identified with $\Re^n$. We first consider a linear mapping (cf. Example B.1). We have the following proposition.

Proposition B.3: Consider a linear mapping $F : \Re^n \mapsto \Re^n$ of the form

$$Fy = b + Ay,$$

where A is an $n \times n$ matrix with components a_{ij}, and b is a vector in $\Re^n$. Denote by $|A|$ the matrix whose components are the absolute values of the components of A and let $\sigma(A)$ and $\sigma\big(|A|\big)$ denote the spectral radii of A and $|A|$, respectively. Then:

(a) $|A|$ has a real eigenvalue λ, which is equal to its spectral radius, and an associated nonnegative eigenvector.

(b) F is a contraction with respect to some weighted sup-norm if and only if $\sigma\big(|A|\big) < 1$. In particular, any substochastic matrix P ($p_{ij} \geq 0$ for all i, j, and $\sum_{j=1}^n p_{ij} \leq 1$, for all i) is a contraction with respect to some weighted sup-norm if and only if $\sigma(P) < 1$.

(c) F is a contraction with respect to the weighted sup-norm

$$\|y\| = \max_{i=1,\ldots,n} \frac{|y_i|}{v(i)}$$

if and only if

$$\frac{\sum_{j=1}^n |a_{ij}|\, v(j)}{v(i)} < 1, \qquad i = 1,\ldots,n.$$

Proof: (a) This is the Perron-Frobenius Theorem; see e.g., [BeT89], Chapter 2, Prop. 6.6.

(b) This follows from the Perron-Frobenius Theorem; see [BeT89], Ch. 2, Cor. 6.2.

(c) This is proved in more general form in the following Prop. B.4. **Q.E.D.**

Consider next a nonlinear mapping $F : \Re^n \mapsto \Re^n$ that has the property

$$|Fy - Fz| \leq P\,|y - z|, \qquad \forall\ y, z \in \Re^n,$$

for some matrix P with nonnegative components and $\sigma(P) < 1$. Here, we generically denote by $|w|$ the vector whose components are the absolute values of the components of w, and the inequality is componentwise. Then we claim that F is a contraction with respect to some weighted sup-norm. To see this note that by the preceding discussion, P is a contraction with respect to some weighted sup-norm $\|y\| = \max_{i=1,\ldots,n} |y_i|/v(i)$, and we have

$$\frac{\big(|Fy - Fz|\big)(i)}{v(i)} \leq \frac{\big(P\,|y - z|\big)(i)}{v(i)} \leq \alpha\,\|y - z\|, \qquad \forall\ i = 1, \ldots, n,$$

for some $\alpha \in (0, 1)$, where $\big(|Fy - Fz|\big)(i)$ and $\big(P\,|y - z|\big)(i)$ are the ith components of the vectors $|Fy - Fz|$ and $P\,|y - z|$, respectively. Thus, F is a contraction with respect to $\|\cdot\|$. For additional discussion of linear and nonlinear contraction mapping properties and characterizations such as the one above, see the book [OrR70].

Linear Mappings on Countable Spaces

The case where X is countable (or, as a special case, finite) is frequently encountered in DP. The following proposition provides some useful criteria for verifying the contraction property of mappings that are either linear or are obtained via a parametric minimization of other contraction mappings.

Proposition B.4: Let $X = \{1, 2, \ldots\}$.

(a) Let $F : \mathcal{B}(X) \mapsto \mathcal{B}(X)$ be a linear mapping of the form

$$(FJ)(i) = b_i + \sum_{j \in X} a_{ij} J(j), \qquad i \in X,$$

where b_i and a_{ij} are some scalars. Then F is a contraction with modulus ρ with respect to the weighted sup-norm (B.3) if and only if

$$\frac{\sum_{j\in X} |a_{ij}|\, v(j)}{v(i)} \le \rho, \qquad i \in X. \tag{B.4}$$

(b) Let $F : \mathcal{B}(X) \mapsto \mathcal{B}(X)$ be a mapping of the form

$$(FJ)(i) = \inf_{\mu\in M} (F_\mu J)(i), \qquad i \in X,$$

where M is parameter set, and for each $\mu \in M$, F_μ is a contraction mapping from $\mathcal{B}(X)$ to $\mathcal{B}(X)$ with modulus ρ. Then F is a contraction mapping with modulus ρ.

Proof: (a) Assume that Eq. (B.4) holds. For any $J, J' \in \mathcal{B}(X)$, we have

$$\begin{aligned}
\|FJ - FJ'\| &= \sup_{i\in X} \frac{\left|\sum_{j\in X} a_{ij}\big(J(j) - J'(j)\big)\right|}{v(i)} \\
&\le \sup_{i\in X} \frac{\sum_{j\in X} \big|a_{ij}\big|\, v(j) \Big(\big|J(j) - J'(j)\big|/v(j)\Big)}{v(i)} \\
&\le \sup_{i\in X} \frac{\sum_{j\in X} \big|a_{ij}\big|\, v(j)}{v(i)}\, \big\|J - J'\big\| \\
&\le \rho\, \|J - J'\|,
\end{aligned}$$

where the last inequality follows from the hypothesis.

Conversely, arguing by contradiction, let's assume that Eq. (B.4) is violated for some $i \in X$. Define $J(j) = v(j)\,\mathrm{sgn}(a_{ij})$ and $J'(j) = 0$ for all $j \in X$. Then we have $\|J - J'\| = \|J\| = 1$, and

$$\frac{\big|(FJ)(i) - (FJ')(i)\big|}{v(i)} = \frac{\sum_{j\in X} |a_{ij}|\, v(j)}{v(i)} > \rho = \rho\, \|J - J'\|,$$

showing that F is not a contraction of modulus ρ.

(b) Since F_μ is a contraction of modulus ρ, we have for any $J, J' \in \mathcal{B}(X)$,

$$\frac{(F_\mu J)(i)}{v(i)} \le \frac{(F_\mu J')(i)}{v(i)} + \rho\, \|J - J'\|, \qquad i \in X,$$

so by taking the infimum over $\mu \in M$,

$$\frac{(FJ)(i)}{v(i)} \le \frac{(FJ')(i)}{v(i)} + \rho\, \|J - J'\|, \qquad i \in X.$$

Reversing the roles of J and J', we obtain

$$\frac{\big|(FJ)(i) - (FJ')(i)\big|}{v(i)} \leq \rho \, \|J - J'\|, \qquad i \in X,$$

and by taking the supremum over i, the contraction property of F is proved. **Q.E.D.**

The preceding proposition assumes that $FJ \in \mathcal{B}(X)$ for all $J \in \mathcal{B}(X)$. The following proposition provides conditions, particularly relevant to the DP context, which imply this assumption.

Proposition B.5: Let $X = \{1, 2, \ldots\}$, let M be a parameter set, and for each $\mu \in M$, let F_μ be a linear mapping of the form

$$(F_\mu J)(i) = b_i(\mu) + \sum_{j \in X} a_{ij}(\mu)\, J(j), \qquad i \in X.$$

(a) We have $F_\mu J \in \mathcal{B}(X)$ for all $J \in \mathcal{B}(X)$ provided $b(\mu) \in \mathcal{B}(X)$ and $V(\mu) \in \mathcal{B}(X)$, where

$$b(\mu) = \big\{b_1(\mu), b_2(\mu), \ldots\big\}, \qquad V(\mu) = \big\{V_1(\mu), V_2(\mu), \ldots\big\},$$

with

$$V_i(\mu) = \sum_{j \in X} \big|a_{ij}(\mu)\big|\, v(j), \qquad i \in X.$$

(b) Consider the mapping F

$$(FJ)(i) = \inf_{\mu \in M} (F_\mu J)(i), \qquad i \in X.$$

We have $FJ \in \mathcal{B}(X)$ for all $J \in \mathcal{B}(X)$, provided $b \in \mathcal{B}(X)$ and $V \in \mathcal{B}(X)$, where

$$b = \big\{b_1, b_2, \ldots\big\}, \qquad V = \{V_1, V_2, \ldots\},$$

with $b_i = \sup_{\mu \in M} b_i(\mu)$ and $V_i = \sup_{\mu \in M} V_i(\mu)$.

Proof: (a) For all $\mu \in M$, $J \in \mathcal{B}(X)$ and $i \in X$, we have

$$(F_\mu J)(i) \leq \big|b_i(\mu)\big| + \sum_{j \in X} \big|a_{ij}(\mu)\big|\, \big|J(j)/v(j)\big|\, v(j)$$

$$\leq \big|b_i(\mu)\big| + \|J\| \sum_{j \in X} \big|a_{ij}(\mu)\big|\, v(j)$$
$$= \big|b_i(\mu)\big| + \|J\|\, V_i(\mu),$$

and similarly $(F_\mu J)(i) \geq -\big|b_i(\mu)\big| - \|J\|\, V_i(\mu)$. Thus

$$\big|(F_\mu J)(i)\big| \leq \big|b_i(\mu)\big| + \|J\|\, V_i(\mu), \qquad i \in X.$$

By dividing this inequality with $v(i)$ and by taking the supremum over $i \in X$, we obtain

$$\|F_\mu J\| \leq \|b_\mu\| + \|J\|\, \|V_\mu\| < \infty.$$

(b) By doing the same as in (a), but after first taking the infimum of $(F_\mu J)(i)$ over μ, we obtain

$$\|FJ\| \leq \|b\| + \|J\|\, \|V\| < \infty.$$

Q.E.D.

References

[ABB02] Abounadi, J., Bertsekas, B. P., and Borkar, V. S., 2002. "Stochastic Approximation for Non-Expansive Maps: Q-Learning Algorithms," SIAM J. on Control and Opt., Vol. 41, pp. 1-22.

[AnM79] Anderson, B. D. O., and Moore, J. B., 1979. Optimal Filtering, Prentice Hall, Englewood Cliffs, N. J.

[BBB08] Basu, A., Bhattacharyya, and Borkar, V., 2008. "A Learning Algorithm for Risk-Sensitive Cost," Math. of OR, Vol. 33, pp. 880-898.

[BBD10] Busoniu, L., Babuska, R., De Schutter, B., and Ernst, D., 2010. Reinforcement Learning and Dynamic Programming Using Function Approximators, CRC Press, N. Y.

[Bau78] Baudet, G. M., 1978. "Asynchronous Iterative Methods for Multiprocessors," Journal of the ACM, Vol. 25, pp. 226-244.

[BeI96] Bertsekas, D. P., and Ioffe, S., 1996. "Temporal Differences-Based Policy Iteration and Applications in Neuro-Dynamic Programming," Lab. for Info. and Decision Systems Report LIDS-P-2349, MIT.

[BeS78] Bertsekas, D. P., and Shreve, S. E., 1978. Stochastic Optimal Control: The Discrete Time Case, Academic Press, N. Y.; may be downloaded from http://web.mit.edu/dimitrib/www/home.html

[BeT89] Bertsekas, D. P., and Tsitsiklis, J. N., 1989. Parallel and Distributed Computation: Numerical Methods, Prentice-Hall, Engl. Cliffs, N. J.; may be downloaded from http://web.mit.edu/dimitrib/www/home.html

[BeT91] Bertsekas, D. P., and Tsitsiklis, J. N., 1991. "An Analysis of Stochastic Shortest Path Problems," Math. of OR, Vol. 16, pp. 580-595.

[BeT96] Bertsekas, D. P., and Tsitsiklis, J. N., 1996. Neuro-Dynamic Programming, Athena Scientific, Belmont, MA.

[BeT08] Bertsekas, D. P., and Tsitsiklis, J. N., 2008. Introduction to Probability, 2nd Ed., Athena Scientific, Belmont, MA.

[BeY07] Bertsekas, D. P., and Yu, H., 2007. "Solution of Large Systems of Equations Using Approximate Dynamic Programming Methods," Lab. for Info. and Decision Systems Report LIDS-P-2754, MIT.

[BeY09] Bertsekas, D. P., and Yu, H., 2009. "Projected Equation Methods for Approximate Solution of Large Linear Systems," J. of Computational and Applied Mathematics, Vol. 227, pp. 27-50.

[BeY10] Bertsekas, D. P., and Yu, H., 2010. "Asynchronous Distributed Policy Iteration in Dynamic Programming," Proc. of Allerton Conf. on Communication, Control and Computing, Allerton Park, Ill, pp. 1368-1374.

[BeY12] Bertsekas, D. P., and Yu, H., 2012. "Q-Learning and Enhanced Policy Iteration in Discounted Dynamic Programming," Math. of OR, Vol. 37, pp. 66-94.

[BeY16] Bertsekas, D. P., and Yu, H., 2016. "Stochastic Shortest Path Problems Under Weak Conditions," Lab. for Information and Decision Systems Report LIDS-2909, January 2016.

[Ber71] Bertsekas, D. P., 1971. "Control of Uncertain Systems With a Set-Membership Description of the Uncertainty," Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, MA (available from the author's website).

[Ber72] Bertsekas, D. P., 1972. "Infinite Time Reachability of State Space Regions by Using Feedback Control," IEEE Trans. Aut. Control, Vol. AC-17, pp. 604-613.

[Ber75] Bertsekas, D. P., 1975. "Monotone Mappings in Dynamic Programming," 1975 IEEE Conference on Decision and Control, pp. 20-25.

[Ber77] Bertsekas, D. P., 1977. "Monotone Mappings with Application in Dynamic Programming," SIAM J. on Control and Opt., Vol. 15, pp. 438-464.

[Ber82] Bertsekas, D. P., 1982. "Distributed Dynamic Programming," IEEE Trans. Aut. Control, Vol. AC-27, pp. 610-616.

[Ber83] Bertsekas, D. P., 1983. "Asynchronous Distributed Computation of Fixed Points," Math. Programming, Vol. 27, pp. 107-120.

[Ber87] Bertsekas, D. P., 1987. Dynamic Programming: Deterministic and Stochastic Models, Prentice-Hall, Englewood Cliffs, N. J.

[Ber96] Bertsekas, D. P., 1996. Lecture at NSF Workshop on Reinforcement Learning, Hilltop House, Harper's Ferry, N. Y.

[Ber98] Bertsekas, D. P., 1998. Network Optimization: Continuous and Discrete Models, Athena Scientific, Belmont, MA.

[Ber09] Bertsekas, D. P., 2009. Convex Optimization Theory, Athena Scientific, Belmont, MA.

[Ber10] Bertsekas, D. P., 2010. "Williams-Baird Counterexample for Q-Factor Asynchronous Policy Iteration,"
http://web.mit.edu/dimitrib/www/Williams-Baird Counterexample.pdf

[Ber11a] Bertsekas, D. P., 2011. "Temporal Difference Methods for General Projected Equations," IEEE Trans. Aut. Control, Vol. 56, pp. 2128-2139.

[Ber11b] Bertsekas, D. P., 2011. "λ-Policy Iteration: A Review and a New Implementation," Lab. for Info. and Decision Systems Report LIDS-P-2874, MIT; appears in Reinforcement Learning and Approximate Dynamic Programming for Feedback Control, by F. Lewis and D. Liu (eds.), IEEE Press, 2012.

[Ber11c] Bertsekas, D. P., 2011. "Approximate Policy Iteration: A Survey and Some New Methods," J. of Control Theory and Applications, Vol. 9, pp. 310-335; a somewhat expanded version appears as Lab. for Info. and Decision Systems Report LIDS-2833, MIT, 2011.

[Ber12a] Bertsekas, D. P., 2012. Dynamic Programming and Optimal Control, Vol. II, 4th Edition: Approximate Dynamic Programming, Athena Scientific, Belmont, MA.

[Ber12b] Bertsekas, D. P., 2012. "Weighted Sup-Norm Contractions in Dynamic Programming: A Review and Some New Applications," Lab. for Info. and Decision Systems Report LIDS-P-2884, MIT.

[Ber14] Bertsekas, D. P., 2014. "Robust Shortest Path Planning and Semicontractive Dynamic Programming," Lab. for Information and Decision Systems Report LIDS-P-2915, MIT, Feb. 2014 (revised Jan. 2015 and June 2016); arXiv preprint arXiv:1608.01670; to appear in Naval Research Logistics.

[Ber15] Bertsekas, D. P., 2015. "Regular Policies in Abstract Dynamic Programming," Lab. for Information and Decision Systems Report LIDS-P-3173, MIT, May 2015; arXiv preprint arXiv:1609.03115; to appear in SIAM J. on Control and Optimization.

[Ber16a] Bertsekas, D. P., 2016. "Affine Monotonic and Risk-Sensitive Models in Dynamic Programming," Lab. for Information and Decision Systems Report LIDS-3204, MIT, June 2016; arXiv preprint arXiv:1608.01393.

[Ber16b] Bertsekas, D. P., 2016. "Proximal Algorithms and Temporal Differences for Large Linear Systems: Extrapolation, Approximation, and Simulation," Report LIDS-P-3205, MIT, Oct. 2016; arXiv preprint arXiv:1610.1610.05427.

[Ber16c] Bertsekas, D. P., 2016. Nonlinear Programming, 3rd Edition, Athena Scientific, Belmont, MA.

[Ber17a] Bertsekas, D. P., 2017. Dynamic Programming and Optimal Control, Vol. I, 4th Edition, Athena Scientific, Belmont, MA.

[Ber17b] Bertsekas, D. P., 2017. "Value and Policy Iteration in Deterministic Optimal Control and Adaptive Dynamic Programming," IEEE Transactions on Neural Networks and Learning Systems, Vol. 28, pp. 500-509.

[Ber17c] Bertsekas, D. P., 2017. "Stable Optimal Control and Semicontractive Dynamic Programming," Report LIDS-P-3506, MIT, May 2017; to appear in SIAM J. on Control and Optimization.

[Ber17d] Bertsekas, D. P., 2017. "Proper Policies in Infinite-State Stochastic Shortest Path Problems", Report LIDS-P-3507, MIT, May 2017; to appear in IEEE Transactions on Aut. Control.

[Ber17e] Bertsekas, D. P., 2017. "Proximal Algorithms and Temporal Differences for Solving Fixed Point Problems," to appear in Computational Optimization and Applications J.

[Bla65] Blackwell, D., 1965. "Positive Dynamic Programming," Proc. Fifth Berkeley Symposium Math. Statistics and Probability, pp. 415-418.

[BoM99] Borkar, V. S., Meyn, S. P., 1999. "Risk Sensitive Optimal Control: Existence and Synthesis for Models with Unbounded Cost," SIAM J. Control and Opt., Vol. 27, pp. 192-209.

[BoM00] Borkar, V. S., Meyn, S. P., 1900. "The O.D.E. Method for Convergence of Stochastic Approximation and Reinforcement Learning," SIAM J. Control and Opt., Vol. 38, pp. 447-469.

[BoM02] Borkar, V. S., Meyn, S. P., 2002. "Risk-Sensitive Optimal Control for Markov Decision Processes with Monotone Cost," Math. of OR, Vol. 27, pp. 192-209.

[Bor98] Borkar, V. S., 1998. "Asynchronous Stochastic Approximation," SIAM J. Control Opt., Vol. 36, pp. 840-851.

[Bor08] Borkar, V. S., 2008. Stochastic Approximation: A Dynamical Systems Viewpoint, Cambridge Univ. Press, N. Y.

[CFH07] Chang, H. S., Fu, M. C., Hu, J., Marcus, S. I., 2007. Simulation-Based Algorithms for Markov Decision Processes, Springer, N. Y.

[CaM88] Carraway, R. L., and Morin, T. L., 1988. "Theory and Applications of Generalized Dynamic Programming: An Overview," Computers and Mathematics with Applications, Vol. 16, pp. 779-788.

[CaR13] Canbolat, P. G., and Rothblum, U. G., 2013. "(Approximate) Iterated Successive Approximations Algorithm for Sequential Decision Processes," Annals of Operations Research, Vol. 208, pp. 309-320.

[Cao07] Cao, X. R., 2007. Stochastic Learning and Optimization: A Sensitivity-Based Approach, Springer, N. Y.

[ChM69] Chazan D., and Miranker, W., 1969. "Chaotic Relaxation," Linear Algebra and Applications, Vol. 2, pp. 199-222.

[ChS87] Chung, K.-J., and Sobel, M. J., 1987. "Discounted MDPs: Distribution Functions and Exponential Utility Maximization," SIAM J. Control and Opt., Vol. 25, pp. 49-62.

[CoM99] Coraluppi, S. P., and Marcus, S. I., 1999. "Risk-Sensitive and Minimax Control of Discrete-Time, Finite-State Markov Decision Processes," Automatica, Vol. 35, pp. 301-309.

[DFV00] de Farias, D. P., and Van Roy, B., 2000. "On the Existence of Fixed Points for Approximate Value Iteration and Temporal-Difference Learning," J. of Optimization Theory and Applications, Vol. 105, pp. 589-608.

[DeM67] Denardo, E. V., and Mitten, L. G., 1967. "Elements of Sequential Decision Processes," J. Indust. Engrg., Vol. 18, pp. 106-112.

[DeR79] Denardo, E. V., and Rothblum, U. G., 1979. "Optimal Stopping, Exponential Utility, and Linear Programming," Math. Programming, Vol. 16, pp. 228-244.

[Den67] Denardo, E. V., 1967. "Contraction Mappings in the Theory Underlying Dynamic Programming," SIAM Review, Vol. 9, pp. 165-177.

[Der70] Derman, C., 1970. Finite State Markovian Decision Processes, Academic Press, N. Y.

[DuS65] Dubins, L., and Savage, L. M., 1965. How to Gamble If You Must, McGraw-Hill, N. Y.

[FeM97] Fernandez-Gaucherand, E., and Marcus, S. I., 1997. "Risk-Sensitive Optimal Control of Hidden Markov Models: Structural Results," IEEE Trans. Aut. Control, Vol. AC-42, pp. 1418-1422.

[Fei02] Feinberg, E. A., 2002. "Total Reward Criteria," in E. A. Feinberg and A. Shwartz, (Eds.), Handbook of Markov Decision Processes, Springer, N. Y.

[FiV96] Filar, J., and Vrieze, K., 1996. Competitive Markov Decision Processes, Springer, N. Y.

[FlM95] Fleming, W. H., and McEneaney, W. M., 1995. "Risk-Sensitive Control on an Infinite Time Horizon," SIAM J. Control and Opt., Vol. 33, pp. 1881-1915.

[Gos03] Gosavi, A., 2003. Simulation-Based Optimization: Parametric Optimization Techniques and Reinforcement Learning, Springer, N. Y.

[GuS17] Guillot, M., and Stauffer, G., 2017. "The Stochastic Shortest Path Problem: A Polyhedral Combinatorics Perspective," Univ. of Grenoble Report.

[HCP99] Hernandez-Lerma, O., Carrasco, O., and Perez-Hernandez. 1999. "Markov Control Processes with the Expected Total Cost Criterion: Optimality, Stability, and Transient Models," Acta Appl. Math., Vol. 59, pp. 229-269.

[Hay08] Haykin, S., 2008. Neural Networks and Learning Machines, (3rd Edition), Prentice-Hall, Englewood-Cliffs, N. J.

[HeL99] Hernandez-Lerma, O., and Lasserre, J. B., 1999. Further Topics on Discrete-Time Markov Control Processes, Springer, N. Y.

[HeM96] Hernandez-Hernandez, D., and Marcus, S. I., 1996. "Risk Sensitive Control of Markov Processes in Countable State Space," Systems and Control Letters, Vol. 29, pp. 147-155.

[HiW05] Hinderer, K., and Waldmann, K.-H., 2005. "Algorithms for Countable State Markov Decision Models with an Absorbing Set," SIAM J. of Control and Opt., Vol. 43, pp. 2109-2131.

[HoM72] Howard, R. S., and Matheson, J. E., 1972. "Risk-Sensitive Markov Decision Processes," Management Science, Vol. 8, pp. 356-369.

[JBE94] James, M. R., Baras, J. S., Elliott, R. J., 1994. "Risk-Sensitive Control and Dynamic Games for Partially Observed Discrete-Time Nonlinear Systems," IEEE Trans. Aut. Control, Vol. AC-39, pp. 780-792.

[JaC06] James, H. W., and Collins, E. J., 2006. "An Analysis of Transient Markov Decision Processes," J. Appl. Prob., Vol. 43, pp. 603-621.

[Jac73] Jacobson, D. H., 1973. "Optimal Stochastic Linear Systems with Exponential Performance Criteria and their Relation to Deterministic Differential Games," IEEE Transactions on Automatic Control, Vol. AC-18, pp. 124-131.

[Kal60] Kalman, R. E., 1960. "Contributions to the Theory of Optimal Control," Bol. Soc. Mat. Mexicana, Vol. 5, pp.102-119.

[Kuc72] Kucera, V., 1972. "The Discrete Riccati Equation of Optimal Control," Kybernetika, Vol. 8, pp. 430-447.

[Kuc73] Kucera, V., 1973. "A Review of the Matrix Riccati Equation," Kybernetika, Vol. 9, pp. 42-61.

[LaR95] Lancaster, P., and Rodman, L., 1995. Algebraic Riccati Equations, Clarendon Press, Oxford, UK.

[Mey07] Meyn, S., 2007. Control Techniques for Complex Networks, Cambridge Univ. Press, N. Y.

[Mit64] Mitten, L. G., 1964. "Composition Principles for Synthesis of Optimal Multistage Processes," Operations Research, Vol. 12, pp. 610-619.

[Mit74] Mitten, L. G., 1964. "Preference Order Dynamic Programming," Management Science, Vol. 21, pp. 43 - 46.

[Mor82] Morin, T. L., 1982. "Monotonicity and the Principle of Optimality," J. of Math. Analysis and Applications, Vol. 88, pp. 665-674.

[OrR70] Ortega, J. M., and Rheinboldt, W. C., 1970. Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, N. Y.

[PaB99] Patek, S. D., and Bertsekas, D. P., 1999. "Stochastic Shortest Path Games," SIAM J. on Control and Opt., Vol. 36, pp. 804-824.

[Pal67] Pallu de la Barriere, R., 1967. Optimal Control Theory, Saunders, Phila; republished by Dover, N. Y., 1980.

[Pat01] Patek, S. D., 2001. "On Terminating Markov Decision Processes with a Risk Averse Objective Function," Automatica, Vol. 37, pp. 1379-1386.

[Pat07] Patek, S. D., 2007. "Partially Observed Stochastic Shortest Path Problems with Approximate Solution by Neuro-Dynamic Programming," IEEE Trans. on Systems, Man, and Cybernetics Part A, Vol. 37, pp. 710-720.

[Pli78] Pliska, S. R., 1978. "On the Transient Case for Markov Decision Chains with General State Spaces," in Dynamic Programming and its Applications, by M. L. Puterman (ed.), Academic Press, N. Y.

[Pow07] Powell, W. B., 2007. Approximate Dynamic Programming: Solving the Curses of Dimensionality, J. Wiley and Sons, Hoboken, N. J; 2nd ed., 2011.

[Put94] Puterman, M. L., 1994. Markovian Decision Problems, J. Wiley, N. Y.

[Rei16] Reissig, G., 2016. "Approximate Value Iteration for a Class of Deterministic Optimal Control Problems with Infinite State and Input Alphabets," Proc. 2016 IEEE Conf. on Decision and Control, pp. 1063-1068.

[Roc70] Rockafellar, R. T., 1970. Convex Analysis, Princeton Univ. Press, Princeton, N. J.

[Ros67] Rosenfeld, J., 1967. "A Case Study on Programming for Parallel Processors," Research Report RC-1864, IBM Res. Center, Yorktown Heights, N. Y.

[Rot79] Rothblum, U. G., 1979. "Iterated Successive Approximation for Sequential Decision Processes," in Stochastic Control and Optimization, by J. W. B. van Overhagen and H. C. Tijms (eds), Vrije University, Amsterdam.

[Rot84] Rothblum, U. G., 1984. "Multiplicative Markov Decision Chains," Math. of OR, Vol. 9, pp. 6-24.

[ScL12] Scherrer, B., and Lesner, B., 2012. "On the Use of Non-Stationary Policies for Stationary Infinite-Horizon Markov Decision Processes," NIPS 2012 - Neural Information Processing Systems, South Lake Tahoe, Ne.

[Sch75] Schal, M., 1975. "Conditions for Optimality in Dynamic Programming and for the Limit of n-Stage Optimal Policies to be Optimal," Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, Vol. 32, pp. 179-196.

[Sch11] Scherrer, B., 2011. "Performance Bounds for Lambda Policy Iteration and Application to the Game of Tetris," Report RR-6348, INRIA, France; J. of Machine Learning Research, Vol. 14, 2013, pp. 1181-1227.

[Sch12] Scherrer, B., 2012. "On the Use of Non-Stationary Policies for Infinite-Horizon Discounted Markov Decision Processes," INRIA Lorraine Report, France.

[Sha53] Shapley, L. S., 1953. "Stochastic Games," Proc. Nat. Acad. Sci. U.S.A., Vol. 39.

[Sob75] Sobel, M. J., 1975. "Ordinal Dynamic Programming," Management Science, Vol. 21, pp. 967-975.

[Str66] Strauch, R., 1966. "Negative Dynamic Programming," Ann. Math. Statist., Vol. 37, pp. 871-890.

[SuB98] Sutton, R. S., and Barto, A. G., 1998. Reinforcement Learning, MIT Press, Cambridge, MA.

[Sze98a] Szepesvari, C., 1998. Static and Dynamic Aspects of Optimal Sequential Decision Making, Ph.D. Thesis, Bolyai Institute of Mathematics, Hungary.

[Sze98b] Szepesvari, C., 1998. "Non-Markovian Policies in Sequential Decision Problems," Acta Cybernetica, Vol. 13, pp. 305-318.

[Sze10] Szepesvari, C., 2010. Algorithms for Reinforcement Learning, Morgan and Claypool Publishers, San Franscisco, CA.

[TBA86] Tsitsiklis, J. N., Bertsekas, D. P., and Athans, M., 1986. "Distributed Asynchronous Deterministic and Stochastic Gradient Optimization Algorithms," IEEE Trans. Aut. Control, Vol. AC-31, pp. 803-812.

[ThS10a] Thiery, C., and Scherrer, B., 2010. "Least-Squares λ-Policy Iteration: Bias-Variance Trade-off in Control Problems," in ICML'10: Proc. of the 27th Annual International Conf. on Machine Learning.

[ThS10b] Thiery, C., and Scherrer, B., 2010. "Performance Bound for Approximate Optimistic Policy Iteration," Technical Report, INRIA, France.

[Tsi94] Tsitsiklis, J. N., 1994. "Asynchronous Stochastic Approximation and Q-Learning," Machine Learning, Vol. 16, pp. 185-202.

[VVL13] Vrabie, V., Vamvoudakis, K. G., and Lewis, F. L., 2013. Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles, The Institution of Engineering and Technology, London.

[VeP87] Verdu, S., and Poor, H. V., 1987. "Abstract Dynamic Programming Models under Commutativity Conditions," SIAM J. on Control and Opt., Vol. 25, pp. 990-1006.

[Wat89] Watkins, C. J. C. H., Learning from Delayed Rewards, Ph.D. Thesis, Cambridge Univ., England.

[Whi80] Whittle, P., 1980. "Stability and Characterization Conditions in Negative Programming," Journal of Applied Probability, Vol. 17, pp. 635-645.

[Whi82] Whittle, P., 1982. Optimization Over Time, Wiley, N. Y., Vol. 1, 1982, Vol. 2, 1983.

[Whi90] Whittle, P., 1990. Risk-Sensitive Optimal Control, Wiley, Chichester.

[WiB93] Williams, R. J., and Baird, L. C., 1993. "Analysis of Some Incremental Variants of Policy Iteration: First Steps Toward Understanding Actor-Critic Learning Systems," Report NU-CCS-93-11, College of Computer Science, Northeastern University, Boston, MA.

[Wil71] Willems, J., 1971. "Least Squares Stationary Optimal Control and the Algebraic Riccati Equation," IEEE Trans. on Automatic Control, Vol. 16, pp. 621-634.

[YuB10] Yu, H., and Bertsekas, D. P., 2010. "Error Bounds for Approximations from Projected Linear Equations," Math. of OR, Vol. 35, pp. 306-329.

[YuB12] Yu, H., and Bertsekas, D. P., 2012. "Weighted Bellman Equations and their Applications in Dynamic Programming," Lab. for Info. and Decision Systems Report LIDS-P-2876, MIT.

[YuB13a] Yu, H., and Bertsekas, D. P., 2013. "Q-Learning and Policy Iteration Algorithms for Stochastic Shortest Path Problems," Annals of Operations Research, Vol. 208, pp. 95-132.

[YuB13b] Yu, H., and Bertsekas, D. P., 2013. "On Boundedness of Q-Learning Iterates for Stochastic Shortest Path Problems," Math. of OR, Vol. 38, pp. 209-227.

[YuB15] Yu, H., and Bertsekas, D. P., 2015. "A Mixed Value and Policy Iteration Method for Stochastic Control with Universally Measurable Policies," Math. of OR, Vol. 40, pp. 926-968.

[Yu11] Yu, H., 2011. "Stochastic Shortest Path Games and Q-Learning," Lab. for Info. and Decision Systems Report LIDS-P-2875, MIT.

[Yu15] Yu, H., 2015. "On Convergence of Value Iteration for a Class of Total Cost Markov Decision Processes," SIAM J. on Control and Optimization, Vol. 53, pp. 1982-2016.

[Zac64] Zachrisson, L. E., 1964. "Markov Games," in Advances in Game Theory, by M. Dresher, L. S. Shapley, and A. W. Tucker, (eds.), Princeton Univ. Press, Princeton, N. J., pp. 211-253.

INDEX