References

Alden, J. M. and Smith, R. L.

[1] Rolling horizon procedures in nonhomogeneous Markov decision processes. *Oper. Res.* **40** (1992), 183–194.

Altman, E.

[1] *Constrained Markov Decision Processes*. Chapman & Hall/CRC, Boca Raton, FL, to appear.

Anderson, E. J. and Nash, P.

[1] *Linear Programming in Infinite-Dimensional Spaces*. Wiley, Chichester, U.K., 1987.

Ash, R. B.

[1] *Real Analysis and Probability*. Academic Press, New York, 1972.

Balder, E.

[1] *Lectures on Young Measures*, Cahiers de mathématiques de la décision. CEREMADE, Université Paris IX-Dauphine, Paris, 1995.

Baykal-Gürsay, M. and Ross, K. W.

[1] Variability sensitive Markov decision processes. *Math. Oper. Res.* **17** (1992), 558–571.

Bensoussan, A.

[1] Stochastic control in discrete time and applications to the theory of production. *Math. Programm. Study* **18** (1982), 43–60.

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

Bertsekas, D. P. and Shreve, S. E.
[1] *Stochastic Optimal Control: The Discrete Time Case*. Academic Press, New York, 1978.

Bes, C. and Lasserre, J. B.
[1] An on-line procedure in discounted infinite-horizon stochastic optimal control. *J. Optim. Theory Appl.* **50** (1986), 61–67.

Bes, C. and Sethi, S. P.
[1] Concepts of forecast and decision horizons: applications to dynamic stochastic optimization problems. *Math. Oper. Res.* **13** (1988), 295–310.

Bhattacharya, R. N. and Majumdar, M.
[1] Controlled semi-Markov models—the discounted case. *J. Statist. Plann. and Inference* **21** (1989), 365–381.

Billingsley, P.
[1] *Convergence of Probability Measures*. Wiley, New York, 1968.

Blackwell, D.
[1] Memoryless strategies in finite-stage dynamic programming. *Ann. Math. Statist.* **35** (1964), 863–865.

Bourbaki, N.
[1] *Intégration*, Chap. IX. Hermann, Paris, 1969.

Brezis, H.
[1] *Analyse Fonctionnelle: Théorie et Applications*, 4^e tirage. Masson, Paris, 1993.

Brown, B. W.
[1] On the iterative method of dynamic programming on a finite space discrete time Markov process. *Ann. Math. Statist.* **33** (1965), 719–726.

Cavazos-Cadena, R.
[1] Finite-state approximations to denumerable discounted Markov decision processes. *Appl. Math. Optim.* **14** (1986), 1–26.

Cavazos-Cadena, R. and Montes-de-Oca, R.
[1] Optimal stationary policies in controlled Markov chains with the expected total-reward criterion. Preprint, Departamento de Matemáticas, UAM-Iztapalapa, México, 1997.

Chen, C.-T.
[1] *Linear System Theory and Design*. Holt, Rinehart and Winston, New York, 1984.

Craven, B. D. and Koliha, J. J.

[1] Generalizations of Farkas' theorem. *SIAM J. Math. Anal. Appl.* **8** (1977), 983–997.

Dekker, R., Hordijk, A. and Spieksma, F.

[1] On the relation between recurrence and ergodicity properties in denumerable Markov chains. *Math. Oper. Res.* **19** (1994), 539–559.

Derman, C.

[1] *Finite State Markovian Decision Processes.* Academic Press, New York, 1970.

Derman, C. and Strauch, R. E.

[1] A note on memoryless rules for controlling sequential control processes. *Ann. Math. Statist.* **37** (1966), 276–278.

Doob, J. L.

[1] *Measure Theory.* Springer-Verlag, New York, 1994.

Duflo, M.

[1] *Méthodes Récursives Aléatoires.* Masson, Paris, 1990.

Dutta, P. K.

[1] What do discounted optima converge to? A theory of discount rate asymptotics in economic models. *J. Economic Theory* **55** (1991), 64–94.

Dynkin, E. B. and Yushkevich, A. A.

[1] *Controlled Markov Processes.* Springer-Verlag, New York, 1979.

Easley, D. and Spulber, D.F.

[1] Stochastic equilibrium and optimality with rolling plans. *International Econ. Rev.* **22** (1981), 79–103.

Fernández-Gaucherand, E., Ghosh, M. K., and Marcus, S. I.

[1] Controlled Markov processes on the infinite planning horizon: weighted and overtaking cost criteria. *ZOR: Math. Methods in Oper. Res.* **39** (1994), 131–155.

Filar, J. A., Kallenberg, L. C. M. and Huey-Miin, L.

[1] Variance-penalized Markov decision processes. *Math. Oper. Res.* **14** (1989), 147–161.

Flynn, J.

[1] On optimality criteria for dynamic programs with long finite horizons. *J. Math. Anal. Appl.* **76** (1980), 202–208.

Gale, D.

[1] On optimal development in a multi-sector economy. *Rev. of Economic Studies* **34** (1965), 1–19.

Glynn, P. W.

[1] *Simulation output analysis for general state space Markov chains*. Ph. D. Dissertation, Dept. of Operations Research, Stanford University, 1989.

[2] Some topics in regenerative steady-state simulation. *Acta Appl. Math.* **34** (1994), 225–236.

Glynn, P. W. and Meyn, S. P.

[1] A Lyapunov bound for solutions of the Poisson equation. *Ann. Prob.* **24** (1996), 916-931.

González-Hernández, J. and Hernández-Lerma, O.

[1] Envelopes of sets of measures, tightness, and Markov control processes. *Appl. Math. Optim.*, to appear.

Gordienko, E. and Hernández-Lerma, O.

[1] Average cost Markov control processes with weighted norms: value iteration. *Appl. Math.* (Warsaw) **23** (1995), 219–237.

[2] Average cost Markov control processes with weighted norms: existence of canonical policies. *Appl. Math.* (Warsaw) **23** (1995), 199–218.

Gordienko, E., Montes-de-Oca, R. and Minjárez-Sosa, A.

[1] Average cost optimization in Markov control processes with unbounded costs: ergodicity and finite horizon approximation. Preprint, Departamento de Matemáticas, UAM-Iztapalapa, México, 1995.

Hall, P. and Heyde, C. C.

[1] *Martingale Limit Theory and Its Applications*. Academic Press, New York, 1980.

Haviv, M. and Puterman, M. L.

[1] Bias optimality in controlled queueing systems. *J. Appl. Prob.* **35** (1998), 136–150.

Hernández-Lerma, O.

[1] *Adaptive Markov Control Processes*. Springer-Verlag, New York, 1989.

Hernández-Lerma, O., Carrasco, G. and Pérez-Hernández, R.

[1] Markov control processes with the expected total-cost criterion: optimality, stability, and transient models. Reporte Interno, Depto. de Matemáticas, CINVESTAV-IPN, 1998. (Submitted.)

Hernández-Lerma, O. and González-Hernández, J.

[1] Infinite linear programming and multichain Markov control processes in uncountable spaces. *SIAM J. Control Optim.* **36** (1998), 313–335.

Hernández-Lerma, O. and Lasserre, J. B.

[1] *Discrete-Time Markov Control Processes: Basic Optimality Criteria.* Springer-Verlag, New York, 1996.

[2] Invariant probabilities for Feller-Markov chains. *J. Appl. Math. and Stoch. Anal.* **8** (1995), 341–345.

[3] Existence of bounded invariant probability densities for Markov chains. *Stat. Prob. Lett.* **28** (1996), 359–366.

[4] Cone-constrained linear equations in Banach spaces. *J. Convex Anal.* **4** (1997), 149–164.

[5] Existence and uniqueness of fixed points for Markov operators and Markov processes. *Proc. London Math. Soc.* **76** (1998), 711–736.

[6] On the probabilistic multichain Poisson equation. LAAS Report N^o 97155, LAAS-CNRS, Toulouse, 1997.

[7] Existence of solutions to the Poisson equation in L_p spaces. *Proc. IEEE Conference on Decision and Control* (CDC), Kobe, Japan, 1996, vol. 4, pp. 4190–4195.

[8] Error bounds for rolling horizon policies in discrete-time Markov control processes. *IEEE Trans. Autom. Control* **35** (1990), 1118–1124.

[9] Value iteration and rolling plans for Markov control processes with unbounded rewards. *J. Math. Anal. Appl.* **177** (1993), 38–55.

[10] A forecast horizon and a stopping rule for general Markov decision processes. *J. Math. Anal. Appl.* **132** (1988), 388–400.

[11] Policy iteration for average cost Markov control processes on Borel spaces. *Acta Appl. Math.* **47** (1997), 125–154.

[12] New criteria for positive Harris recurrence of Markov chains. LAAS Report N^o 97431, LAAS-CNRS, Toulouse, 1997.

[13] Ergodic theorems and ergodic decomposition for Markov chains. *Acta Appl. Math.* **54** (1998), 99–199.

[14] Linear programming and average optimality for Markov control processes in Borel spaces—unbounded costs. *SIAM J. Control Optim.* **32** (1994), 480–500.

[15] Approximation schemes for infinite linear programs. *SIAM J. Optim.* **8** (1998), 973–988.

[16] Linear programming approximations for Markov control processes in metric spaces. *Acta Appl. Math.* **51** (1998), 123–139.

Hernández-Lerma, O., Montes-de-Oca, R. and Cavazos-Cadena, R.

[1] Recurrence conditions for Markov decision processes with Borel state space: a survey. *Ann. Oper. Res.* **28** (1991), 29–46.

Hernández-Lerma, O. and Vega-Amaya, O.

[1] Infinite-horizon Markov control processes with undiscounted cost criteria: from average to overtaking optimality. *Appl. Math.* (Warsaw) **25** (1998), 153–178.

Hernández-Lerma, O., Vega-Amaya, O. and Carrasco, G.

[1] Sample-path optimality and variance-minimization of average cost Markov control processes. *SIAM J. Control Optim.*, to appear.

Hinderer, K.

[1] *Foundations of Non-Stationary Dynamic Programming with Discrete-Time Parameter*. Lecture Notes in Oper. Res. and Math. Syst. **33**, Springer-Verlag, Berlin, 1970.

Hinderer, K. and Hübner, G.

[1] An improvement of J. F. Shapiro's turnpike theorem for the horizon of finite stage discrete dynamic programs. In *Trans. 7th Prague Conf. on Information Theory, Statist. Dec. Func. and Random Proc.*, Vol. A, 1974 (Academia, Prague, 1977), pp. 245–255.

Hopp, W.

[1] Identifying forecast horizons in non-homogeneous Markov decision processes. *Oper. Res.* **37** (1989), 339–344.

Hordijk, A.

[1] *Dynamic Programming and Markov Potential Theory*, 2nd ed. Mathematical Centre Tracts No. **51**, Mathematisch Centrum, Amsterdam, 1977.

Hordijk, A. and Lasserre, J. B.

[1] Linear programming formulation of MDPs in countable state space: the multichain case. *Z. Oper. Res.* **40** (1994), 91–108.

Hordijk, A. and Spieksma, F.

[1] A new formula for the deviation matrix. In *Probability, Statistics and Optimization* (F. P. Kelly, ed.), Wiley, New York, 1994, pp. 497–507.

Johansen, L.

[1] *Lectures on Macroeconomic Planning*. North-Holland, Amsterdam, 1977.

Kallenberg, L. C. M.

[1] *Linear Programming and Finite Markovian Control Problems*. Mathematical Centre Tracts No. **148**, Mathematisch Centrum, Amsterdam, 1983.

Kartashov, N. V.

[1] Criteria for uniform ergodicity and strong stability of Markov chains with a common phase space. *Theory Probab. and Math. Statist.* **30** (1985), 71–89.

[2] Inequalities in theorems of ergodicity and stability for Markov chains with common phase space. I. *Theory Probab. Appl.* **30** (1985), 247–259.

[3] Inequalities in theorems of ergodicity and stability for Markov chain with common phase space. II. *Theory Probab. Appl.* **30** (1985), 507–515.

[4] Strongly stable Markov chains. *J. Soviet Math.* **34** (1986), 1493–1498.

[5] *Strong Stable Markov Chains*. VSP, Utrecht, The Netherlands, 1996.

Kleinman, D.

[1] An easy way to stabilize a linear control system. *IEEE Trans. Autom. Control* **15** (1970), p. 692.

Kurano, M.

[1] Markov decision processes with a minimum-variance criterion. *J. Math. Anal. Appl.* **123** (1987), 572–583.

Kurano, M. and Kawai, M.

[1] Existence of optimal stationary policies in discounted decision processes: approaches by occupation measures. *Computers Math. Appl.* **27** (1994), 95–101.

Kwon, W. H., Bruckstein, A. M. and Kailath, T.

[1] Stabilizing state feedback design via the moving horizon method. *Internat. J. Control* **37** (1983), 631–643.

Lasota, A. and Mackey, M. C.

[1] *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*, 2nd ed. Springer-Verlag, New York, 1994.

Lasserre, J. B.

[1] Existence and uniqueness of an invariant probability measure for a class of Feller-Markov chains. *J. Theoret. Prob.* **9** (1996), 595–612.

[2] Invariant probabilities for Markov chains on a metric space. *Stat. Prob. Lett.*, to appear.

[3] Sample-path average optimality for Markov control processes. *IEEE Trans. Autom. Control*, to appear.

Lippman, S. A.

[1] On dynamic programming with unbounded rewards. *Manage. Sci.* **21** (1975), 1225–1233.

Luenberger, D. G.

[1] *Optimization by Vector Space Methods*. Wiley, New York, 1969.

Makowski, A. M. and Shwartz, A.
[1] On the Poisson equation for countable Markov chains. Tech. Rpt., Dept. of Electrical Engineering, University of Maryland, 1994.

Mandl, P.
[1] On the variance in controlled Markov chains. *Kybernetika* **7** (1971), 1–12.
[2] A connection between controlled Markov chains and martingales. *Kybernetika* **9** (1973), 237–241.

Mandl, P. and Lausmanová, M.
[1] Two extensions of asymptotic methods in controlled Markov chains. *Ann. Oper. Res.* **28** (1991), 67–80.

McKenzie, L. W.
[1] Turnpike theory. *Econometrica* **44** (1976), 841–865.

Metivier, M. and Priouret, P.
[1] Théorèmes de convergence presque sure pour une classe d'algorithmes stochastique à pas decroissant. *Prob. Th. Rel. Fields* **74** (1987), 403–428.

Meyn, S. P.
[1] The policy iteration algorithm for average reward Markov decision processes with general state space. *IEEE Trans. Autom. Control* **42** (1997), 1663–1679.

Meyn, S. P. and Tweedie, R. L.
[1] *Markov Chains and Stochastic Stability*. Springer-Verlag, London, 1993.
[2] Computable bounds for geometric convergence rates of Markov chains. *Ann. Appl. Prob.* **4** (1994), 981–1011.

Mokkadem, A.
[1] Sur un modèle autorégressif nonlinéaire. Ergodicité et ergodicité géométrique. *J. Time Series Anal.* **8** (1987), 195–205.

Neveu, J.
[1] *Mathematical Foundations of the Calculus of Probability*. Holden-Day, San Francisco, 1965.

Nowak, A. S.
[1] Stationary overtaking optimal strategies in Markov decision processes with general state space. Preprint, Institute of Mathematics, Technical University of Wroclaw, Poland, 1992.

Nowak, A. S. and Vega-Amaya, O.
[1] A counterexample on overtaking optimality. Preprint, Institute of Mathematics, Technical University of Wroclaw, Poland; Departamento de Matemáticas, Universidad de Sonora, México, 1997.

Nummelin, E.

[1] *General Irreducible Markov Chains and Non-Negative Operators*. Cambridge University Press, Cambridge, 1984.

[2] On the Poisson equation in the potential theory of a single kernel. *Math. Scand.* **68** (1991), 59–82.

Orey, S.

[1] *Limit Theorems for Markov Chain Transition Probabilities*. Van Nostrand Reinhold, London, 1971.

Parthasarathy, K. R.

[1] *Probability Measures on Metric Spaces*. Academic Press, New York, 1967.

Piunovski, A. B.

[1] General Markov models with the infinite horizon. *Problems of Control and Infor. Theory* **18** (1989), 169–182.

Pliska, S. R.

[1] On the transient case for Markov decision chains with general state spaces. In: Puterman [2], pp. 335–349.

Puterman, M. L.

[1] *Markov Decision Processes*. Wiley, New York, 1994.

[2] (Editor) *Dynamic Programming and Its Applications*. Academic Press, New York, 1979.

Quelle, G.

[1] Dynamic programming of expectation and variance. *J. Math. Anal. Appl.* **55** (1976), 239–252.

Ramsey, F. P.

[1] A mathematical theory of savings. *Economic J.* **38** (1928), 543–559.

Rempala, R.

[1] Forecast horizon in a dynamic family of one-dimensional control problems. *Diss. Math.* **315** (1991).

Revuz, D.

[1] *Markov Chains*, revised ed. North-Holland, Amsterdam, 1984.

Rieder, U.

[1] Measurable selection theorems for optimization problems. *Manuscripta Math.* **24** (1978), 115–131.

[2] On optimal policies and martingales in dynamic programming. *J. Appl. Prob.* **13** (1976), 507–518.

[3] On Howard's policy improvement method. *Math. Operationsforsch. Statist., Ser. Optimization* **8** (1977), 227–236.

Robertson, A. P. and Robertson, W.

[1] *Topological Vector Spaces*. Cambridge University Press, Cambridge, UK, 1964.

Ross, S. M.

[1] *Applied Probability Models with Optimization Applications*. Holden-Day, San Francisco, 1970.

Royden, H. L.

[1] *Real Analysis*, 2nd ed. Macmillan, New York, 1968.

Rudin, W.

[1] *Real and Complex Analysis*, 3rd ed. McGraw-Hill, New York, 1986.

Schäl, M.

[1] Conditions for optimality and for the limit of n-stage optimal policies to be optimal. *Z. Wahrs. Verw. Geb.* **32** (1975), 179–196.

Schweitzer, P. J.

[1] On undiscounted Markovian decision processes with compact action spaces. *RAIRO Rech. Opér./Oper. Res.* **19** (1985), 71–86.

Shapiro, J. M.

[1] Turnpike planning horizons for a Markovian decision model. *Manage. Sci.* **14** (1968), 292–300.

Spieksma, F. and Tweedie, R. L.

[1] Strengthening ergodicity to geometric ergodicity for Markov chains. *Stoch. Models* **10** (1994), 45–75.

Strauch, R.

[1] Negative dynamic programming. *Ann. Math. Statist.* **37** (1966), 871–890.

Syski, R.

[1] Ergodic potential. *Stoch. Proc. Appl.* **7** (1978), 311–336.

Tijms, H. C.

[1] *Stochastic Models: An Algorithmic Approach.* Wiley, New York, 1994.

Tijms, H. C. and Wessels, J.

[1] (Editors) *Markov Decision Theory*. Mathematical Centre Tracts No. **93**, Mathematisch Centrum, Amsterdam, 1977.

Tweedie, R. L.

[1] Sufficient conditions for ergodicity and geometric ergodicity of Markov chains on a general state space. *Stoch. Proc. Appl.* **3** (1975), 385–403.

[2] The existence of moments for stationary Markov chains. *J. Appl. Probab.* **20** (1983), 191–196.

van Hee, K. M. Hordijk, A., and van der Wal, J.
[1] Successive approximations for convergent dynamic programming. In: Tijms and Wessels [1], pp. 183–211.

van Nunen, J. A. E. E. and Wessels, J.
[1] A note on dynamic programming with unbounded rewards. *Manage. Sci.* **24** (1978), 576–580.
[2] Markov decision processes with unbounded rewards. In: Tijms and Wessels [1], pp. 1–24.

Vega-Amaya, O.
[1] Overtaking optimality for a class of production-inventory systems. Preprint, Departamento de Matemáticas, Universidad de Sonora, México, 1996.
[2] *Markov Control Processes In Borel Spaces: Undiscounted Criteria.* Doctoral Thesis, UAM-Iztapalapa, México, 1998. (In Spanish.)
[3] Sample path average optimality of Markov control processes with strictly unbounded cost. *Appl. Math.* (Warsaw), to appear.

Veinott, A. F. (Jr.)
[1] Discrete dynamic programming with sensitive discount optimality criteria. *Ann. Math. Statist.* **40** (1969), 1635–1660.
[2] On finding optimal policies in discrete dynamic programming with no discounting. *Ann. Math. Statist.* **37** (1965), 1284–1294.

Vershik, A. M.
[1] Some remarks on the infinite-dimensional problems of linear programming. *Russian Math. Surveys* **29** (1970), 117–124.

Vershik, A. M. and Temel't, V.
[1] Some questions concerning the approximation of the optimal value of infnite-dimensional problems in linear programming. *Siberian Math. J.* **9** (1968), 591–601.

von Weizsäcker, C. C.
[1] Existence of optimal programs of accumulation for an infinite horizon. *Rev. of Economic Studies* **32** (1965), 85–104.

Wakuta, K.
[1] Arbitrary state semi-Markov decision processes with unbounded rewards. Optimization **18** (1987), 447–454.

Wessels, J.
[1] Markov programming by successive approximations with respect to weighted supremum norms. *J. Math. Anal. Appl.* **58** (1977), 326–335.

Yosida, K.
[1] *Functional Analysis*, 6th ed. Springer-Verlag, Berlin, 1980.

Yushkevich, A. A.
[1] Blackwell optimality in Borelian continuous-in-action Markov decision processes. *SIAM J. Control Optim.* **35** (1997), 2157–2182.

Abbreviations

a.a.	almost all
a.e.	almost everywhere
a.s.	almost surely
i.i.d.	independent and identically distributed
l.s.c.	lower semicontinuous
u.s.c.	upper semicontinuous
p.m.	probability measure
i.p.m.	invariant probability measure
AC	average cost
ACOE	Average Cost Optimality Equation
ACOI	Average Cost Optimality Inequality
ADO	asymptotic discount optimality
DC	discounted cost
DCOE	discounted cost optimality equation
DP	dynamic programming
ETC	expected total cost
IFS	iterated function system
LCSM	locally compact separable metric
LP	linear programming
LLN	Law of Large Numbers
MCM	Markov control model
MCP	Markov control process
O.O.	overtaking optimality
P.E.	Poisson equation

PI	policy iteration
PIA	policy iteration algorithm
RH	rolling horizon
VI	value iteration

Glossary of notation

$\square$	end of proof or example or remark
$:=$	equality by definition
I_B	indicator function of a set B, defined as

$$I_B(x) := \begin{cases} 1 & \text{if } x \in B, \\ 0 & \text{otherwise.} \end{cases}$$

$r^+ := \max(r, 0)$, $r^- := -\min(r, 0)$

Chapter 7

Section 7.1

X	Borel (state) space
$\mathcal{B}(X)$	Borel σ-algebra

Section 7.2

w	weight function
$\|u\|$	sup norm of a function u
$\|u\|_w$	w-norm of a function u
$\mathbb{B}(X)$	Banach space of bounded measurable functions on X
$\mathbb{B}_w(X)$	Banach space of measurable functions on X with finite w-norm
$\|\mu\|_{TV}$	total variation norm of a measure μ
$\|\mu\|_w$	w-norm of a measure μ
$\mathbb{M}(X)$	Banach space of signed measures μ on X with $\|\mu\|_{TV} < \infty$
$\mathbb{M}_w(X)$	Banach space of signed measures μ on X with $\|\mu\|_w < \infty$
$Q(B\|x)$	signed kernel on X

$Qu(x) := \int_X u(y)Q(dy|x)$

$\mu Q(B) := \int_X Q(B|x)\mu(dx)$

$\|Q\|_w$	w-norm of a signed kernel Q

δ_x	Dirac measure at $x \in X$
QR	composition of the signed kernels Q and R

$Q^n := QQ^{n-1}$

$Q^0(B|x) := \delta_x(B)$

Section 7.3

$P(B|x)$	transition probability function of a Markov chain
$P^t(B|x)$	t-step transition probability
τ_B	hitting time of the set B
η_B	occupation time of the set B

$L(x,B) := P_x(\tau_B < \infty)$

$U(x,B) := E_x(\eta_B)$

$\nu \ll \mu$	the measure ν is absolutely continuous with respect to the measure μ
λ	maximal irreducibility measure
$\mathcal{B}(X)^+$	family of sets $B \in \mathcal{B}(X)$ with $\lambda(B) > 0$

$\int v d\mu \equiv \mu(v) \equiv \langle \mu, v \rangle$

$C_b(X)$	Banach space of continuous bounded functions on X

Chapter 8

Section 8.2

$\mathcal{M} := (X, A, \{A(x)|x \in X\}, Q, c)$ Markov control model

$\mathbb{K}$	set of feasible state-action pairs
$\mathbb{F}$	set of decision functions (or selectors)
Φ	set of randomized decision functions
H_t	family of admissible histories up to time t
π	control policy
φ^∞	randomized stationary policy, $\varphi \in \Phi$
f^∞	deterministic stationary policy, $f \in F$
Π	set of all control policies
Π_{RM}	set of randomized Markov policies
Π_{RS}	set of randomized stationary policies φ^∞
Π_D	set of deterministic policies
Π_{DM}	set of deterministic Markov policies
Π_{DS}	set of deterministic stationary policies f^∞

$c(x,\varphi) \equiv c_\varphi(x) := \int_A c(x,a)\varphi(da|x)$ for $\varphi \in \Phi$

$c(x,f) \equiv c_f(x) := c(x, f(x))$ for $f \in F$

$Q(\cdot|x,\varphi) \equiv Q_\varphi(\cdot|x) := \int_A Q(\cdot|x,a)\varphi(da|x)$ for $\varphi \in \Phi$

$Q(\cdot|x,f) \equiv Q_f(\cdot|x) := Q(\cdot|x, f(x))$ for $f \in F$

$(\Omega, \mathcal{F})$	canonical measurable space
P_ν^π	p.m. on $(\Omega, \mathcal{F})$ determined by the policy π and the initial distribution ν

$P_x^\pi := P_\nu^\pi$ if $\nu = \delta_x$

E_ν^π	expectation with respect to P_ν^π

$E_x^\pi := E_\nu^\pi$ if $\nu = \delta_x$

Section 8.3

$V(\pi,x)$	α-discounted cost ($0 < \alpha < 1$) when using the policy π, given the initial state x
$V^*(x)$	α-discount value function
$V_n(\pi,x)$	α-discounted n-stage expected cost
$v_n(x)$	α-value iteration (α-VI) function, $n = 1, 2, \ldots$
$\|Q\|_w$	w-norm of the transition law Q

T_α	DP operator

Section 8.4

$\mathbb{F}_n$	set of α-VI decision function, $n = 1, 2, \ldots$
$\mathbb{F}_*$	set of α-discount optimal decision functions
$D(x, a)$	α-discount discrepancy function
$A_*(x)$	α-discount optimal control actions in the state x
$A_n(x)$	α-VI optimal control actions in the state x; $n = 1, 2, \ldots$
$D_n(x, a)$	α-VI discrepancy function, $n = 1, 2, \ldots$

Section 8.5

$\mathbb{L}(X)$	family of l.s.c. functions on X

$\mathbb{L}_w(X) := \mathbb{L}(X) \cap \mathbb{B}_w(X)$

$C(X)$	family of continuous functions on X

$C_w(X) := C(X) \cap \mathbb{B}_w(X)$

$C_b(X)$	family of continuous bounded functions on X

Chapter 9

Section 9.1

$V_1(\pi, x)$	expected total cost (ETC) when using the policy π, given the initial state x
$V_1^*(x)$	ETC value function

Section 9.2

$\overline{R}$	extended real numbers

$r^+ := \max(r, 0)$

$r^- := \max(-r, 0) = -\min(r, 0)$

Section 9.3

$J_n(\pi, x)$	n-stage ETC when using the policy π, given the initial state x
$J_n^*(x)$	n-stage optimal ETC

$V_\alpha(\pi, x) \equiv V(\pi, x)$ α-discounted cost, $0 < \alpha < 1$

$V_\alpha^*(x) \equiv V^*(x)$ α-discount value function

$$V_1^{(+)}(\pi, x) := E_x^\pi\left(\sum_{t=0}^\infty c_t^+\right)$$

$$V_1^{(-)}(\pi, x) := E_x^\pi\left(\sum_{t=0}^\infty c_t^-\right)$$

$V_1^n(\pi, x)$	ETC from time n onwards when using the policy π, given the initial state $x_0 = x$

Section 9.4.

$V_1(\pi, \nu)$	ETC when using the policy π, given the initial distribution ν
μ_ν^π	ETC-expected occupation measure on $X \times A$ when using the policy π, given the initial distribution ν
$\widehat{\mu}_\nu^\pi$	marginal of μ_ν^π on X
$\mu_{\nu,t}^\pi$	distribution of (x_t, a_t) when using the policy π, given the initial distribution ν, $t = 0, 1, \ldots$
$\widehat{\mu}_{\nu,t}^\pi$	marginal of $\mu_{\nu,t}^\pi$ on X
$\pi^{(1)}$	1-shift policy determined by π

Section 9.5

$T := T_1$	dynamic programming operator (when $\alpha = 1$)
$\mathcal{U}$	see Definition 9.5.1
$D_1(x, a)$	ETC-discrepancy function
$\pi^{(n)}$	n-shift policy determined by π $(n = 0, 1, \ldots)$
M_n^*	see (9.5.12)

Section 9.6

$Q_\varphi(\cdot|x) \equiv Q(\cdot|x,\varphi)$
$:= \int_A Q(\cdot|x,a)\varphi(da|x)$

$Q_t(\cdot|x) := Q_{\varphi_t}(\cdot|x)$

$Q_\pi^t := Q_0 Q_1 \cdots Q_{t-1}$ for $t = 1, 2, \ldots$

$\|Q_\varphi\|_w$	w-norm of Q_φ

Chapter 10

Section 10.1

$OC(\pi,x)$	opportunity cost of policy π, given the initial state x
$D(\pi,x)$	Dutta's criterion
$J(\pi,x)$	expected average cost (AC) when using the policy π, given the initial state x
$J^*(x)$	optimal expected AC

Section 10.2

$Q_f^t(\cdot\|x) \equiv Q^t(\cdot\|x,f)$	t-step transition probability, $t = 0, 1, \ldots$
$\|Q_f\|_w$	w-norm of Q_f
μ_f	i.p.m. of Q_f

$L_1(\mu) := L_1(X, \mathcal{B}(X), \mu)$

$\mu_f(u) := \int_X u\, d\mu_f$

$c_f(x) \equiv c(x,f) := c(x, f(x))$

$J(f) := \mu_f(c_f)$

h_f	bias function of $f \in \mathbb{F}$

Section 10.3

$\mathbb{F}_{AC}$	family of AC-optimal decision functions
$\mathbb{F}_{ca}$	family of canonical decision functions
$\mathbb{F}_{bias}$	family of bias-optimal decision functions
(ρ^*, h^*)	solution of the ACOE
$J_n(\pi,x,h)$	n-stage ETC with terminal cost function h
$J_n^*(x,h)$	value function for $J_n(\pi,x,h)$
λ	irreducibility measure
$\widehat{h}(x)$	optimal bias function
$A^*(x)$	AC-canonical control actions at state x

Section 10.4

$u_\alpha(\cdot) := V_\alpha^*(\cdot) - V_\alpha^*(z)$

$\rho(\alpha) := (1-\alpha)V_\alpha^*(z)$

Section 10.7

$\mathcal{M}_{bias}$	see (10.7.3)

Chapter 11

Section 11.1

$J_n^0(\pi,\nu)$	n-stage sample-path cost when using the policy π, given the initial distribution ν
$J^0(\pi,\nu)$	long-run sample-path AC
$J^I(\pi,\nu)$	limit-infimum expected AC
Var (π,ν)	limiting average variance when using the policy π, given the initial distribution ν
$\rho_{\min}$	see (11.1.12)
$\widehat{\rho}$	see (11.1.15)
ρ^*	see (11.1.23)
$\mathcal{P}(X)$	set of probability measures on X

$\mathcal{P}_\delta(X) := \{\delta_x | x \in X\}$

Section 11.2

$P^{(n)}(\cdot\|x)$	expected average occupation measure
$\sigma^2(c,x)$	limiting average variance for a Markov chain
σ_c^2	see (11.2.11)
$\psi(x)$	see (11.2.12)
Y_t	see (11.2.17)

$M_n := \sum_{t=1}^{n} Y_t$

$\mathcal{F}_t := \sigma\{x_0, \ldots, x_t\}$

Section 11.3

$\mathcal{P}_w(X)$	see (11.3.2)
(ρ^*, h_*)	solution of the ACOE
$\underline{J}^0(\pi, x)$	lim-inf sample path AC
$\psi(x, a)$	see (11.3.5)
$\psi_f(x) := \psi(x, f(x))$	
(σ_*^2, V_*)	see (11.3.17)
$v(x) := w(x)^{1/2}$	
$\mathcal{F}_t(\pi, x)$	see (11.3.23)
$Y_t(\pi, x)$	see (11.3.24)
$M_n(\pi, x)$	see (11.3.25)
$\widehat{D}(x, a)$	AC-discrepancy function
$\Psi_f(X)$	see (11.3.35)
$\sigma^2(f) := \text{Var}\,(f^\infty, \cdot)$	
$\mathcal{M}$var	MCM for the variance minimization problem

Chapter 12

Section 12.2

$(\mathcal{X}, \mathcal{Y})$	dual pair of vector spaces $\mathcal{X}$, $\mathcal{Y}$
$\sigma(\mathcal{X}, \mathcal{Y})$	weak topology on $\mathcal{X}$
$\sigma(\mathcal{Y}^*, \mathcal{Y})$	weak* topology on $\mathcal{Y}^*$
$C_0(S)$	Banach space of continuous functions vanishing at infinity
$\langle \mu, u \rangle := \int u d\mu$	
G^*	adjoint of linear map G
w_0	weight function on X
w	weight function on $X \times Y$
$\mathbb{M}_w(X)_+$	positive cone in $\mathbb{M}_w(X)$
$\mathbb{B}_w(X)_+$	positive cone in $\mathbb{B}_w(X)$
K^*	dual cone of the positive cone K
$\widehat{\mu}$	marginal on X of the measure μ on $X \times Y$
$\mathbb{P}$	primal linear program
$\inf \mathbb{P}$	value of $\mathbb{P}$
$\min \mathbb{P}$	optimal value of $\mathbb{P}$
$\mathbb{P}^*$	dual of $\mathbb{P}$
$\sup \mathbb{P}^*$	value of $\mathbb{P}^*$
$\max \mathbb{P}^*$	optimal value of $\mathbb{P}^*$

Section 12.3

$w(x, a)$	weight function on $\mathbb{K}$; see (12.3.5)
$w_0(x)$	weight function on X; see (12.3.5)
$L_0\mu$	see (12.3.12)
$L_1\mu$	see (12.3.13)
$L\mu$	see (12.3.14)
$L^*(\rho, \mu)$	see (12.3.15)
(P)	AC-related primal linear program
(P*)	dual of (P)
$\mathcal{L}$	perturbation of L

Section 12.5

$\mathcal{C}(X)$	a countable dense subset of $C_0(X)$
$\{\mathcal{C}_k\}$	increasing sequence of finite sets $\mathcal{C}_k \uparrow \mathcal{C}(X)$
$\mathbb{P}(\mathcal{C}_k)$	aggregation of constraints of (P)
$\{\varepsilon_k\}$	sequence of numbers $\varepsilon_k \downarrow 0$
$\mathbb{P}(\mathcal{C}_k, \varepsilon_k)$	aggregation-relaxation of (P)
D	a countable dense subset of $\mathbb{K}$
$\{D_n\}$	increasing sequence of finite sets $D_n \uparrow D$
$\Delta_n := \mathcal{P}(D_n)$	
$\Delta := \cup_{n=1}^{\infty} \Delta_n$	
$\mathbb{P}(\mathcal{C}_k, \varepsilon_k, \Delta_n)$	aggregation-relaxation-inner approximation of (P)

Section 12.6

$E_j := \{(x, a) \in \mathbb{K} | c(x, a) \leq j\}$

Index

The numbers in this index refer to sections

Applications of Mathematics

(continued from page ii)

33 Embrechts/Klüppelberg/Mikosch, **Modelling Extremal Events** (1997)
34 Duflo, **Random Iterative Models** (1997)
35 Kushner/Yin, **Stochastic Approximation Algorithms and Applications** (1997)
36 Musiela/Rutkowski, **Martingale Methods in Financial Modeling: Theory and Application** (1997)
37 Yin/Zhang, **Continuous-Time Markov Chains and Applications** (1998)
38 Dembo/Zeitouni, **Large Deviations Techniques and Applications**, Second Ed. (1998)
39 Karatzas/Shreve, **Methods of Mathematical Finance** (1998)
40 Fayolle/Iasnogorodski/Malyshev, **Random Walks in the Quarter Plane** (1999)
41 Aven/Jensen, **Stochastic Models in Reliability** (1999)
42 Hernández-Lerma/Lasserre, **Further Topics on Discrete-Time Markov Control Processes** (1999)