

# Contents

<b>Acknowledgments</b>	<b>iii</b>
<b>Preface</b>	<b>v</b>
<b>Reader’s Guide and Reference Apparatus</b>	<b>vii</b>
Notation glossary . . . . .	vii
Chapter dependency graph . . . . .	viii
Theorem index: major rigorous statements . . . . .	viii
Algorithm index: major computational templates . . . . .	ix
Concept map . . . . .	ix
<b>1 Introduction</b>	<b>1</b>
1.1 Learning, Dynamics, and Feedback . . . . .	1
1.2 Classical Control Theory and Its Mathematical Foundations . . . . .	2
1.3 Machine Learning, Approximation, and Optimization . . . . .	3
1.4 Reinforcement Learning, Dynamic Programming, and Adaptive Control . . . . .	4
1.5 Data-Driven and Learning-Based Control . . . . .	5
1.6 Deep Learning as a Dynamical-System and Control Problem . . . . .	5
1.7 Safety, Robustness, and Generalization in Feedback . . . . .	7
1.8 Organization of the Book . . . . .	7
1.9 World Models, Autonomous Intelligence, and the Control Viewpoint . . . . .	9
1.10 Control–ML Dictionary and Diagram . . . . .	13
<b>2 Supervised Learning</b>	<b>15</b>
2.1 Literature Review and Historical Context . . . . .	15
2.2 Mathematical Formulation of Supervised Learning . . . . .	17
2.2.1 General Concepts . . . . .	17
2.2.2 Convolutional Neural Networks . . . . .	23
2.2.3 Optimization Theory . . . . .	25
2.3 Gradient Methods . . . . .	27
2.3.1 The Gradient Descent Algorithm . . . . .	27
2.3.2 A Decreasing Sequence of Energies . . . . .	29
2.3.3 Newton’s Method . . . . .	31
2.4 Control–ML Dictionary and Diagram . . . . .	35
2.5 Exercises . . . . .	35

<b>3</b>	<b>Nonparametric Methods and Reproducing Kernels</b>	<b>37</b>
3.1	Introduction: Nonparametric Learning, Regularization, and Kernels	37
3.2	Supervised Learning with Nonparametric Methods	38
3.2.1	Conceptual framework	38
3.3	Reproducing Kernel Hilbert Spaces	39
3.3.1	Machine-learning meaning	39
3.3.2	Definition	39
3.3.3	Immediate properties	40
3.3.4	Uniqueness of the reproducing kernel	40
3.3.5	Density of kernel sections	40
3.4	Constructing an RKHS from a Kernel	42
3.4.1	Machine-learning meaning	42
3.4.2	Positive definite kernels	42
3.4.3	The associated RKHS	42
3.4.4	Comments	45
3.5	The Representer Theorem	45
3.5.1	Machine-learning meaning	45
3.5.2	Solution of the regularized least-squares problem	45
3.6	Examples of RKHSs	46
3.6.1	Machine-learning meaning	46
3.6.2	Finite-dimensional input spaces	46
3.6.3	Application of the representer theorem	47
3.6.4	Sobolev spaces and smoothness penalties	47
3.6.5	Changing the scalar product changes the kernel	48
3.6.6	The standard scalar product on $H_1(a,b)$	49
3.7	Mercer Expansion	50
3.7.1	Machine-learning meaning	50
3.7.2	Compact integral operators	50
3.7.3	Mercer expansion of the kernel	50
3.7.4	The RKHS associated with the Mercer expansion	50
3.8	Feature-Map Kernels and Stochastic Processes	51
3.8.1	Machine-learning meaning	51
3.8.2	Definition of a feature-map kernel	51
3.8.3	Characterization of $HK(E)$	51
3.8.4	The quotient interpretation	52
3.8.5	The separable case	53
3.8.6	Application to stochastic processes	53
3.9	Kernel Interpretation of Support Vector Machines	54
3.9.1	Machine-learning meaning	54
3.9.2	Feature-space formulation	54
3.9.3	Kernel form	55
3.10	Control-ML Dictionary and Diagram	55
3.11	Exercises	56

<b>4</b>	<b>Probability and Stochastic Processes</b>	<b>59</b>
4.1	A Review of Probability Theory . . . . .	60
4.1.1	Basic concepts . . . . .	60
4.1.2	Lebesgue integration on $\mathbb{R}$ . . . . .	61
4.1.3	Random variables . . . . .	63
4.1.4	Moments and functional spaces . . . . .	64
4.1.5	Independence and conditioning . . . . .	66
4.1.6	Limit theorems . . . . .	67
4.2	Stochastic Processes . . . . .	68
4.2.1	Filtrations and stopping times . . . . .	68
4.2.2	Definition and path properties . . . . .	68
4.2.3	Martingales . . . . .	69
4.3	Markov Chains . . . . .	70
4.3.1	Transition kernels . . . . .	70
4.3.2	Markov chains and dynamic programming . . . . .	70
4.3.3	Chapman–Kolmogorov equation . . . . .	71
4.4	Diffusion Processes . . . . .	71
4.4.1	Wiener process . . . . .	71
4.4.2	The Itô integral . . . . .	72
4.4.3	Itô calculus . . . . .	73
4.4.4	Stochastic differential equations . . . . .	74
4.4.5	Markov properties and Kolmogorov equations . . . . .	75
4.4.6	Probabilistic representations of PDEs . . . . .	76
4.4.7	Weak solutions and Girsanov’s theorem . . . . .	76
4.4.8	Backward stochastic differential equations . . . . .	78
4.5	Control–ML Dictionary and Diagram . . . . .	79
4.6	Exercises . . . . .	80
<b>5</b>	<b>Stochastic Approximation and Learning Algorithms</b>	<b>83</b>
5.1	Stochastic Gradient Methods . . . . .	84
5.1.1	Motivation . . . . .	84
5.1.2	Deterministic convergence lemmas . . . . .	85
5.1.3	Robbins–Monro theorem . . . . .	87
5.1.4	Kiefer–Wolfowitz finite-difference method . . . . .	91
5.2	Central Limit Theorem for Stochastic Gradient . . . . .	96
5.2.1	Mean-square rate . . . . .	97
5.2.2	Linearization . . . . .	97
5.2.3	Asymptotic normality . . . . .	98
5.3	Application to Supervised Learning . . . . .	98
5.3.1	Empirical risk . . . . .	98
5.3.2	Gradient formula . . . . .	99
5.3.3	One-neuron illustration . . . . .	99
5.4	Support Vector Machines . . . . .	100
5.4.1	Linear classification . . . . .	100
5.4.2	Dual formulation . . . . .	101
5.4.3	Kernel form . . . . .	101
5.5	Control–ML Dictionary and Diagram . . . . .	102
5.6	Exercises . . . . .	102

<b>6</b>	<b>Overview of Control Theory</b>	<b>105</b>
6.1	Deterministic Control Theory . . . . .	106
6.1.1	Description of the Problem . . . . .	106
6.1.2	Computation of the Gâteaux Differential . . . . .	108
6.1.3	Necessary Condition of Optimality . . . . .	109
6.1.4	Preliminaries to Dynamic Programming . . . . .	111
6.1.5	Hamilton–Jacobi–Bellman Equation . . . . .	112
6.1.6	Link Between Dynamic Programming and the Pontryagin Maximum Principle	114
6.1.7	Linear–Quadratic Case . . . . .	116
6.2	Stochastic Control Theory . . . . .	118
6.2.1	Description of the Problem . . . . .	118
6.2.2	Computation of the Gâteaux Derivative . . . . .	119
6.2.3	Stochastic Dynamic Programming . . . . .	121
6.2.4	Link Between Dynamic Programming and the Stochastic Maximum Principle	127
6.2.5	Stochastic Linear–Quadratic Case . . . . .	129
6.3	Control–ML Dictionary and Diagram . . . . .	131
6.4	Exercises . . . . .	132
<b>7</b>	<b>Generative Modeling and Controllability</b>	<b>135</b>
7.1	Introduction and Literature . . . . .	135
7.2	Diffusion-Based Generative Modeling . . . . .	136
7.2.1	The Machine Learning Problem . . . . .	136
7.2.2	Forward Diffusions and Time Reversal . . . . .	137
7.2.3	The Reverse-Time Diffusion . . . . .	138
7.2.4	A Gaussian Diffusion Example . . . . .	140
7.2.5	The Score Function . . . . .	140
7.2.6	Stochastic Control Interpretation . . . . .	141
7.2.7	Score Estimation, Reverse Sampling, and Control . . . . .	142
7.3	Controllability Theory . . . . .	143
7.3.1	Conceptual Role in Machine Learning . . . . .	143
7.3.2	The Deterministic Linear Case . . . . .	143
7.3.3	Controllability of a Backward Stochastic Differential Equation . . . . .	144
7.3.4	Controllability of a Forward Stochastic Differential Equation . . . . .	146
7.4	Concluding Remarks . . . . .	148
7.5	Control–ML Dictionary and Diagram . . . . .	148
7.6	Exercises . . . . .	149
<b>8</b>	<b>Dynamic Programming Foundations of Reinforcement Learning</b>	<b>151</b>
8.1	Introduction: reinforcement learning and its literature . . . . .	151
8.2	Reinforcement learning and Markov decision processes . . . . .	152
8.2.1	The learning problem before the mathematics . . . . .	152
8.2.2	Controlled transition kernels . . . . .	153
8.2.3	Costs, value functions, and Bellman equations . . . . .	153
8.3	The bounded discounted case . . . . .	154
8.3.1	Conceptual overview . . . . .	154
8.3.2	Four layers of assumptions in the bounded discounted case . . . . .	154
8.3.3	The Bellman operator . . . . .	156
8.3.4	Optimal feedbacks . . . . .	160

8.4	Bellman equations in the unbounded case . . . . .	163
8.4.1	Why unbounded costs require new ideas . . . . .	163
8.4.2	Assumptions and ceiling functions . . . . .	163
8.4.3	The ordered set of Bellman solutions . . . . .	163
8.4.4	Interpretation of the minimum solution . . . . .	167
8.5	Policy iteration and gradient-based improvement . . . . .	169
8.5.1	Policy iteration as evaluation plus improvement . . . . .	169
8.5.2	The exact policy-iteration algorithm . . . . .	169
8.5.3	Gradient-based action improvement . . . . .	171
8.5.4	Combining approximate evaluation and improvement . . . . .	171
8.5.5	Simulation-based approximation . . . . .	172
8.6	Learning Bellman Equations from Trajectories . . . . .	172
8.6.1	Temporal-difference learning . . . . .	172
8.6.2	Q-learning . . . . .	173
8.6.3	Actor–critic methods . . . . .	174
8.6.4	Exploration versus exploitation . . . . .	174
8.6.5	Function approximation error . . . . .	174
8.6.6	Stability and safety in reinforcement learning . . . . .	175
8.6.7	Notes on rigorous convergence theory . . . . .	175
8.7	Subsolutions and linear programming . . . . .	176
8.7.1	Conceptual role of subsolutions . . . . .	176
8.7.2	The largest subsolution . . . . .	176
8.7.3	The linear-programming formulation . . . . .	177
8.8	Explicit solutions: the linear-quadratic-Gaussian model . . . . .	177
8.8.1	Why the linear-quadratic-Gaussian case is special . . . . .	177
8.8.2	The model and transition kernel . . . . .	177
8.8.3	Bellman equation and quadratic ansatz . . . . .	177
8.8.4	Riccati equation as matrix dynamic programming . . . . .	178
8.8.5	Minimum and maximum Riccati solutions . . . . .	179
8.8.6	A uniqueness condition . . . . .	182
8.9	Control–ML Dictionary and Diagram . . . . .	184
8.10	Exercises . . . . .	185
<b>9</b>	<b>Bayesian Learning in Inventory Control</b> . . . . .	<b>187</b>
9.1	Introduction: Bayesian Learning and Sequential Decisions . . . . .	187
9.2	A General Belief-State MDP Formulation . . . . .	189
9.3	A Benchmark Model in Inventory Control . . . . .	189
9.3.1	Modeling idea: an MDP with known demand . . . . .	189
9.3.2	Demand, filtration, and policies . . . . .	190
9.4	The Optimal Control Problem . . . . .	190
9.4.1	Cost structure: holding, shortage, and ordering . . . . .	190
9.4.2	Dynamic programming concept before the equation . . . . .	191
9.4.3	Bellman equation . . . . .	191
9.5	Solving the Known-Demand Bellman Equation . . . . .	191
9.5.1	Concept: removing the immediate piecewise-linear cost . . . . .	191
9.5.2	Change of unknown function . . . . .	191
9.5.3	Concept: the marginal value equation . . . . .	192
9.5.4	The derivative fixed point . . . . .	192

9.5.5	Construction of the solution . . . . .	192
9.5.6	Base-stock policy . . . . .	193
9.6	Learning an Unknown Demand Distribution . . . . .	193
9.6.1	Concept: prior, likelihood, posterior, and predictive demand . . . . .	193
9.6.2	Bayesian updating . . . . .	193
9.6.3	A simple conjugate-prior example . . . . .	194
9.6.4	Concept: the belief-state MDP . . . . .	195
9.6.5	Transition operator in belief space . . . . .	195
9.7	Dynamic Programming with Bayesian Learning . . . . .	196
9.7.1	Concept: optimality with a learned state . . . . .	196
9.7.2	Bellman equation . . . . .	196
9.7.3	Concept: unnormalized beliefs . . . . .	197
9.7.4	Unnormalized Bellman equation . . . . .	197
9.8	The Functionals $Z$ and $Z$ -prime . . . . .	197
9.8.1	Concept: transforming the value functional . . . . .	197
9.8.2	Concept: a derivative equation in belief space . . . . .	198
9.9	Adaptive Base-Stock Levels . . . . .	201
9.9.1	Concept: learning changes the target inventory . . . . .	201
9.10	A Value-Iteration Algorithm . . . . .	202
9.10.1	Concept: fixed-point iteration and marginal value learning . . . . .	202
9.11	A Two-Model Bayesian Inventory Problem . . . . .	202
9.11.1	Concept: posterior odds and Bayes factors . . . . .	202
9.11.2	Exponential demand models . . . . .	203
9.11.3	First iteration . . . . .	203
9.12	Comparison with Bayesian adaptive methods . . . . .	203
9.13	Conclusion . . . . .	204
9.14	Control–ML Dictionary and Diagram . . . . .	204
9.15	Exercises . . . . .	205
<b>10</b>	<b>Deep Learning as Optimal Control</b> . . . . .	<b>207</b>
10.1	Introduction and Literature . . . . .	207
	Chapter notation . . . . .	209
10.2	Deep Learning with a Finite Number of Layers . . . . .	210
10.2.1	Conceptual Overview . . . . .	210
10.2.2	The Finite-Depth Model . . . . .	211
10.2.3	Loss Function and Regularization . . . . .	212
10.2.4	Backpropagation of the Gradient . . . . .	212
10.2.5	Formal Derivation . . . . .	214
10.3	Deep Learning with an Infinite Number of Layers . . . . .	215
10.3.1	Conceptual Overview . . . . .	215
10.3.2	The Continuous-Depth Model . . . . .	216
10.3.3	Gâteaux Derivative . . . . .	217
10.3.4	Computation of the Gradient . . . . .	217
10.3.5	Formal Derivation . . . . .	218
10.3.6	A Scalar Optimal-Control Example . . . . .	218
10.3.7	Existence and uniqueness for large regularization . . . . .	219
10.4	From Neural ODEs to Neural PDEs . . . . .	225
10.4.1	Neural ODEs versus Neural PDEs . . . . .	226

10.4.2	CNN vector fields and convolutional finite differences . . . . .	226
10.4.3	Parabolic Neural PDEs . . . . .	227
10.4.4	Dual system for the parabolic coefficient-control problem . . . . .	228
10.4.5	Hyperbolic Neural PDEs . . . . .	229
10.4.6	Elliptic Neural PDEs . . . . .	230
10.4.7	Adjoint formula for the elliptic Neural PDE . . . . .	231
10.4.8	Summary: finite-dimensional versus infinite-dimensional depth models . . . . .	233
10.5	Concluding Remarks . . . . .	234
10.6	Control–ML Dictionary and Diagram . . . . .	234
10.7	Exercises . . . . .	235
<b>11</b>	<b>Learning Dynamical Systems under Full and Partial Observation</b>	<b>237</b>
11.1	Introduction and Literature . . . . .	237
	Comparison with classical and modern identification viewpoints . . . . .	237
	Chapter notation . . . . .	239
11.2	Learning Linear Dynamical Systems from Full State Observations . . . . .	240
11.2.1	Conceptual Overview . . . . .	240
11.2.2	Full State Observation . . . . .	240
11.2.3	Least-Squares Identification . . . . .	241
11.2.4	A Regularized Machine-Learning Formulation . . . . .	241
11.2.5	Dual Problem . . . . .	243
11.3	Learning with Partial Information . . . . .	243
11.3.1	Conceptual Overview . . . . .	243
11.3.2	Model and Regularized Reconstruction . . . . .	245
11.3.3	Gradient of the Functional . . . . .	247
11.3.4	Necessary Conditions of Optimality . . . . .	249
11.3.5	Second Derivative . . . . .	250
11.3.6	Gradient Descent . . . . .	251
11.3.7	A structure-preserving descent method . . . . .	255
11.3.8	Convergence of the structure-preserving method . . . . .	256
11.3.9	Duality for the partial-information problem . . . . .	261
11.4	Concluding Remarks . . . . .	264
11.5	Control–ML Dictionary and Diagram . . . . .	264
11.6	Exercises . . . . .	265