ADAPTATION IN DYNAMICAL SYSTEMS

In the context of this book, adaptation is taken to mean a feature of a system aimed at achieving the best possible performance when mathematical models of the environment and the system itself are not fully available. This has applications ranging from theories of visual perception and the processing of information to the more technical problems of friction compensation and adaptive classification of signals in fixed-weight recurrent neural networks.

Largely devoted to the problems of adaptive regulation, tracking and identification, this book presents a unifying system-theoretic view on the problem of adaptation in dynamical systems. Special attention is given to systems with nonlinearly parametrized models of uncertainty. Concepts, methods, and algorithms given in the text can be successfully employed in wider areas of science and technology. The detailed examples and background information make this book suitable for a wide range of researchers and graduates in cybernetics, mathematical modeling, and neuroscience.

IVAN TYUKIN is an RCUK Academic Fellow in the Department of Mathematics, University of Leicester. His research and scientific interests cover many areas, including the analysis, modeling, and synthesis of systems with fragile, nonlinear, chaotic, and meta-stable dynamics.

ADAPTATION IN DYNAMICAL SYSTEMS

IVAN TYUKIN

University of Leicester and
Saint-Petersburg State Electrotechnical University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314-321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi - 110025, India

103 Penang Road, #05-06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9780521198196

First published 2011

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging in Publication data
Tyukin, Ivan.
Adaptation in Dynamical Systems / Ivan Tyukin.
p. cm
Includes bibliographical references and index.
ISBN 978-0-521-19819-6 (hardback)
1. Dynamics. 2. Control theory–Mathematical models.
3. Neurosciences–Mathematics. I. Title.
QA845.T94 2011
515´.39–dc22 2010051427

ISBN 978-0-521-19819-6 Hardback

Contents

Preface

Adaptation is amongst the most familiar and wide spread phenomena in nature. Since the early days of the nineteenth century it has puzzled researchers in broad areas of science. Since it had often been observed in responsive behaviors of biological systems, adaptation was initially understood as a regulatory mechanism that helps an animal to survive in a changing environment. Later the notion of adaptation was adopted in wider fields of science and engineering.

As a theoretical discipline it began to emerge as a branch of control theory during the first half of the twentieth century. Its beginning was marked by publications discussing basic principles of adaptation and its merits for engineering. Imprecise technology and mechanisms were, perhaps, amongst the strongest practical motivations for such a theory at that time. Various notions of adaptation were adopted by engineers and theoreticians in order to grasp, understand, and implement relevant features of this phenomenon in practice. The first applications of the new theory were simple schemes for extremal control of mechanical systems; these systems could be described by just a few linear ordinary differential equations. Since then adaptive controllers have evolved to encompass substantially more complex devices. The controlling devices themselves can now be viewed as nonlinear dynamical systems with specific input–output properties. Methods for the design and analysis of such systems are currently recognized by many in terms of the theory of adaptive control and systems identification.

Because the initial motivation to develop a theory of adaptation was driven mainly by the demands of mechanical engineering and the need for robust design of otherwise imprecise machines, the domain of application of the theories of adaptation was naturally restricted to the realm of artificial devices and engineering. The focus of the developing theory was restricted, in particular, to the problems of control of a relatively narrow class of well-studied and modeled mechanical systems, many of which were stable in the Lyapunov sense, for which the values of some parameters and variables are unknown and cannot be measured explicitly.

Yet, the potential role of the theory of adaptation was much wider and broader. It has become evident recently that there exists a demand for a systematic theory of adaptation outside of the domain of engineering.

Understanding basic mechanisms and principles of adaptation and regulation is recognized as relevant in physics, chemistry, biology, and brain sciences (Sontag 2004; Fradkov 2005). Because of the huge complexity of the phenomena studied in these domains, using the standard language of each particular science for systematic studies of the phenomenon of adaptation might not be adequate. Therefore in these areas system-theoretic views, irrespective of the particular subjects of study, have exceptional potential.

Apart from in the natural sciences, the needs for further development of the theory of adaptation are evident in handling complex artificial systems. This is especially true when changes in the working environment cannot be predicted a priori or there is a substantial degree of uncertainty about the system's internal state. Although there is a large literature on the theory of adaptive systems, both in the theoretical and in the applied domain, there are several issues preventing explicit application of classical recipes of adaptive control in these fields. These issues with classical schemes are

- the necessity to have a precise mathematical model of a controlled system,
- the requirement that models of uncertainties are linear or convex with respect to unknown or uncertain variables,
- the assumption that the target dynamics is stable in the Lyapunov sense,
- the assumption that a corresponding Lyapunov function for the target motions is available (Sastry and Bodson 1989; Narendra and Annaswamy 1989; Krstić *et al.* 1995; Ljung 1999; Eykhoff 1975; Bastin *et al.* 1992; Fradkov *et al.* 1999).

Every one of these requirements alone limits the role of the existing theory of adaptive systems in solving relevant problems in science. Altogether they constitute the "standard" approach which applies to several canonical cases, which are limited even within the realm of engineering.

The purpose of this work is to contribute towards extending the existing theory of adaptation and adaptive control beyond the scope of its usual applications in engineering to new and non-conventional areas, such as neuroscience and mathematical modeling of biological systems. It is hoped that this extension will create additional opportunities for control theorists to apply their expertise in novel and still developing fields of science; it will also help to expand the synthetic and analytical functions of systems and control theory into the natural sciences.

The focus of this book on the analysis of possible adaptation mechanisms in systems with nonlinear parametrization and unstable target dynamics was influenced by the author's work in the Laboratory for Perceptual Dynamics, RIKEN Brain Science Institute, Japan from November 2001 to March 2007. Neural systems of living organisms, and ultimately the human brain, were the source of inspiration. It became clear very quickly that the standard tools and methods in the arsenal of conventional adaptive control theory do not offer an acceptable explanation for the versatility and robustness of neural systems working in an uncertain environment. The aim therefore was to enhance the theory by making it suitable for the analysis and synthesis of adaptive schemes for nonlinear dynamical systems:

- with potentially Lyapunov-unstable and non-equilibrium target dynamics;
- when explicit definition of the target sets is not possible;
- using minimal, *qualitative*, macro-information about the system, and also allowing substantial uncertainty about the specific mathematical model of the system;
- allowing uncertainty models that are maximally adequate to describe the physical laws of processes and phenomena in the system.

The necessary ingredients of this extended theory of adaptation follow naturally from the logic of its development: from basic principles of the system's organization in the presence of uncertainties to specific laws of regulation. These ingredients include

(1) *methods for analysis* of basic input–output properties of the nonlinear systems; they should allow incomplete knowledge of equations describing the system dynamics; and they also should apply both to stable and to unstable systems;
(2) *principles and methods* of adaptation to disturbances that are unknown a priori and unavailable for measurement; the principles should rely exclusively on the fundamental physical properties of the systems considered; and the adaptation mechanisms should be able to realize these principles using adequate physical models of uncertainties and requiring a minimal amount of measurement information.

The following topics received particular attention: analysis of the completeness, realizability, and state boundedness of interconnections of uncertain dynamical systems; conditions ensuring convergence of the system's state to the target sets and their neighborhoods; designing laws of adaptive regulation and parameter estimation of nonlinearly parametrized models; characterizing the quality of the transient dynamics in systems with uncertainties; and parametric, signal/functional perturbations. In order to provide the reader with the necessary background and also to support our own argumentation a brief review of major classical concepts of adaptation is included.

The content of the book is based largely on the work I had the privilege to carry out together with my colleagues and co-authors.[1] The structure of the book can be summarized as follows. The text is organized into three large parts. The first part (Chapters 1–3) contains mainly introductory and preliminary results. Proofs of lemmas and theorems presented in this introductory part are kept within the main text.

In Chapter 1 we provide an informal discussion of the notion of adaptation followed by an overview of the range of specific problems considered in the text.

Chapter 2 contains background and preliminary results such as basic notions of stability, a very brief introduction to the method of Lyapunov functions, and a particularly important result on the exponential stability of the origin for a class of linear systems of ordinary equations with skew-symmetric matrices.

In Chapter 3 we review and analyze conventional approaches to the problem of adaptive control of nonlinear systems. We formulate the main theoretical and practical issues arising in these standard approaches (Fomin *et al.* 1981; Fradkov 1990; Narendra and Annaswamy 1989; Krstić *et al.* 1995) and their mathematical statements of the problem. These issues include the ambiguity of standard mathematical notions of an adaptive system, performance measures, limitations on defining the system's target sets,[2] restricted classes of the uncertainty models, and requirements for precise knowledge of the mathematical model of a system.

The second part, Chapters 4 and 5, presents the main theoretical results developed in the monograph. In order to preserve the integrity of the text, proofs of statements formulated in this part are given in appendices at the ends of these chapters.

In Chapter 4 we consider nonlinear systems defined in terms of their "input-to-output" and "input-to-state" characterizations given by mappings, or operators in functional, L_p, spaces. We introduce mathematical tools for the analysis of interconnections of dynamical systems with input–output (input–state) operators that are locally bounded in state and provide a formal statement of the problem for functional synthesis of an adaptive system. We demonstrate how this problem can be solved. The solution to the problem of functional synthesis of an adaptive system allows us to formulate various principles of its organization at the macroscopic level: the separation principle, the bottle-neck principle, and the emergence of weakly attracting sets in the interconnections of systems with contracting and wandering dynamics. The latter result is based on Tyukin *et al.* (2008a).

[1] This includes earlier texts such as Tyukin and Terekhov (2008).

[2] One of the most severe restrictions is the requirement for the target dynamics to be globally stable in the sense of Lyapunov. In addition, there is a necessity to specify target sets of the adaptive system a priori. The latter condition either requires prior identification of the system, which contradicts the very essence of adaptive behavior, or leads to enforcing motions that are not necessarily inherent and, hence, optimal to a physical system itself.

In Chapter 5 we utilize the principles derived in the previous chapter in order to provide an adequate statement of the problem of adaptive control and regulation of nonlinear dynamical systems. Its distinctive features are that the uncertainty models are allowed to be nonlinearly parametrized, mathematical models of the system need not be known precisely, the target dynamics is not restricted exclusively to globally Lyapunov-stable motions, and the target sets could be defined implicitly – as invariant sets of an auxiliary dynamical system. Generally, the problem is stated as that of *regulating the influence of uncertainties on the target dynamics to some functional space*. This allows one to refrain from explicit use of the method of Lyapunov functions and, hence, avoid its limitations.

We also consider several specific problems that have substantial theoretical and practical interest:

- adaptive regulation to invariant sets;
- adaptation in interconnected systems;
- state and parameter inference for systems with nonlinear parametrization of uncertainty.

In order to solve these problems two synthesis strategies were developed: the method of the *virtual adaptation algorithm* presented in Tyukin *et al.* (2007b) and the strategy based on purposeful introduction of unstable attracting sets into the system's state space (Tyukin *et al.* 2008a).

In the third part of the book (Chapters 6–8) we illustrate how the theory can be used to solve a number of practical problems of control, processing of information, and identification in mechanics, experimental biophysics, and computer and cognitive science. In particular, we consider the problem of adaptive classification in neural networks with fixed weights, the problem of identifying the dynamics of neuronal cells, and the problem of invariant recognition of spatially distributed information. We discuss why existing techniques cannot be successfully applied to solve these problems, or their application yields practically inefficient outcomes. The content of this part is based on Tyukin *et al.* (2008b), Tyukin *et al.* (2009), and Fairhurst *et al.* (2010).

This book would never have seen the light of day without the continuous support, help, and encouragement I received from many people with whom I have had the honor of working. I would like to express my deep gratitude to Professor V. A. Terekhov, my teacher, friend, and co-author, for his help, fruitful and motivating discussions of the philosophical foundations of the problem of adaptation, and unlimited patience. I am grateful to my colleagues and co-authors Cees van Leeuwen, Danil Prokhorov, Henk Nijmeijer, Erik Steur, David Fairhurst, Alexey Semyanov, and Inseon Song who contributed to the development of the ideas in the monograph. I am grateful to Dr Steven Holt and his colleagues for proof-reading

and editing the monograph at the final stage of production. Finally, I am indebted to my dear wife Tanya, who contributed to the applied side of the project, assisted with the artwork, and also helped me enormously to summarize the results during the later stage of the production of the manuscript. My own personal role was limited to mere listening, interpretation, and writing. As is unfortunately the case in scientific endeavors, errors are inevitable companions. Even though I tried to avoid these unwelcome companions, my own journey is unlikely to be an exception, for which I fully accept sole responsibility. I would therefore be extremely grateful to readers, should they wish to help by contacting me when an error is found.

Notational conventions

Throughout the text the following notational conventions apply.

- Symbol $\mathbb{R}$ defines the field of real numbers and $\mathbb{R}_{\geq c} = \{x \in \mathbb{R} | x \geq c\}$; $\mathbb{N}$ defines the set of natural numbers; and $\mathbb{Z}$ denotes the set of whole numbers or integers.
- Symbol $\mathbb{R}^n$ stands for an n-dimensional linear space over the field of reals.
- $\mathcal{C}^k$ denotes the space of functions that are at least k times differentiable.
- Symbol $\mathcal{K}$ denotes the class of all strictly increasing functions $\kappa : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\kappa(0) = 0$; symbol $\mathcal{K}_\infty$ denotes the class of all functions $\kappa \in \mathcal{K}$ such that $\lim_{s\to\infty} \kappa(s) = \infty$.
- Let Ω be a set, then by $\mathcal{S}\{\Omega\}$ we denote the set of all subsets of Ω.
- $\|\mathbf{x}\|$ denotes the Euclidian norm of $\mathbf{x} \in \mathbb{R}^n$.
- The notation $|\cdot|$ stands for the absolute value of a scalar.
- The notation $\mathrm{sign}(\cdot)$ denotes the signum function.
- By $L_p^n[t_0, T]$, where $t_0 \geq 0$, $T \geq t_0$, $p \geq 1$, we denote the space of all functions $\mathbf{f} : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ such that

$$\|\mathbf{f}\|_{p,[t_0,T]} = \left(\int_{t_0}^{T} \|\mathbf{f}(\tau)\|^p \, d\tau \right)^{1/p} < \infty.$$

- The notation $\|\mathbf{f}\|_{p,[t_0,T]}$ denotes the $L_p^n[t_0, T]$-norm of $\mathbf{f}(t)$.
- By $L_\infty^n[t_0, T]$, $t_0 \geq 0$, $T \geq t_0$, we denote the space of all functions $\mathbf{f} : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ such that

$$\|\mathbf{f}\|_{\infty,[t_0,T]} = \mathrm{ess\,sup}\{\|\mathbf{f}(t)\|, t \in [t_0, T]\} < \infty,$$

  and $\|\mathbf{f}\|_{\infty,[t_0,T]}$ stands for the $L_\infty^n[t_0, T]$-norm of $\mathbf{f}(t)$.
- Let $\mathcal{A}$ be a set in $\mathbb{R}^n$, $\mathbf{x} \in \mathbb{R}^n$, and let $\|\cdot\|$ be the usual Euclidean norm in $\mathbb{R}^n$. By the symbol $\|\cdot\|_{\mathcal{A}}$ we denote the following induced norm:

$$\|\mathbf{x}\|_{\mathcal{A}} = \inf_{\mathbf{q}\in\mathcal{A}} \{\|\mathbf{x} - \mathbf{q}\|\}.$$

- Let $\Delta \in \mathbb{R}_{\geq 0}$, then the notation $\|\mathbf{x}\|_{\mathcal{A}_\Delta}$ stands for the following equality:

$$\|\mathbf{x}\|_{\mathcal{A}_\Delta} = \begin{cases} \|\mathbf{x}\|_{\mathcal{A}} - \Delta, & \|\mathbf{x}\|_{\mathcal{A}} > \Delta, \\ 0, & \|\mathbf{x}\|_{\mathcal{A}} \leq \Delta. \end{cases}$$

- The symbol $\|\cdot\|_{\mathcal{A}_\infty,[t_0,t]}$ is defined as follows:

$$\|\mathbf{x}(\tau)\|_{\mathcal{A}_\infty,[t_0,t]} = \sup_{\tau \in [t_0,t]} \|\mathbf{x}(\tau)\|_{\mathcal{A}}.$$

- Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ be given. The function $\mathbf{f}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$ is said to be locally bounded if for any $\|\mathbf{x}\| < \delta$, $\delta \in \mathbb{R}_{>0}$ there exists a constant $D(\delta) > 0$ such that $\|\mathbf{f}(\mathbf{x})\| \leq D(\delta)$.
- Let Γ be an $n \times n$ square matrix, then $\Gamma > 0$ denotes a positive definite (symmetric) matrix. (Γ^{-1} is the inverse of Γ). By $\Gamma \geq 0$ we denote a positive semi-definite matrix.
- We reserve $\|\mathbf{x}\|^2_\Gamma$ to denote the quadratic form $\mathbf{x}^T \Gamma \mathbf{x}$, where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{x}^T$ is the transpose of $\mathbf{x}$.
- Symbols $\lambda_{\min}(\Gamma)$ and $\lambda_{\max}(\Gamma)$ stand for the minimal and maximal eigenvalues of Γ, respectively.
- By the symbol I we denote the identity matrix.
- The solution of a system of differential equations $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t, \boldsymbol{\theta}, \mathbf{u}(t))$, $\mathbf{u} : \mathbb{R}_{\geq 0} \to \mathbb{R}^m$, $\boldsymbol{\theta} \in \mathbb{R}^d$ passing through point $\mathbf{x}_0$ at $t = t_0$ will be denoted for $t \geq t_0$ as $\mathbf{x}(t, \mathbf{x}_0, t_0, \boldsymbol{\theta}, \mathbf{u})$, or simply as $\mathbf{x}(t)$ if it is clear from the context what the values of $\mathbf{x}_0$ and $\boldsymbol{\theta}$ are and how the function $\mathbf{u}(t)$ is defined.
- Let $\mathbf{u} : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}_{\geq 0} \to \mathbb{R}^m$ be a function of state $\mathbf{x}$, parameters $\hat{\boldsymbol{\theta}}$, and time t. Let in addition both $\mathbf{x}$ and $\hat{\boldsymbol{\theta}}$ be functions of t. Then, when the arguments of $\mathbf{u}$ are clearly defined by the context, we will simply write $\mathbf{u}(t)$ instead of $\mathbf{u}(\mathbf{x}(t), \hat{\boldsymbol{\theta}}(t), t)$.
- When dealing with vector fields and partial derivatives we will use the following extended notion of the Lie derivative of a function. Let it be the case that $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{x}$ can be partitioned as follows: $\mathbf{x} = \mathbf{x}_1 \oplus \mathbf{x}_2$, where $\mathbf{x}_1 \in \mathbb{R}^q$, $\mathbf{x}_1 = (x_{11}, \ldots, x_{1q})^T$, $\mathbf{x}_2 \in \mathbb{R}^p$, $\mathbf{x}_2 = (x_{21}, \ldots, x_{2p})^T$, $q + p = n$, and $\oplus$ denotes concatenation of two vectors. We define $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^n$ such that $\mathbf{f}(\mathbf{x}, \boldsymbol{\theta}, t) = \mathbf{f}_1(\mathbf{x}, \boldsymbol{\theta}, t) \oplus \mathbf{f}_2(\mathbf{x}, \boldsymbol{\theta}, t)$, where $\mathbf{f}_1 : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^q$, $\mathbf{f}_1(\cdot) = (f_{11}(\cdot), \ldots, f_{1q}(\cdot))^T$, $\mathbf{f}_2 : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^p$, and $\mathbf{f}_2(\cdot) = (f_{21}(\cdot), \ldots, f_{2p}(\cdot))^T$. Then $L_{\mathbf{f}_i(\mathbf{x},\boldsymbol{\theta},t)} \psi(\mathbf{x}, t)$, $i \in \{1, 2\}$, denotes the Lie derivative of the function $\psi(\mathbf{x}, t)$ with respect to the vector field $\mathbf{f}_i(\mathbf{x}, \boldsymbol{\theta}, t)$:

$$L_{\mathbf{f}_i(\mathbf{x},\boldsymbol{\theta},t)} \psi(\mathbf{x}, t) = \sum_{j=1}^{\dim \mathbf{x}_i} \frac{\partial \psi(\mathbf{x}, t)}{\partial x_{ij}} f_{ij}(\mathbf{x}, \boldsymbol{\theta}, t).$$

- Let $\mathbf{f}, \mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be differentiable vector fields. Then the symbol $[\mathbf{f}, \mathbf{g}]$ stands for the Lie bracket:
$$[\mathbf{f}, \mathbf{g}] = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\mathbf{g} - \frac{\partial \mathbf{g}}{\partial \mathbf{x}}\mathbf{f}.$$
The adjoint representation of the Lie bracket is defined as
$$\mathrm{ad}^0_f \mathbf{g} = \mathbf{g}, \qquad \mathrm{ad}^k_{\mathbf{f}} \mathbf{g} = [\mathbf{f}, \mathrm{ad}^{k-1}_{\mathbf{f}} \mathbf{g}].$$