

Natural Computing Series

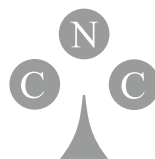
Thomas Bäck
Christophe Foussette
Peter Krause



Contemporary Evolution Strategies

 Springer

Natural Computing Series



Series Editors: G. Rozenberg

Th. Bäck A.E. Eiben J.N. Kok H.P. Spink

Leiden Center for Natural Computing

Advisory Board: S. Amari G. Brassard K.A. De Jong C.C.A.M. Gielen T. Head
L. Kari L. Landweber T. Martinetz Z. Michalewicz M.C. Mozer E. Oja
G. Păun J. Reif H. Rubin A. Salomaa M. Schoenauer H.-P. Schwefel C. Torras
D. Whitley E. Winfree J.M. Zurada

For further volumes:

<http://www.springer.com/series/4190>

Thomas Bäck • Christophe Foussette
Peter Krause

Contemporary Evolution Strategies

 Springer

Thomas Bäck
Leiden University
Leiden
The Netherlands

Christophe Foussette
Peter Krause
divis intelligent solutions GmbH
Dortmund
Germany

Series Editors

G. Rozenberg (Managing Editor)

Th. Bäck, J.N. Kok, H.P. Spaink
Leiden Center for Natural Computing
Leiden University
Leiden, The Netherlands

A.E. Eiben
Vrije Universiteit Amsterdam
The Netherlands

ISSN 1619-7127 Natural Computing Series

ISBN 978-3-642-40136-7

DOI 10.1007/978-3-642-40137-4

Springer Heidelberg New York Dordrecht London

ISBN 978-3-642-40137-4 (eBook)

Library of Congress Control Number: 2013950140

© Springer-Verlag Berlin Heidelberg 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

1	Introduction	1
1.1	Background and Motivation	1
1.2	Structure of the Book	4
1.3	Notation	4
1.4	Source Code	5
2	Evolution Strategies	7
2.1	Introduction	7
2.1.1	Optimization	7
2.1.2	Evolution Strategies as a Specialization of Evolutionary Algorithms	8
2.1.3	Mutation in \mathbb{R}^n	10
2.2	Algorithms	13
2.2.1	From the (1+1)-ES to the CMA-ES	14
2.2.2	Modern Evolution Strategies	21
2.3	Further Aspects of ES	43
2.3.1	Constraint Handling	44
2.3.2	Beyond Real-Valued Search Spaces	44
2.3.3	Multiobjective Optimization	45
3	Taxonomy of Evolution Strategies	47
3.1	Development Strands of Modern Evolution Strategies	47
3.1.1	Overview	47
3.1.2	Restart Heuristics	49
3.1.3	Methods for Adapting Mutation Parameters	50
3.1.4	Methods for Avoiding Function Evaluations	51
3.2	Characteristics of Modern Evolution Strategies	51
3.2.1	Computational Effort	51
3.2.2	Convergence Behavior	52
3.3	Recommendations for Practical Use	53
3.3.1	Global Optimization	53
3.3.2	High-Dimensional Search Spaces	53

4 Empirical Analysis	55
4.1 Measuring Efficiency	55
4.1.1 The FCE Measure	56
4.1.2 The ERT Measure	57
4.2 Experiments	57
4.2.1 Selection of Algorithms	57
4.2.2 Technical Aspects	58
4.2.3 Analysis	61
4.3 Results	63
4.3.1 Ranks by FCE	63
4.3.2 Discussion of Results	63
4.4 Further Analysis for $n = 100$	68
5 Summary	85
Bibliography	87

List of Figures

Fig. 2.1	Mutation ellipsoids representing $N(\mathbf{0}, \mathbf{I})$, $N(\mathbf{0}, \text{diag}(\delta^2))$ and $N(\mathbf{0}, \mathbf{C})$ (from <i>left to right</i>)	12
Fig. 3.1	Development strands of modern evolution strategies.....	49
Fig. 4.1	Example of a convergence plot	56
Fig. 4.2	Ranking for $C_t = 100 \cdot n$ with <i>best-of-1</i> approach. 1: $(1 + 1)$ -ES, 2: (μ, λ) -MSC-ES, 3: (μ_w, λ) -CMA-ES, 4: Active-CMA-ES, 5: $(1, 4_m^s)$ -CMA-ES, 6: xNES, 7: DR1, 8: DR2, 9: DR3, A: (μ, λ) -CMSA-ES, B: $(1 + 1)$ -Cholesky-CMA-ES, C: LS-CMA-ES, D: $(1 + 1)$ -Active-CMA-ES, E: sep-CMA-ES	64
Fig. 4.3	Rankings for $C_t = 100 \cdot n$ with <i>best-of-5</i> approach. 1: $(1 + 1)$ -ES, 2: (μ, λ) -MSC-ES, 3: (μ_w, λ) -CMA-ES, 4: Active-CMA-ES, 5: $(1, 4_m^s)$ -CMA-ES, 6: xNES, 7: DR1, 8: DR2, 9: DR3, A: (μ, λ) -CMSA-ES, B: $(1 + 1)$ -Cholesky-CMA-ES, C: LS-CMA-ES, D: $(1 + 1)$ -Active-CMA-ES, E: sep-CMA-ES	65
Fig. 4.4	Rankings for $C_t = 50 \cdot n$ with <i>best-of-1</i> approach. 1: $(1 + 1)$ -ES, 2: (μ, λ) -MSC-ES, 3: (μ_w, λ) -CMA-ES, 4: Active-CMA-ES, 5: $(1, 4_m^s)$ -CMA-ES, 6: xNES, 7: DR1, 8: DR2, 9: DR3, A: (μ, λ) -CMSA-ES, B: $(1 + 1)$ -Cholesky-CMA-ES, C: LS-CMA-ES, D: $(1 + 1)$ -Active-CMA-ES, E: sep-CMA-ES	65
Fig. 4.5	Rankings for $C_t = 50 \cdot n$ with <i>best-of-5</i> approach. 1: $(1 + 1)$ -ES, 2: (μ, λ) -MSC-ES, 3: (μ_w, λ) -CMA-ES, 4: Active-CMA-ES, 5: $(1, 4_m^s)$ -CMA-ES, 6: xNES, 7: DR1, 8: DR2, 9: DR3, A: (μ, λ) -CMSA-ES, B: $(1 + 1)$ -Cholesky-CMA-ES, C: LS-CMA-ES, D: $(1 + 1)$ -Active-CMA-ES, E: sep-CMA-ES	66

Fig. 4.6	Rankings for $C_t = 25 \cdot n$ with <i>best-of-1</i> approach. 1: (1 + 1)-ES, 2: (μ, λ) -MSC-ES, 3: (μ_W, λ) -CMA-ES, 4: Active-CMA-ES, 5: $(1, 4_m^s)$ -CMA-ES, 6: xNES, 7: DR1, 8: DR2, 9: DR3, A: (μ, λ) -CMSA-ES, B: (1 + 1)-Cholesky-CMA-ES, C: LS-CMA-ES, D: (1 + 1)-Active-CMA-ES, E: sep-CMA-ES	66
Fig. 4.7	Rankings for $C_t = 25 \cdot n$ with <i>best-of-5</i> approach. 1: (1 + 1)-ES, 2: (μ, λ) -MSC-ES, 3: (μ_W, λ) -CMA-ES, 4: Active-CMA-ES, 5: $(1, 4_m^s)$ -CMA-ES, 6: xNES, 7: DR1, 8: DR2, 9: DR3, A: (μ, λ) -CMSA-ES, B: (1 + 1)-Cholesky-CMA-ES, C: LS-CMA-ES, D: (1 + 1)-Active-CMA-ES, E: sep-CMA-ES	67
Fig. 4.8	Convergence plot for test function f_1 (sphere function) showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs.....	71
Fig. 4.9	Convergence plot for test function f_2 showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the <i>solid line</i> represents their mean.....	71
Fig. 4.10	Convergence plot for test function f_3 showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the <i>solid line</i> represents their mean.....	72
Fig. 4.11	Convergence plot for test function f_4 showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the <i>solid line</i> represents their mean.....	72

Fig. 4.12 Convergence plot for test function f_5 showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the *solid line* represents their mean..... 73

Fig. 4.13 Convergence plot for test function f_6 showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the *solid line* represents their mean..... 73

Fig. 4.14 Convergence plot for test function f_7 showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the *solid line* represents their mean..... 74

Fig. 4.15 Convergence plot for test function f_8 showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the *solid line* represents their mean..... 74

Fig. 4.16 Convergence plot for test function f_9 showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the *solid line* represents their mean..... 75

Fig. 4.17 Convergence plot for test function f_{10} showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the *solid line* represents their mean..... 75

Fig. 4.18	Convergence plot for test function f_{11} showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the <i>solid line</i> represents their mean.....	76
Fig. 4.19	Convergence plot for test function f_{12} showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the <i>solid line</i> represents their mean.....	76
Fig. 4.20	Convergence plot for test function f_{13} showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the <i>solid line</i> represents their mean.....	77
Fig. 4.21	Convergence plot for test function f_{14} showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the <i>solid line</i> represents their mean.....	77
Fig. 4.22	Convergence plot for test function f_{15} showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the <i>solid line</i> represents their mean.....	78
Fig. 4.23	Convergence plot for test function f_{16} showing the order of magnitude of fitness value normalized w.r.t. the fitness of the initial search point for the number of function evaluations $\{100, 200, \dots, 1,000\}$. Error bars reflect the 20 % respectively 80 % quantiles of the 1,000 conducted runs and the <i>solid line</i> represents their mean.....	78