
Evolution Strategies

FACHHOCHSCHULE VORARLBERG
MASTER INFORMATIK

DOZENT

HANS-GEORG BEYER

VORGELEGT VON

GRIESSER, RIEDMANN

DORNBIRN, 12.02.2017

Kurzreferat

Diese Übung dient zum Vergleich der zuvor implementierten evolutionären Strategien. Dazu wird zuerst der Einfluss der Parameter μ und τ untersucht, dann die $(\mu/\mu_I, \lambda)$ - σ -SA-ES mit der (1+1)-ES mit 1/5 Regel verglichen und zuletzt der Algorithmus am verrauschten Kugelmodell und der RASTRIGIN-Funktion getestet. Anschließend wird noch die CMSA Strategie für die Linsenoptimierung evaluiert.

Inhaltsverzeichnis

1	Evaluierung $(\mu/\mu_I, \lambda)$-σ-SA-ES	4
1.1	Einfluss von μ	4
1.2	Einfluss von τ	5
1.3	Vergleich (1+1)-ES mit 1/5 Regel	8
1.4	Verrauschtes Kugelmodell	9
1.5	RASTRIGIN-Funktion	12
2	CMSA-ES Linsenoptimierung	18

1. Evaluierung $(\mu/\mu_I, \lambda)$ - σ -SA-ES

1.1 Einfluss von μ

Der Einfluss von μ auf die $(\mu/\mu_I, 10)$ - σ -SA-ES wird am Kugelmodell mit $N = 100$ untersucht. Abbruchbedingungen sind $\sigma = 10^{-5}$ oder 20.000 Generationen. Abbildung 1 zeigt die Fitness-Dynamiken für $\mu = 1$ bis $\mu = 9$.

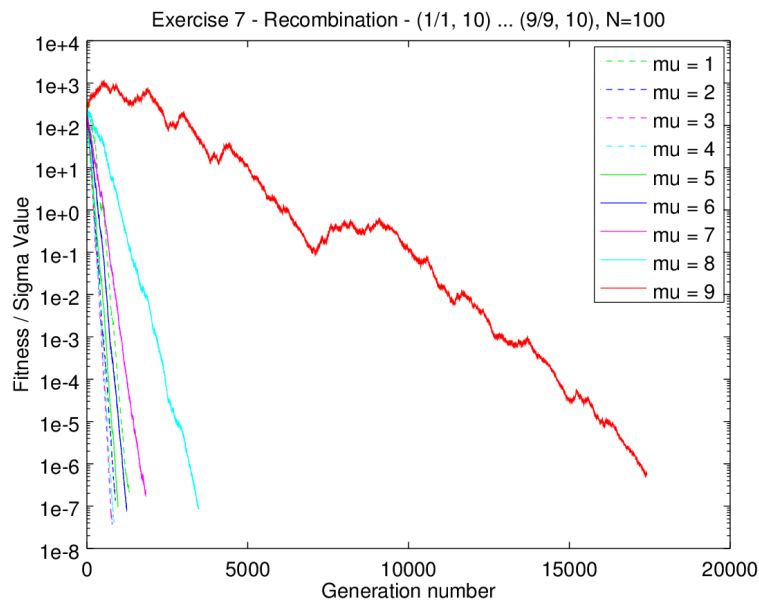


Abbildung 1: Aufgabe 7 - Rekombination - (1/1, 10) ... (9/9, 10), N=100

Daraus ist ersichtlich, dass sich die Laufzeit bei Erhöhung von μ im Verhältnis zu λ zuerst verbessert und ab einem gewissen Punkt wieder verschlechtert. Das Optimum liegt in etwa bei $\mu = 3$, also bei $\mu/\lambda = 1/3$.

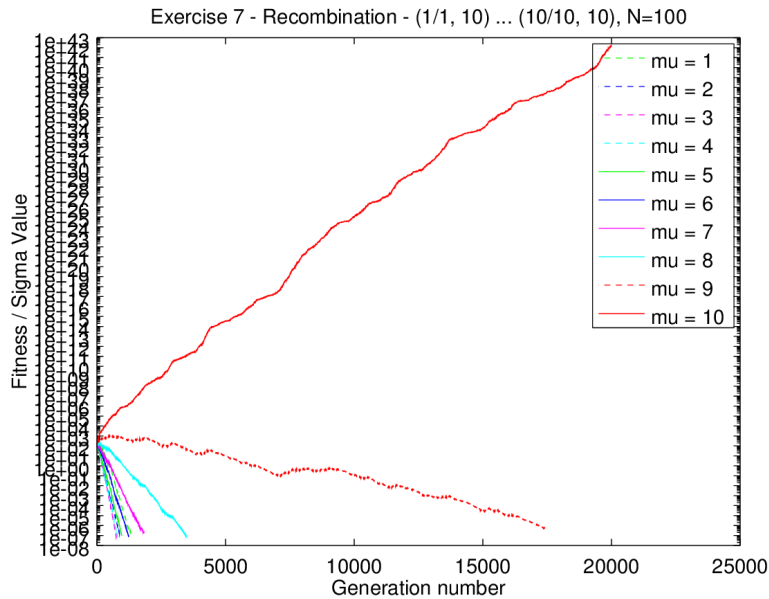


Abbildung 2: Aufgabe 7 - Rekombination - (1/1, 10) ... (10/10, 10), N=100

Abbildung 2 zeigt das Verhalten bei $\mu = 10$. Wird $\mu = \lambda$ gewählt, fällt der Teil der Selektion im Algorithmus weg, da immer alle Nachkommen für die Mutation des nächsten Eltern hergenommen werden und es tritt keine Verbesserung ein, sondern eine Verschlechterung der Fitness.

1.2 Einfluss von τ

Der Einfluss von τ auf die (3/3, 10)- σ -SA-ES wird am Kugelmodell mit $N = 100$ untersucht. Abbruchbedingungen sind $\sigma = 10^{-5}$ oder 4.000 Generationen. Die Abbildungen 3, 4 und 5 zeigen, dass, je größer τ wird, desto schneller konvergiert der Algorithmus. Das kommt daher, dass τ der sogenannte Lernparameter ist und die Geschwindigkeit und Genauigkeit der Selbstadaption angibt.

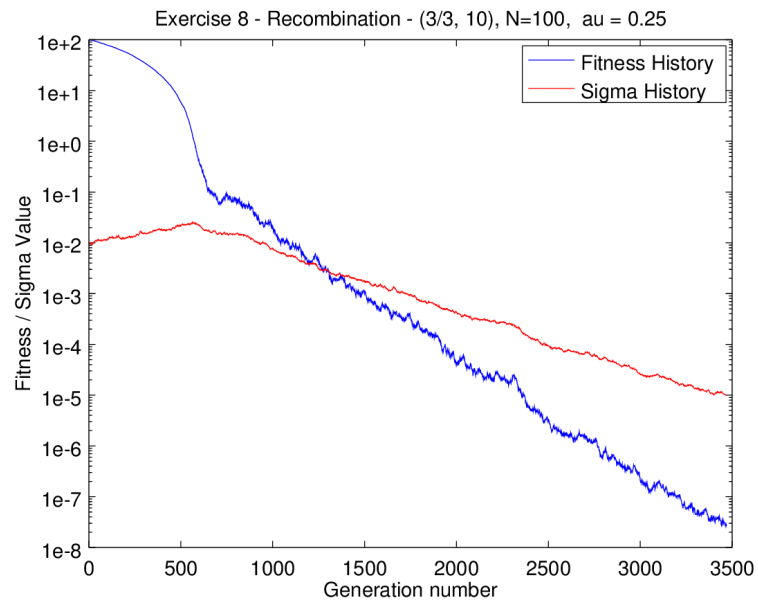


Abbildung 3: Aufgabe 8 - Rekombination - (3/3, 10), N=100, $\tau = 0.25$

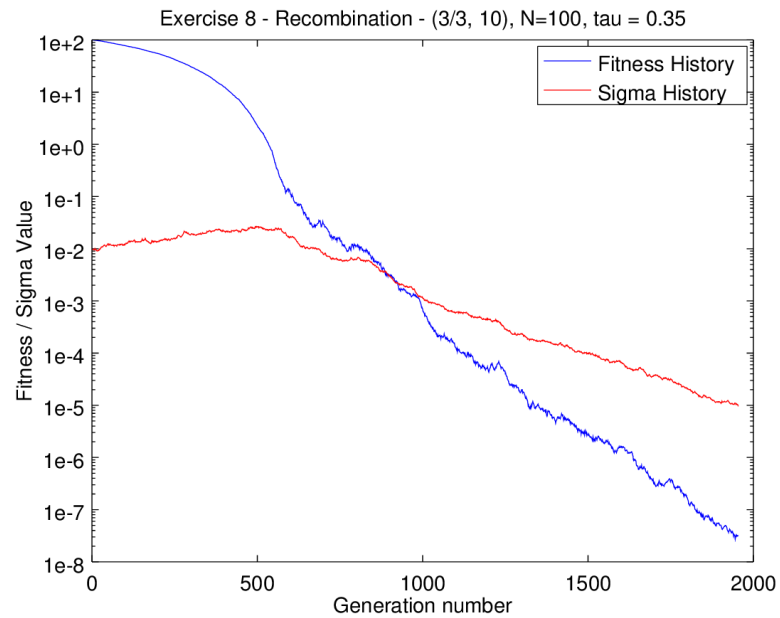


Abbildung 4: Aufgabe 8 - Rekombination - (3/3, 10), N=100, $\tau = 0.35$

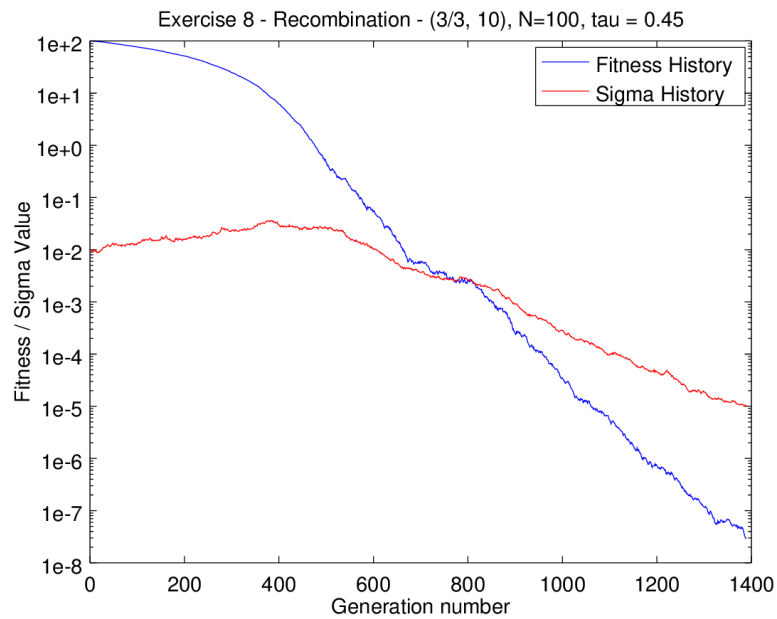


Abbildung 5: Aufgabe 8 - Rekombination - (3/3, 10), N=100, $\tau = 0.45$

Abbildung 6 zeigt das Verhalten bei $\tau = 0$. Es ist zu erkennen, dass Sigma konstant bleibt, da der Lernparameter 0 ist. Dadurch wird auch irgendwann keine Verbesserung mehr gefunden.

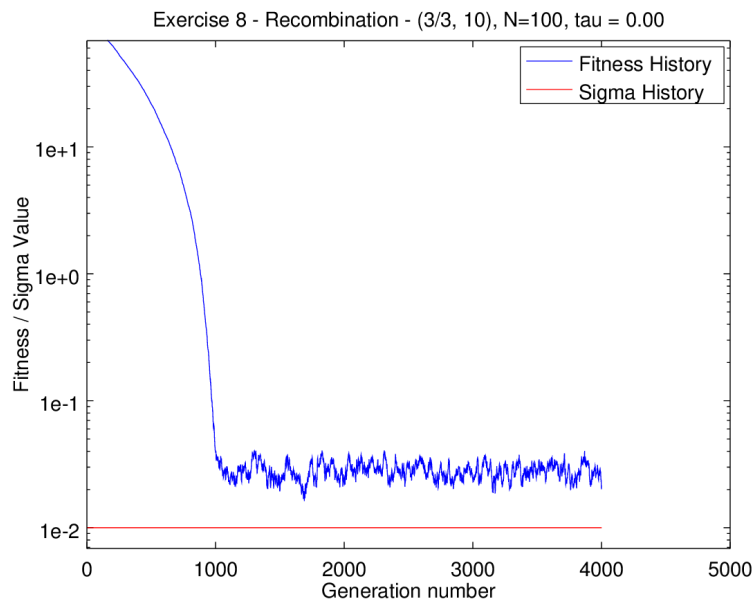


Abbildung 6: Aufgabe 8 - Rekombination - (3/3, 10), N=100, $\tau = 0.00$

1.3 Vergleich (1+1)-ES mit 1/5 Regel

Es wird die $(3/3, 10)$ - σ -SA-ES mit der (1+1)-ES mit 1/5 Regel am Kugelmodell mit $N = 100$ verglichen.

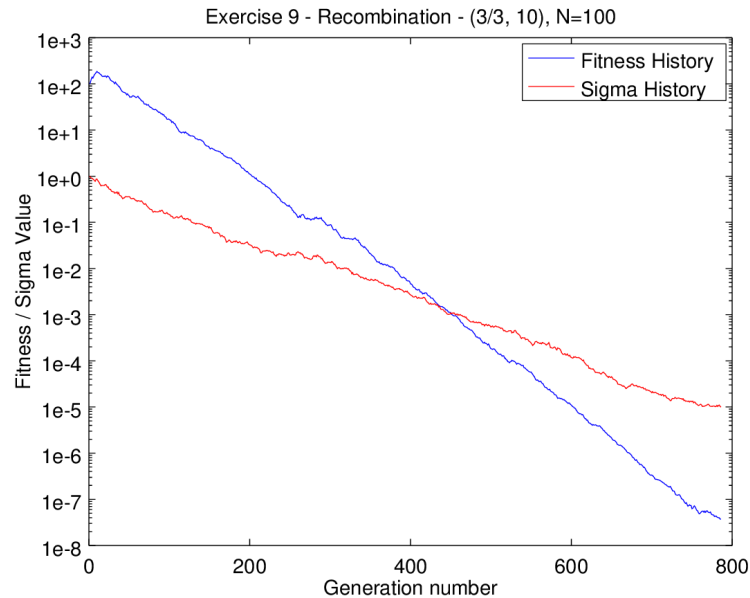


Abbildung 7: Aufgabe 9 - Rekombination - $(3/3, 10)$, $N=100$

- Fitness: 3.7111×10^{-8}
- Generationen: 785
- Funktionswertberechnungen: 7851

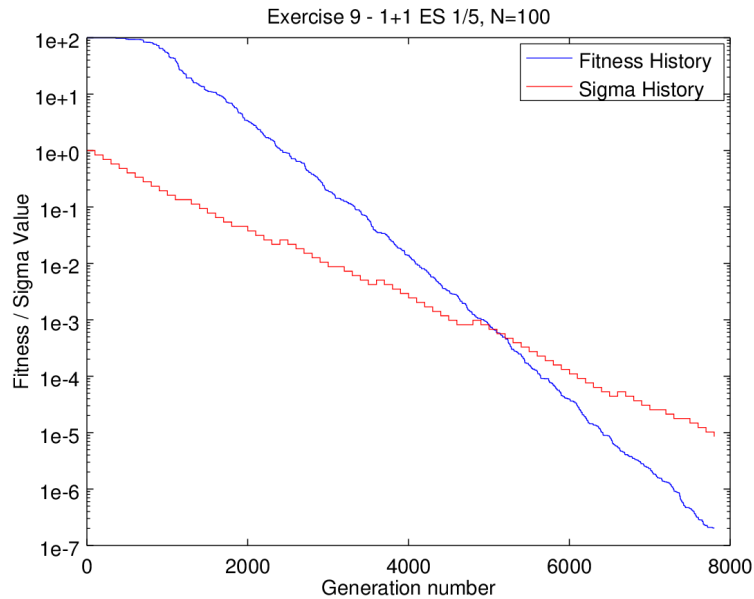


Abbildung 8: Aufgabe 9 - 1+1 ES 1/5, N=100

- Fitness: 2.0220e-07
- Generationen: 7800
- Funktionswertberechnungen: 7801

Aus den Abbildungen 7 und 8 ist ersichtlich, dass die (1+1)-ES mit 1/5 Regel in diesem Fall rund zehnmal so viele Generationen benötigt und das Ergebnis nicht ganz so nah am Optimum liegt, wie das der (3/3, 10)- σ -SA-ES. Die Anzahl der Funktionswertberechnungen ist allerdings ganz ähnlich, sogar etwas höher bei der (3/3, 10)- σ -SA-ES. Allgemein kann gesagt werden, dass die (3/3, 10)- σ -SA-ES am Kugelmodell effizienter ist, da mit annähernd der selben Anzahl an Funktionswertberechnungen ein besseres Optimum gefunden wurde.

1.4 Verrauschtes Kugelmodell

Es wird die Performance der $(\mu/\mu_I, \lambda)$ - σ -SA-ES am verrauschten Kugelmodell mit $N=30$, $\sigma=1$, $y=(10, \dots, 10)^T$ über 3.000 Generationen untersucht. Das Optimum ist durch das Rauschen bei jeder Funktionswertberechnung anders und kann somit nie genau erreicht werden. Bei unendlich vielen Ausführungen

liegt das Optimum aber im Schnitt bei 0, da das Rauschen um 0 normalverteilt ist. Das bedeutet auch, dass die Fitnesswerte im Schnitt dem Restzielabstand entsprechen.

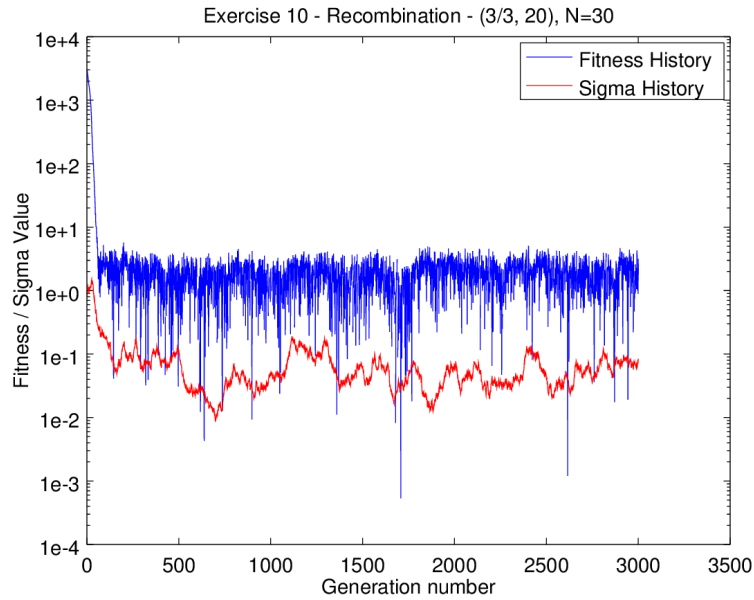


Abbildung 9: Aufgabe 10 - Performance der $(\mu/\mu_I, \lambda)$ - σ -SA-ES am verrauschten Kugelmodell

Initialisierung an der Stelle des Optimums und $\sigma = 10^{-5}$. Der durchschnittliche Restzielabstand bei konstantem μ/λ beträgt:

- $\mu=2, \lambda=20$: 1.4907
- $\mu=5, \lambda=50$: 0.33833
- $\mu=10, \lambda=100$: 0.32637

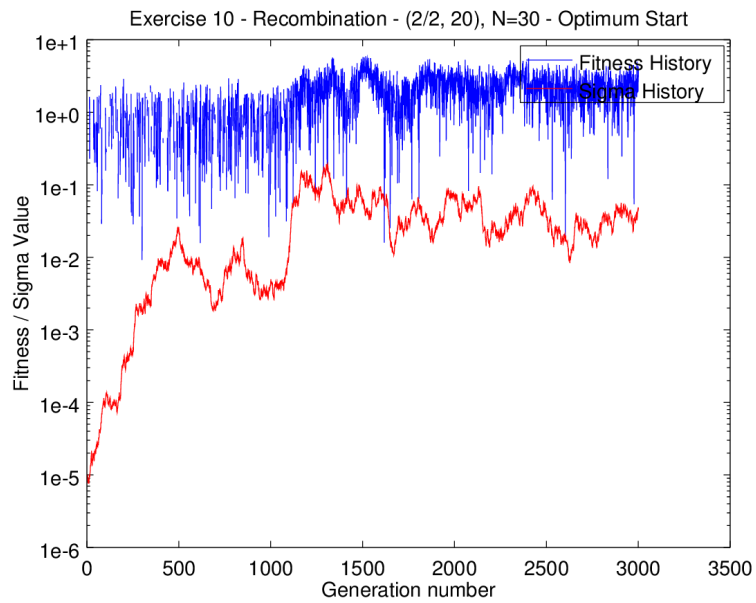


Abbildung 10: Aufgabe 10 - Start beim Optimum mit $(2/2, 20)$ und $N=30$

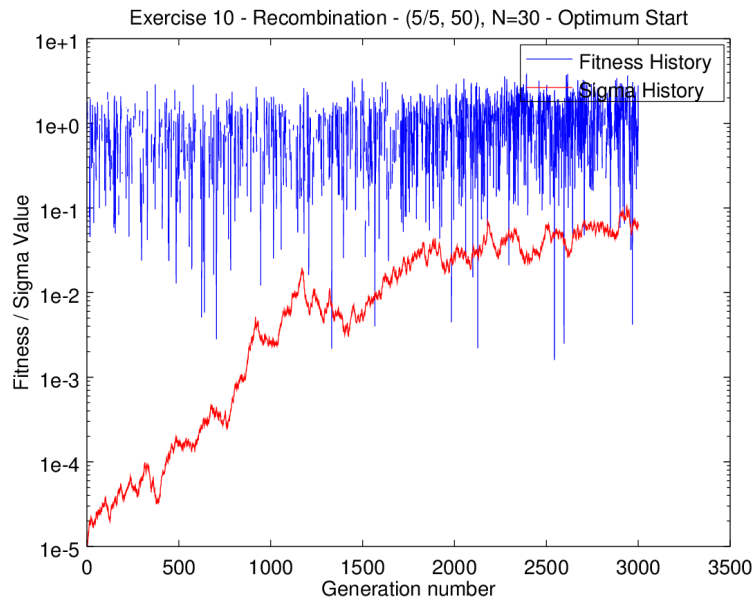


Abbildung 11: Aufgabe 10 - Start beim Optimum mit $(5/5, 50)$ und $N=30$

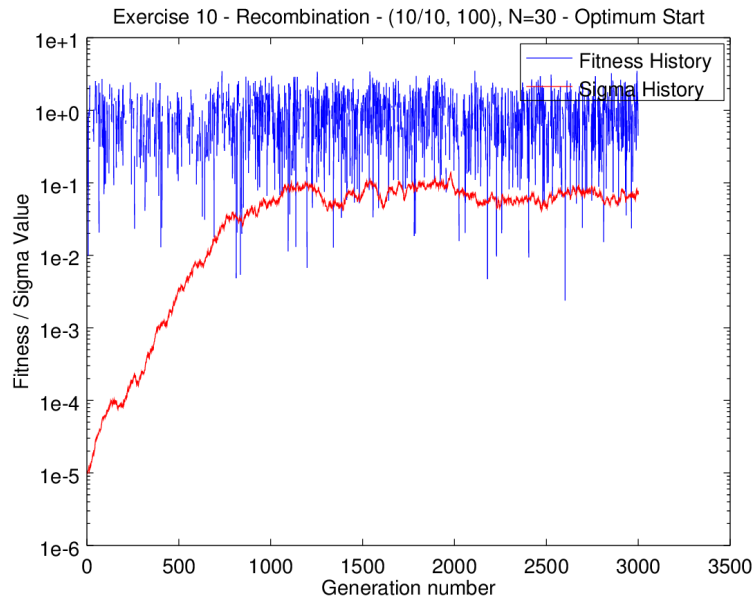


Abbildung 12: Aufgabe 10 - Start beim Optimum mit (10/10, 100) und N=30

In Abbildung 10, 11 und 12 ist zu erkennen, dass die Fitnesswerte aufgrund des Rauschens stark schwanken. Bei zu kleinem μ schwankt der mittlere Restzielabstand stärker, ab einem gewissen μ wird dieser aber nicht mehr wesentlich beeinflusst.

1.5 RASTRIGIN-Funktion

Es wird die Dynamik der $(\mu/\mu_I, \lambda)$ - σ -SA-ES an der RASTRIGIN-Funktion untersucht, mit $N = 30$, $\sigma = 1$ und $y = (10, \dots, 10)^T$. Abbildung 13 zeigt das Verhalten bei $\mu = 3$ und $\sigma = 9$.

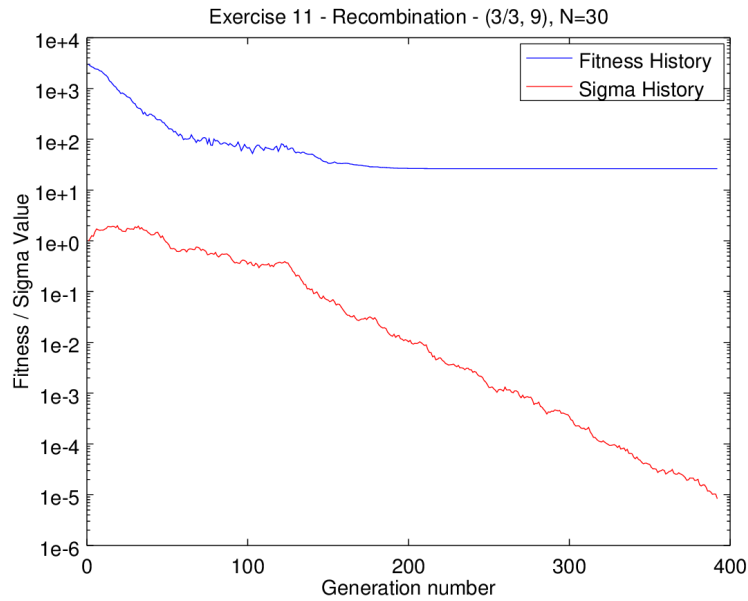


Abbildung 13: Aufgabe 11 - Rekombination - (3/3, 9), N=30

Es wird versucht, das globale Minimum, durch das schrittweise Erhöhen der Nachkommen, zu finden. Begonnen wird mit $\mu = 10$ und wird in 10er-Schritten auf $\mu = 200$ erhöht. Dabei wird das Abschneideverhältnis (μ/λ) konstant bei $1/3$ gehalten. Abbildungen 14, 15, 16 und 17 zeigen die Entwicklungen bei $\mu = 10$, $\mu = 50$, $\mu = 100$ und $\mu = 200$. Tabelle 1 zeigt die Ergebnisse der Fitness und Anzahl Generationen.

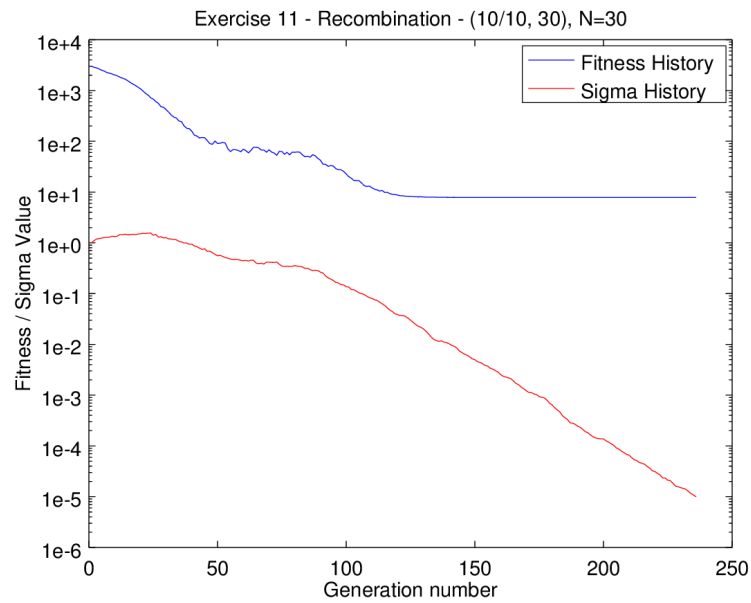


Abbildung 14: Aufgabe 11 - Rekombination - (10/10, 30), N=30

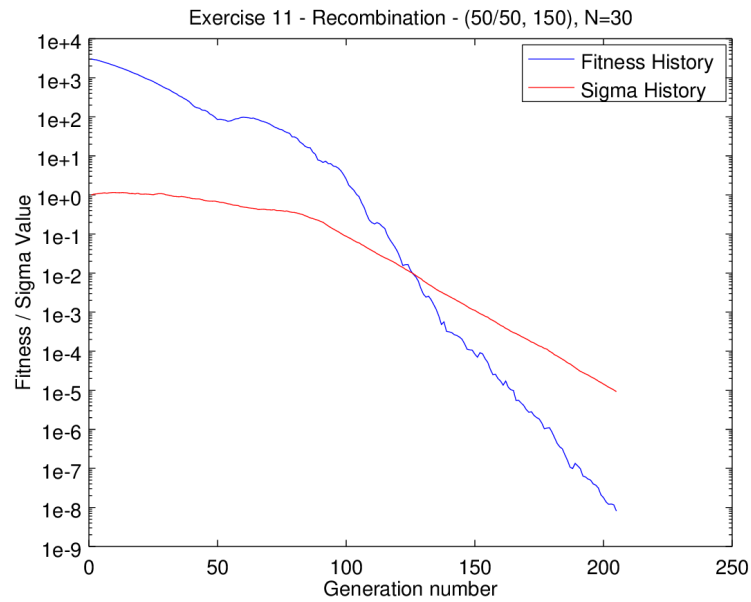


Abbildung 15: Aufgabe 11 - Rekombination - (50/50, 150), N=30

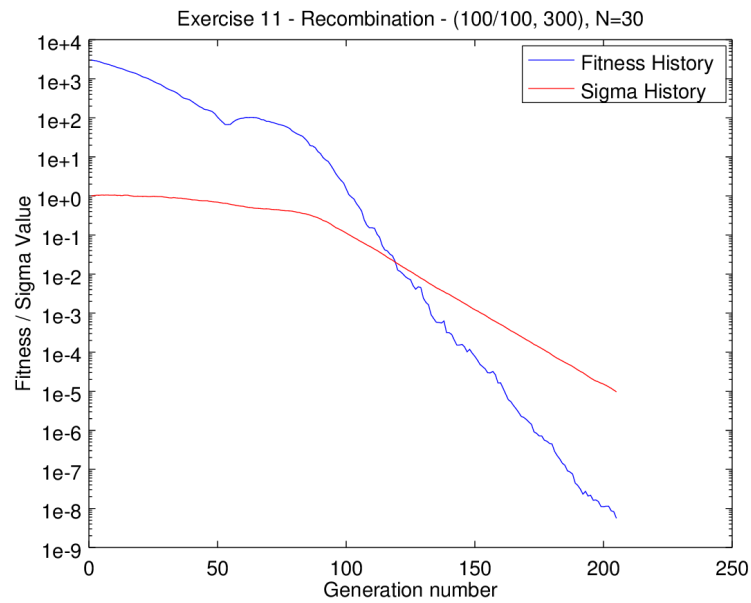


Abbildung 16: Aufgabe 11 - Rekombination - (100/100, 300), N=30

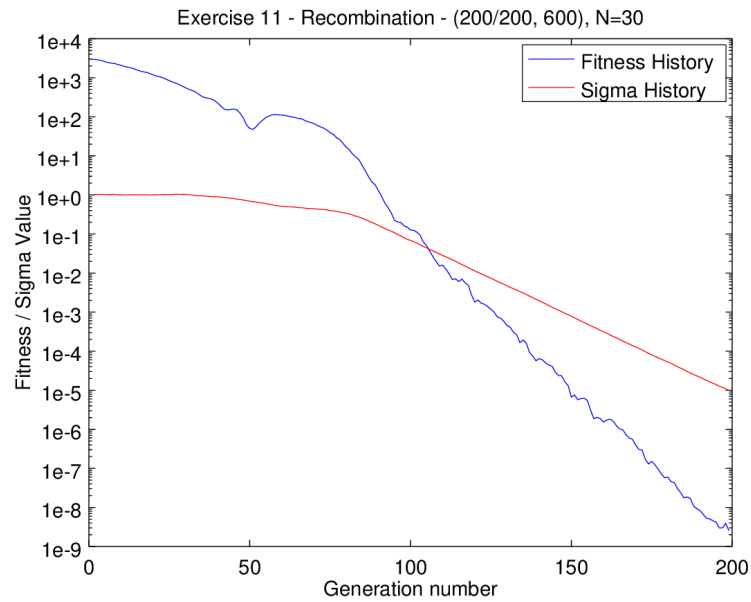


Abbildung 17: Aufgabe 11 - Rekombination - (200/200, 600), N=30

μ	Fitness	Generationen
10	7.8020	235
20	3.9010	214
30	2.9257	208
40	1.9505	204
50	8.0821e-09	204
60	1.9505	204
70	2.9257	204
80	3.4133e-09	203
90	7.1891e-09	199
100	5.6069e-09	204
110	3.1615e-09	197
120	1.8090e-09	197
130	9.7525e-01	197
140	9.7525e-01	198
150	9.7525e-01	195
160	9.7525e-01	196
170	3.1750e-09	198
180	1.7615e-09	197
190	1.8327e-09	198
200	2.5158e-09	198

Tabelle 1: Experimente Schrittweise Erhöhung von μ

Das beste Ergebnis wurde bei $\mu = 180$ gefunden. Daher werden die Experimente für $\mu = 171$ bis $\mu = 190$ in 1er-Schritten wiederholt, wie in Tabelle 2 dargestellt. Bei $\mu = 186$ wird das Minimum 1.5744e-09 gefunden, was dem globalen Minimum von 0 schon sehr nahe kommt.

μ	Fitness	Generationen
171	2.0741e-09	198
172	2.9246e-09	198
173	2.8466e-09	197
174	9.7525e-01	202
175	2.5521e-09	196
176	1.8860e-09	197
177	2.4774e-09	196
178	2.5307e-09	197
179	9.7525e-01	200
180	1.7615e-09	197
181	2.0189e-09	198
182	2.2926e-09	197
183	2.0091e-09	196
184	1.6369e-09	197
185	1.6388e-09	198
186	1.5744e-09	197
187	1.9301e-09	198
188	1.7551e-09	198
189	1.8099e-09	198
190	1.8327e-09	198

Tabelle 2: Experimente Schrittweise Erhöhung von μ

2. CMSA-ES Linsenoptimierung

Um die Auswirkung der Einbindung der Kovarianzmatrix zu zeigen, wird die Optimierung eines Linsenmodells zum einen mit der CMSA-Evolutionsstrategie durchgeführt und zum anderen mit der $(\mu/\mu_I, \lambda) - \sigma$ SA-Evolutionsstrategie. Bei diesem Vergleich zeigt sich, dass die CMSA-ES sich noch einmal von der $(\mu/\mu_I, \lambda) - \sigma$ SA-ES absetzen kann (vgl. Abbildung 18).

Da es eine derartige Verbesserung im Optimierungsverlauf gibt, deutet es darauf hin, dass der hierbei untersuchte Lösungsraum nicht "wohl-geformt" ist und die Einbindung der Kovarianzmatrix also einen Vorteil bringt.

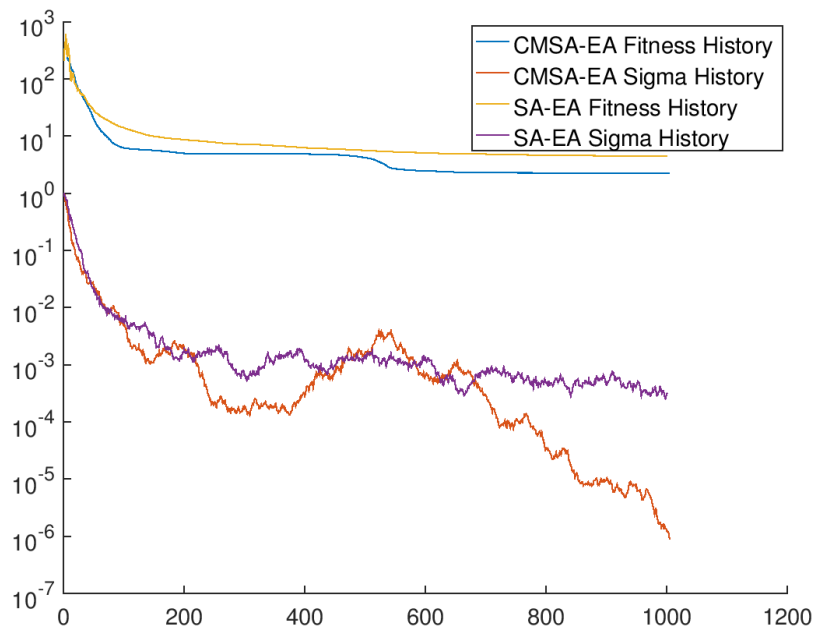


Abbildung 18: Aufgabe 12 - Vergleich der CMSA und SA Evolutionsstrategien am Linsenmodell