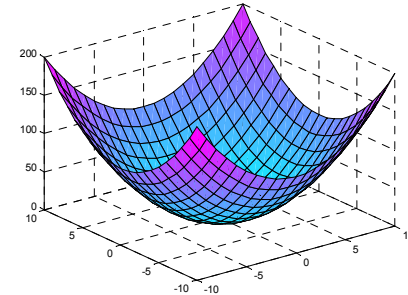


Evolutionsstrategie / Genetischer Algorithmus

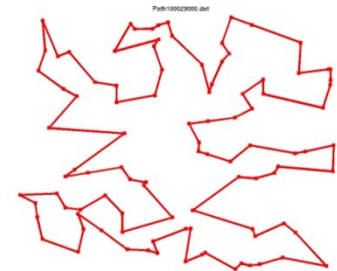
1. Evolutionsstrategien

(kontinuierliche Parameter)



2. Genetische Algorithmen

(diskrete Parameter in kombinatorischen Problemen)



Dr. Markus Olhofer
markus.olhofer(at)honda-ri.de

Evolutionäre Algorithmen

Algorithmus:

t = 0
Initialisiere
Do

- **Reproduktion**
- **Rekombination**
- **Mutation**

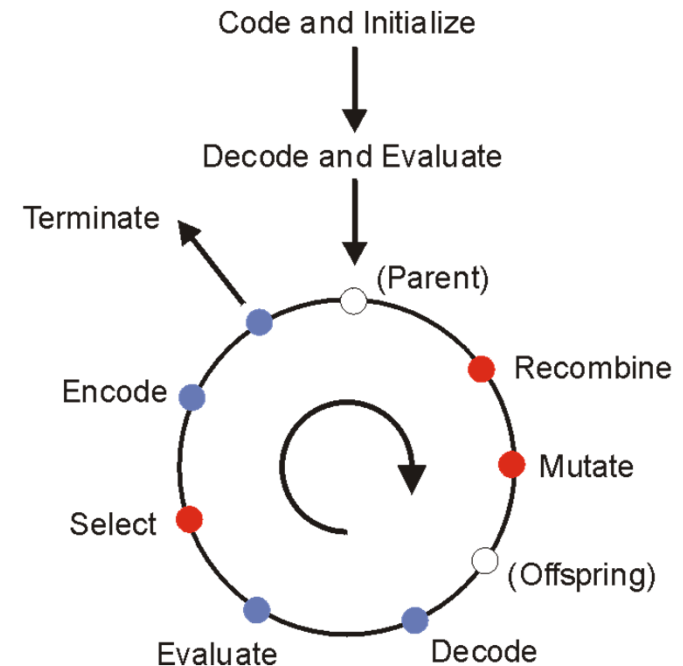
Dekodierung
Evaluierung

- **Selektion**

t := t + 1
Until stop

Evolutionäre Algorithmen

- Evolutionsstrategien
- Genetische Algorithmen
- Genetische Programmierung
- Evolutionäre Programmierung

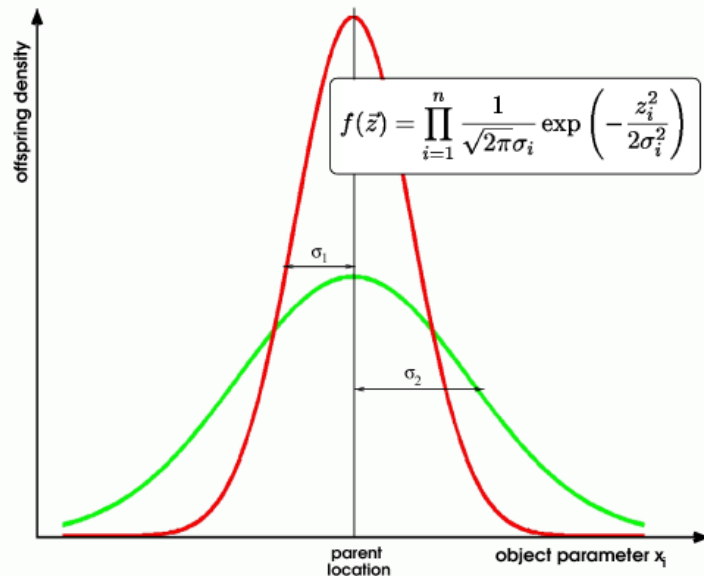


Evolutionsstrategie (ES): Hintergrund

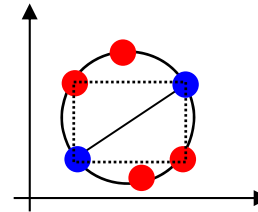
Makroskopisches Modell

Fokussierung auf die Phänotypebene: Verhalten ist Gegenstand der Selektion

- Komponenten des Lösungsvektors sind phänotypische Verhaltensmerkmale
- Annahme: Wie auch immer der genetische Unterbau realisiert ist, die resultierende phänotypische Änderung ist Gauß-verteilt
- Mutation ist der primäre Variationsoperator



- $c = \text{id}_{\mathbb{R}}$
- $g_1(t) \in \mathbb{R}^n$ - $x(t) := g_1(t)$ (Objektvariable)
- r = kontinuierliche Rekombination, z.B.

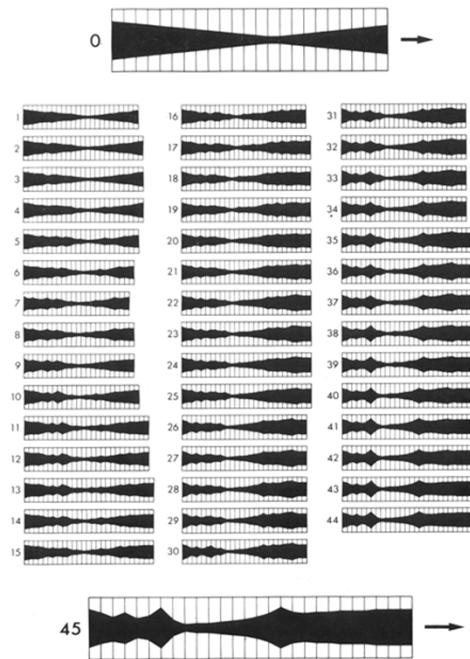


- s = deterministische Selektion
(μ, λ) und ($\mu + \lambda$) Selektion
- m = Mutation durch Addition normalverteilter Zufallszahlen
- Erweiterung mit multi-Rekombination und begrenzte Lebenszeit: ($\mu, \lambda, \kappa, \delta$)
 $\mu = \delta \rightarrow$ „center of mass“ Rekombination

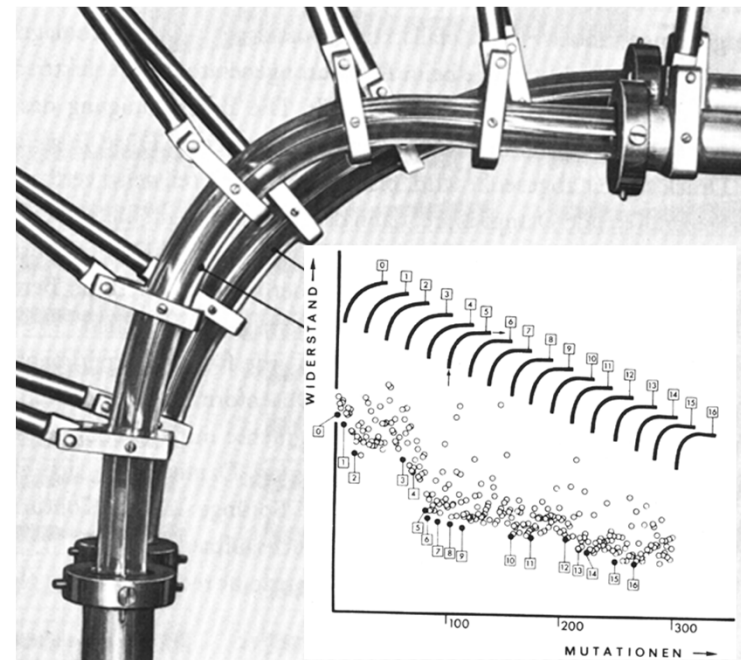
Evolutionsstrategie: Historie (Rechenberg, Bieri, Schwefel, 1965)

Evolutionäre Optimierung von fluid-mechanischem Design

- zunächst ein Elter und einen Nachkommen (1+1) Strategie
 - nur Mutation als Veränderung, anfangs binomial verteilt später normal verteilte Zufallszahlen
 - erste theoretische Analyse des Optimierungsprozesses für einfache Qualitätslandschaften
Korridormodell und Kugelmodell
- Rechenbergs 1/5 Regel: der Quotient aus erfolgreichen und erfolglosen Mutanten sollte 1:5 sein

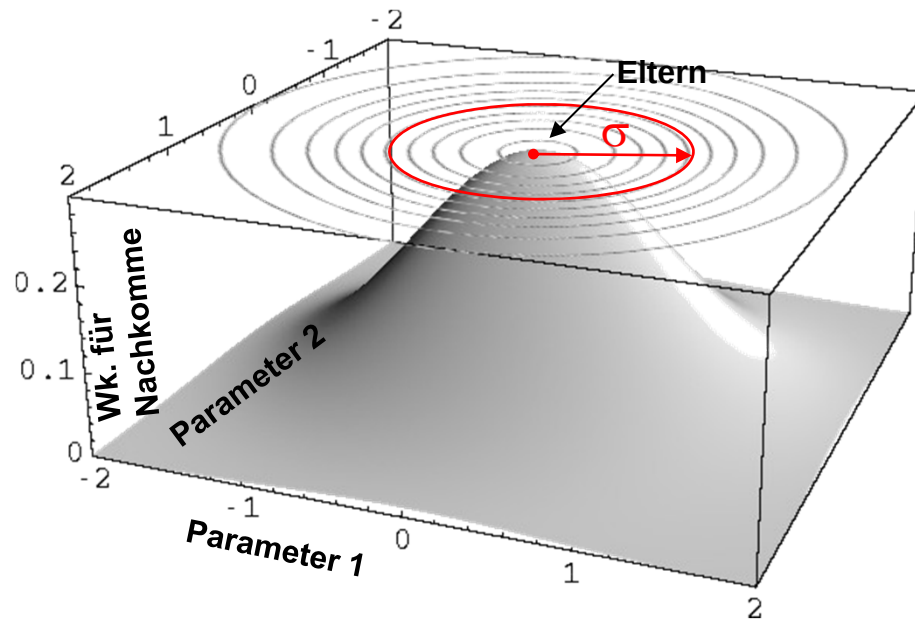


Evolutionäre Optimierung einer 2-Phasen
Überschalldüse (Schwefel et al. 1968)

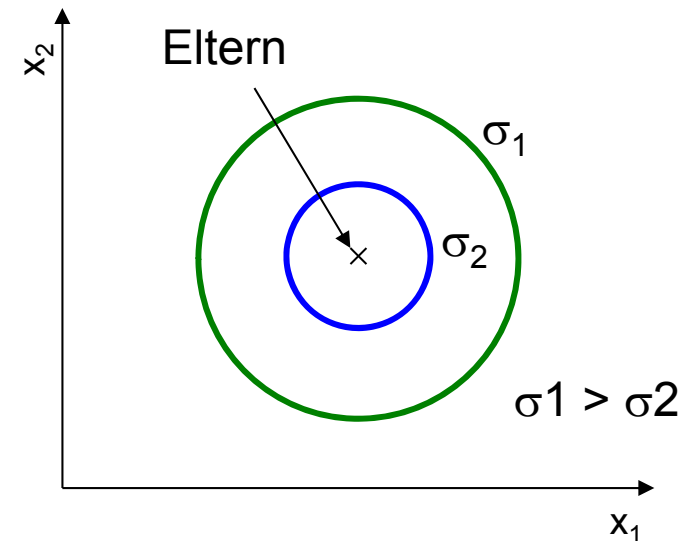


Evolutionäre Optimierung einer 90° Grad
Schlauchbiegung (Rechenberg et al. 1966)

Mutationsweite

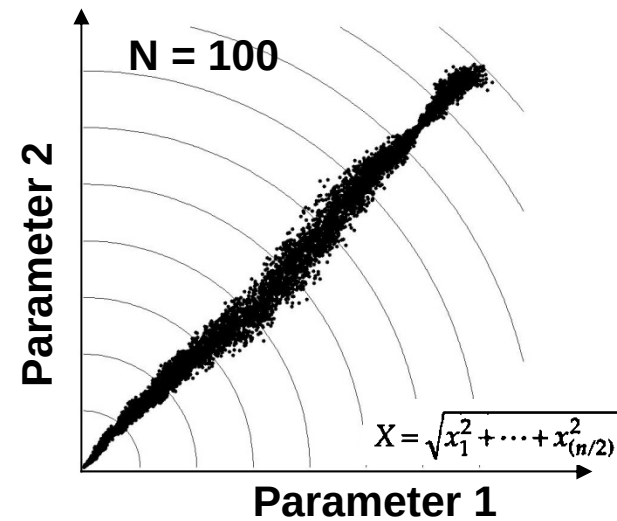
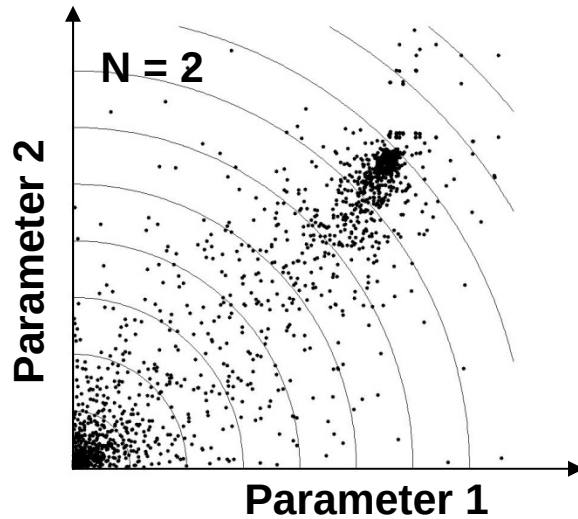


Wahrscheinlichkeit für die Erzeugung eines Nachkommens

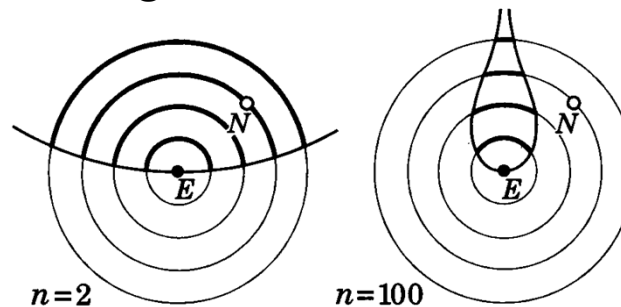


- die Varianzen σ^2 (Strategieparameter) beeinflussen die Mutationstärke (Schrittweite), d.h. die Veränderung die zu den Objektparametern addiert wird
- die optimale Größe der Varianzen ist von der lokalen Topologie des Suchraumes und vom Zustand der Population abhängig

Schrittweitenanpassung



Erfolgswahrscheinlichkeit



aus: Ingo Rechenberg, *Evolutionstrategie '94*, Frommann-Holzboog, Stuttgart 1994 / ISBN 3-7728-1642-8

Anpassung der Mutationsstärke

Anpassung der Schrittweite:

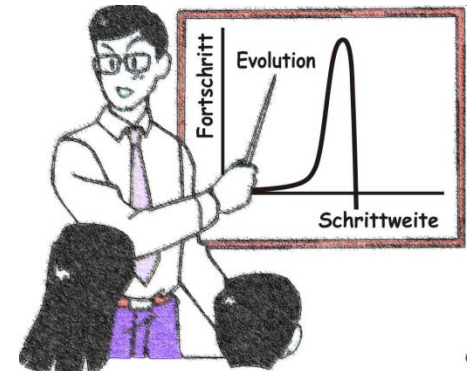
- Optimale Mutationsweite -> Maximaler Fortschritt
- Anpassung der Mutationsstärke ein Optimierungsproblem

1/5 - Erfolgsregel

- basiert auf der Analyse der *Fortschritts*geschwindigkeit für das Korridormodell und Kugelmodell.
- maximaler Fortschritt, wenn 1/5 der Nachkommen besser sind

Algorithmus

- a) bestimmen wie viele Nachkommen besser als die Eltern sind
- b) vergrößern / verkleinern der Schrittweite je nach Ergebnis



Selbstadaptation der Mutationsweite

- **mutative Selbstadaptation**

die Mutationsweiten σ_i^2 werden selbst Gegenstand der evolutionären Optimierung

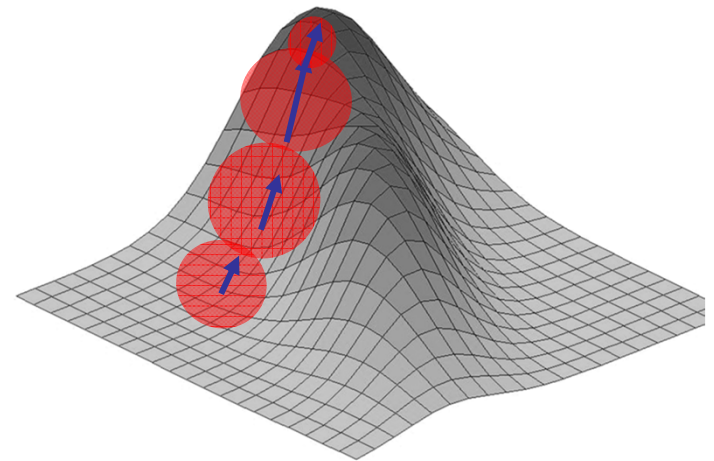
Mutationsoperator:

$$x_i(t+1) = x_i(t) + N(0, \sigma)$$

Nachkomme Eltern zufällige Mutation

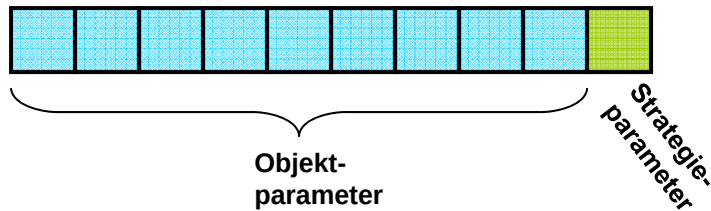
Schrittweite

„Waren die Objektparameter gut, so war auch die Schrittweite gut die zu ihrer Erzeugung genutzt wurde! - (im Mittel)“

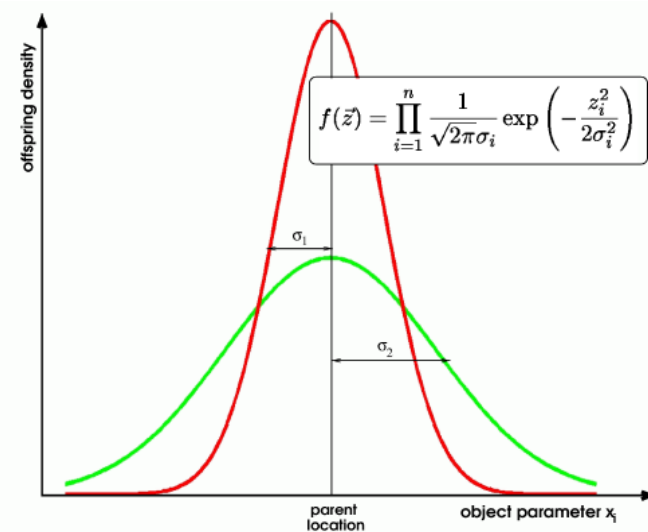


mutative Selbstadaptation

Chromosom:



Idee: Für die Adaptation des Strategieparameters wird der gleiche Prozeß genutzt wie bei der Optimierung der Objektparameter x_i .



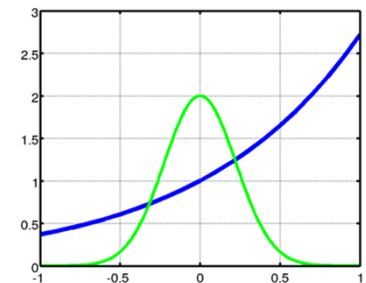
Anpassung der Schrittweite

$$\sigma(t+1) = \sigma(t) \cdot \exp(\xi) \quad \text{mit} \quad \xi \sim N(0, \sigma_\xi)$$

Anpassung der Objektparameter

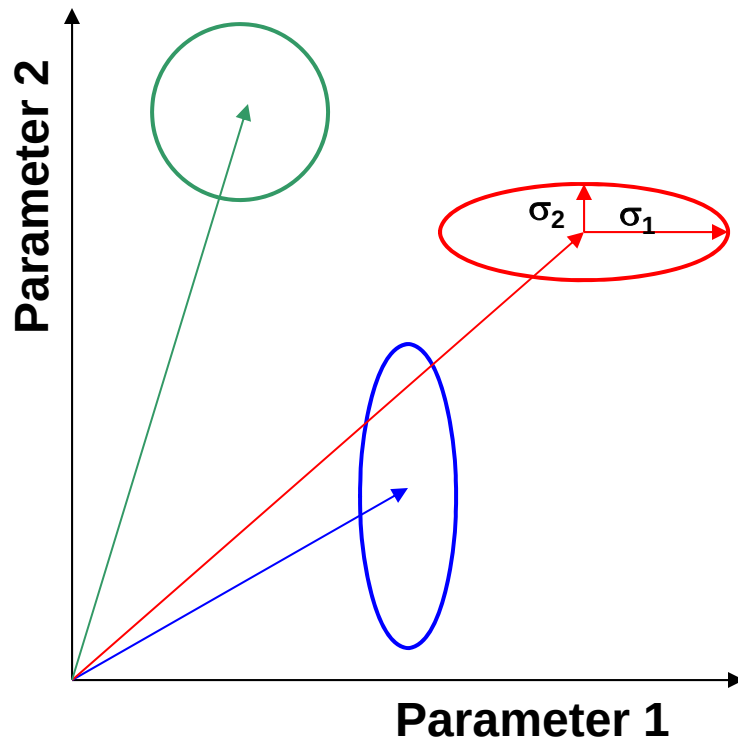
$$x_i(t+1) = x_i(t) + \sigma(t+1)z_i$$

$$z_i \sim N(0, 1) \quad \text{und} \quad i = 1, \dots, n$$



Individuelle Schrittweiten

Jede Variable hat eine eigene Schrittweite



Chromosome:

x_i									
σ_i									

$$\sigma_i(t) = \sigma_i(t-1) \exp(\tau' z) \exp(\tau z_i)$$

$$x(t) = x(t-1) + \vec{z}$$

$$z_i, z \sim N(0,1)$$

$$\vec{z} \sim N(\mathbf{0}, \vec{\sigma}(t)^2).$$

Standard Werte:

$$\tau' = \frac{1}{\sqrt{2n}}$$

globale
Anpassung

$$\tau = \frac{1}{\sqrt{2\sqrt{n}}}$$

lokale
Anpassung

mutative Schrittweitensteuerung

- pseudoCode

```
t = 0
Initialisiere Objektparameter der Eltern Population
Initialisiere Strategieparameter der Eltern Population
do // alle Generationen
    für alle Nachkommen
        wähle Eltern
        Rekombination Objektparameter
        Rekombination Strategieparameter
        Mutation der Objektparameter
        Mutation der Strategieparameter
        Evaluieren
    end
    Selektion (neue Elternpopulation)
    t := t + 1
Until stop
```

Source

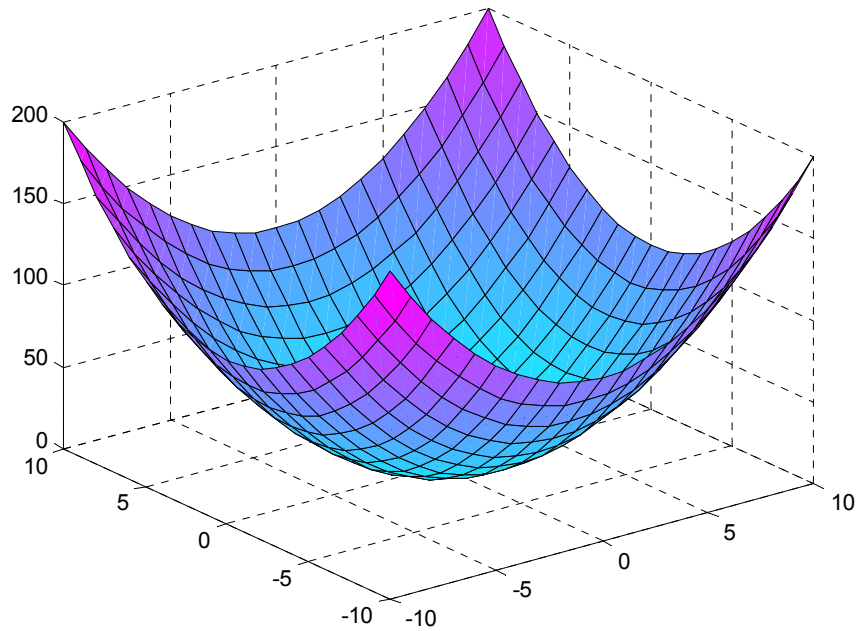
```
#include "Population.h"

using namespace std;

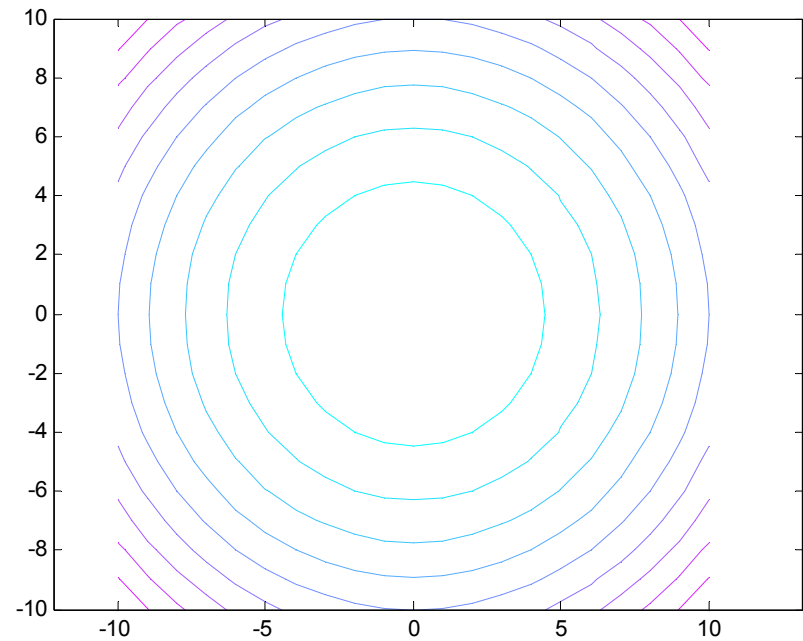
double sphere(const vector <double> &v)
{
    double sum = 0.;
    for (unsigned i = 0; i < v.size(); i++)
        sum += v[i] * v[i];
    return sum;
}

//=====
//
// main program
//
int main( int argc, char **argv )
{
    //
    // constants
    //
```

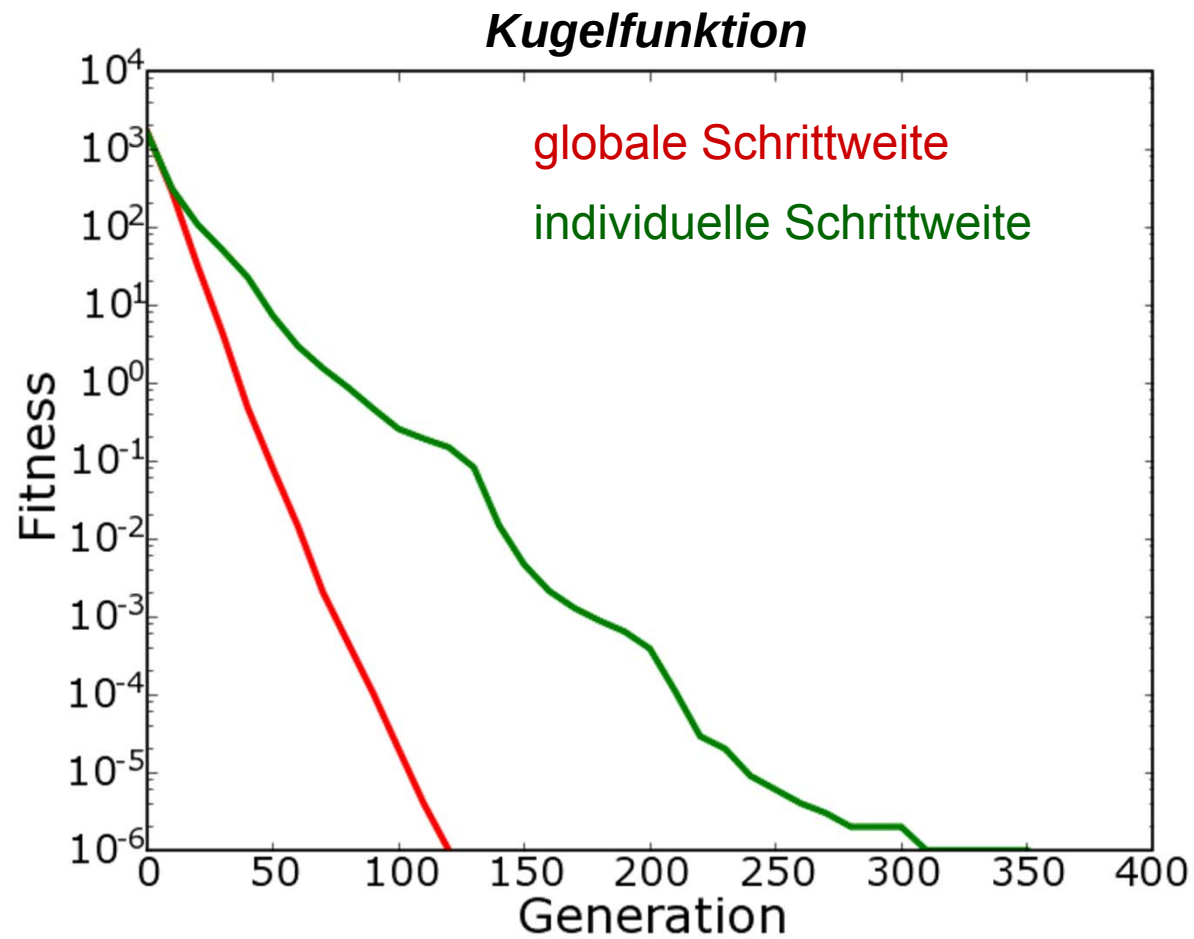
Kugelmodell



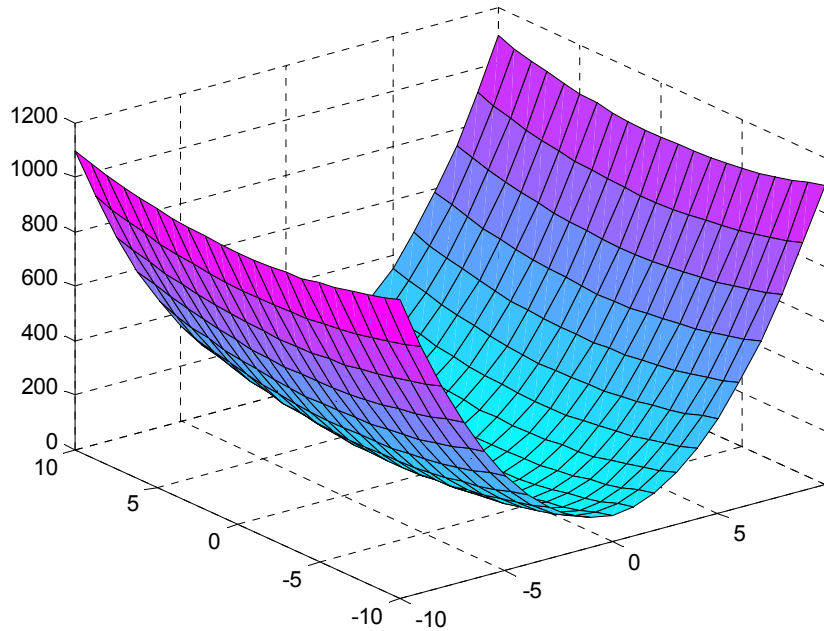
$$Q = \sum_{i=1}^N x_i^2$$



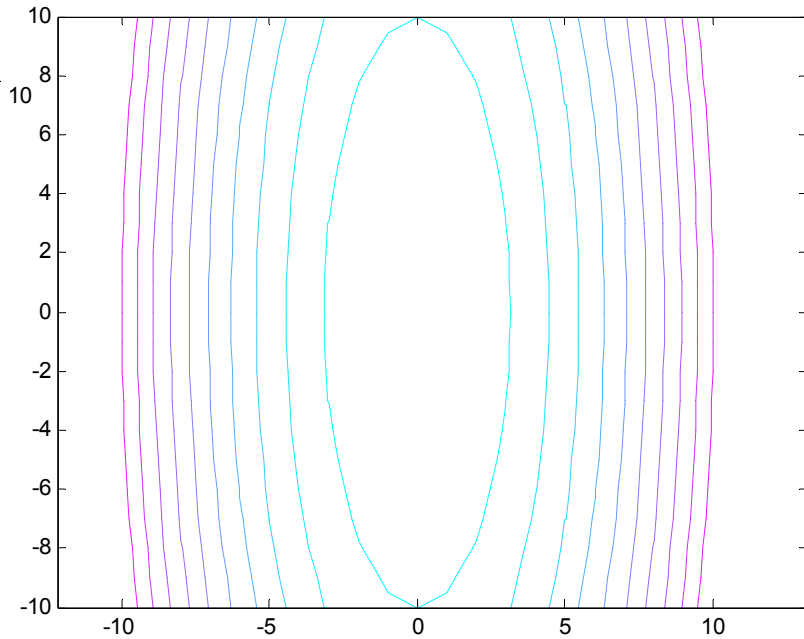
Kugelmodell



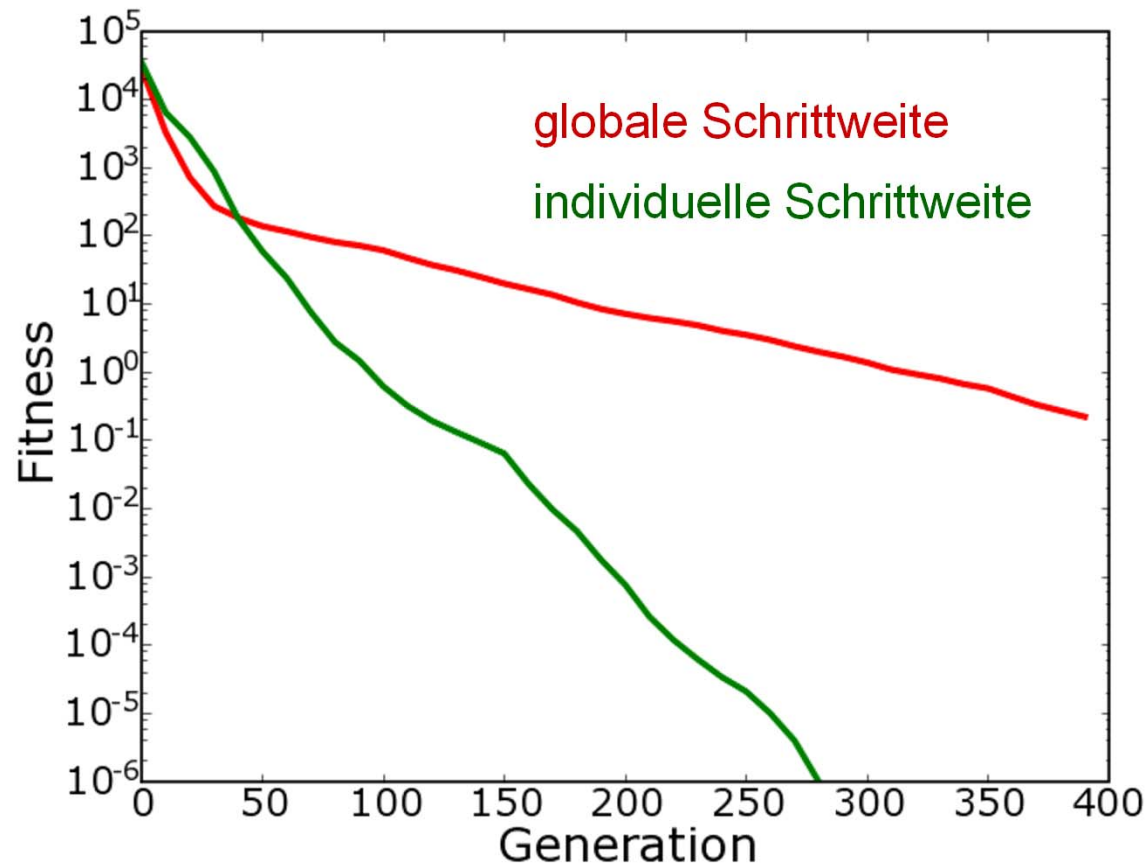
skalierte Kugelmodell



$$Q = \sum_{i=1}^N i x_i^2$$



skaliertes Kugelmodell

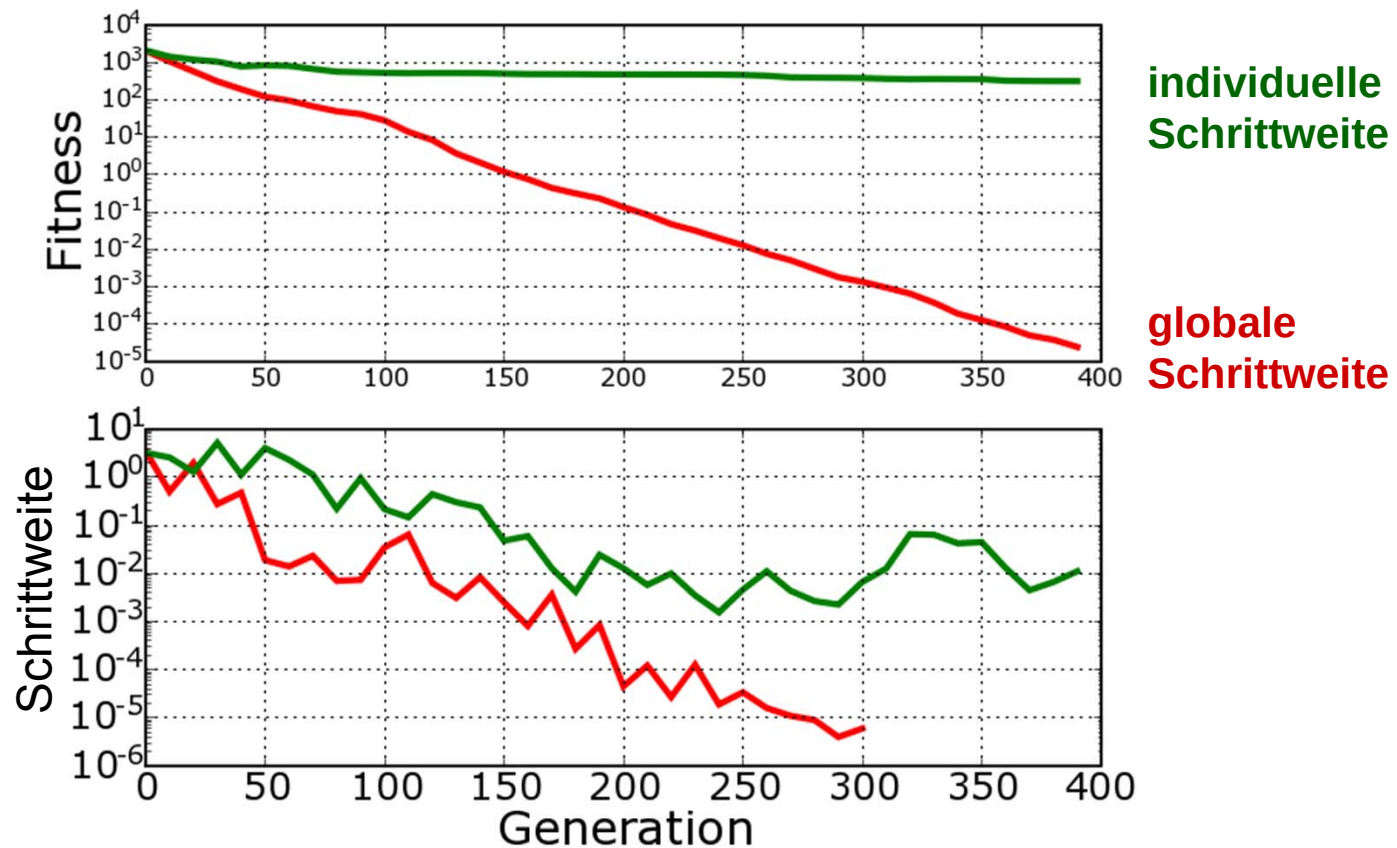


schnelle „grobe“ Anpassung bei globaler Schrittweite

- besser in den ersten Generationen
- schlechter in der Nähe des Optimums

Beispiel – individuelle Schrittweiten

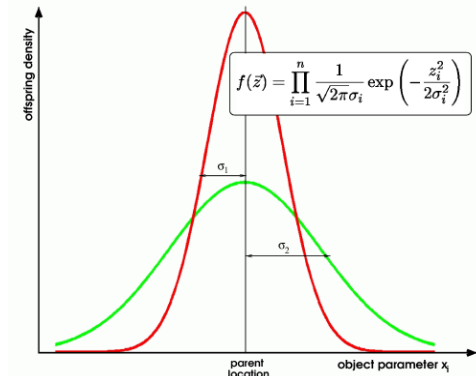
kleine Populationen ES(1,10)



Probleme

Prinzip der mutativen Schrittweitenanpassung

Koppelung von Strategieparameter an das Individuum
Selektionsdruck zu optimaler Schrittweite



Bei großen Schrittweiten kann die *tatsächlich realisierte Schrittweite* sehr klein sein

Würde ein solches Individuum selektiert, wird eine große Schrittweite vererbt, obwohl eine Mutation durch eine kleine Schrittweite realisiert wurde.

Probleme durch stochastische Fluktuationen der Schrittweite verstärkt durch:

- kleine Populationen
- hohe Anzahl an Strategieparametern (individuelle mutative Anpassung)

Derandomisierte Verfahren

1. Anpassung der Schrittweite aufgrund der *realisierten* Schrittweite
2. Dämpfungsfaktor der Adaptation

$$\vec{\sigma}(t+1) = \vec{\sigma}(t) \exp \left(\underbrace{\frac{1}{d}}_{\text{Dämpfung}} \left(\underbrace{|| \vec{z}^{sel} ||}_{\text{realisierte Mutation des selektierten Individuums}} - \underbrace{E[|| N(\vec{0}, \mathbf{I}) ||]}_{\text{erwartete Stärke der Mutation}} \right) \right)$$

$$\vec{x}(t+1) = \vec{x}(t) + \vec{\sigma}(t) \vec{z}$$

$$\vec{z} = (z_1 \dots z_n) \quad \text{mit} \quad z_i \sim N(0, 1) \quad \text{für} \quad i = 1 \dots n$$

Kumulation der Schrittweiten

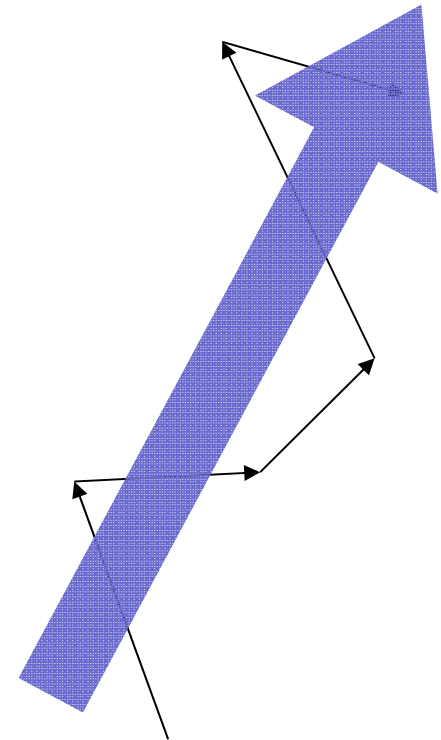
Nutzung der Informationen aus vorhergehenden Generationen

Pfad der Evolution:

$$\vec{s}(t) = (1 - c)\vec{s}(t - 1) + c_u \vec{z}^{sel}(t)$$

Anpassung der Schrittweite:

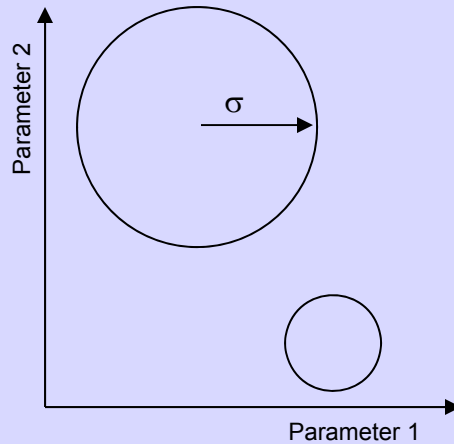
$$\vec{\sigma}(t) = \vec{\sigma}(t - 1) \cdot \exp\left(\frac{|\vec{s}(t)| - \chi_n}{d \cdot \chi_n}\right)$$



Evolutionsspfad

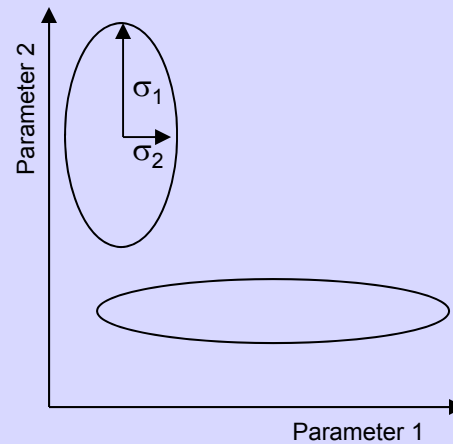
Richtung der Mutation

globale Schrittweite



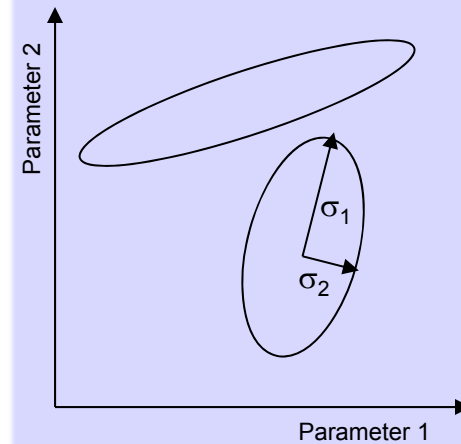
1 Parameter

individuelle Schrittweite



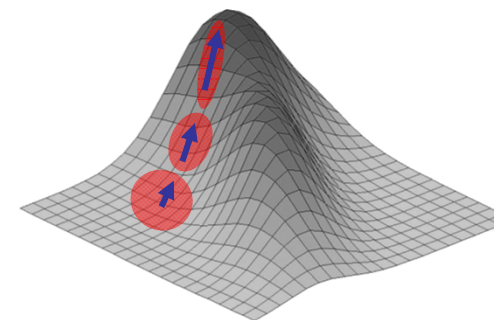
n Parameter

Kovarianzmatrix



$\sim n^2 / 2$ Parameter

Wahrscheinlichkeit für die Erzeugung eines Nachkommen

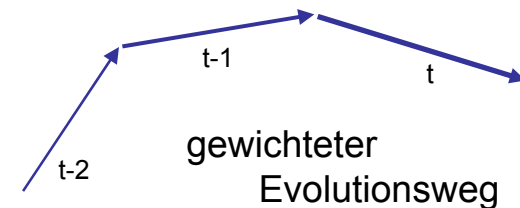


Selbstadaptation: „Derandomisierung“ und Kovarianzmatrix

- eine (de-randomisierte) Zufallszahl für die Anpassung der Objekt- und Strategieparameter; die *tatsächliche* Schrittweite im Parameterraum wird zur Anpassung der Strategieparameter genutzt

➡ wenn die Mutation stärker als erwartet $E(|N(0,1)|)$ aber erfolgreich war, wird die Varianz erhöht und umgekehrt

- *kumulative Schrittweitenadaptation*; der Evolutionsweg (Gedächtnis) wird bei der Adaptation genutzt
Analogie: Ensemble vs. Zeitmittel

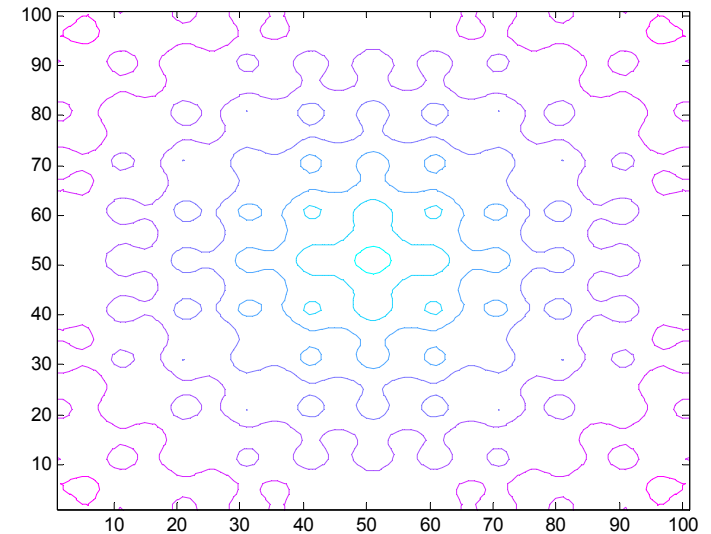
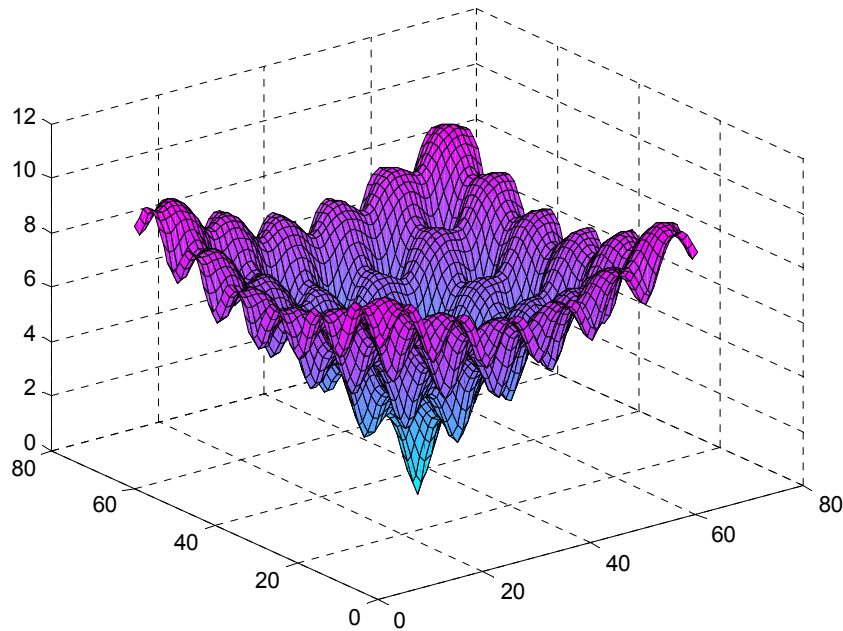


- die Kovarianzmatrix der gaußschen Wahrscheinlichkeitsdichtefunktion wird an die lokale Topologie des Suchraumes angepaßt

$$\vec{\delta} \sim \frac{\sqrt{\det(\Sigma^{-1})}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}(\vec{x} - \vec{\mu})^T \Sigma^{-1}(\vec{x} - \vec{\mu})\right)$$

da Σ^{-1} positiv definit sein muss mit $\det(\Sigma^{-1}) > 0$, sind die verschiedenen Matrixelemente nicht unabhängig und die detaillierte Vorgehensweise bei der Adaptation etwas komplizierter

Ackley's Testfunktion



$$f(x) = -20 \exp \left(-\frac{1}{5} \left(\frac{1}{n} \sum_{1 \leq i \leq n} x_i^2 \right)^{1/2} \right) - \exp \left(\frac{1}{n} \sum_{1 \leq i \leq n} \cos(2\pi x_i) \right)$$

Kovarianzmatrix-Adaptation

Eigenschaften:

- derandomisiert
- Dämpfung der Schrittweitenadaptation
- Anpassung der Kovarianzmatrix basierend auf dem Pfad der Evolution
($n^2 + n$)/2 freie Parameter
- zusätzl. globale Schrittweite

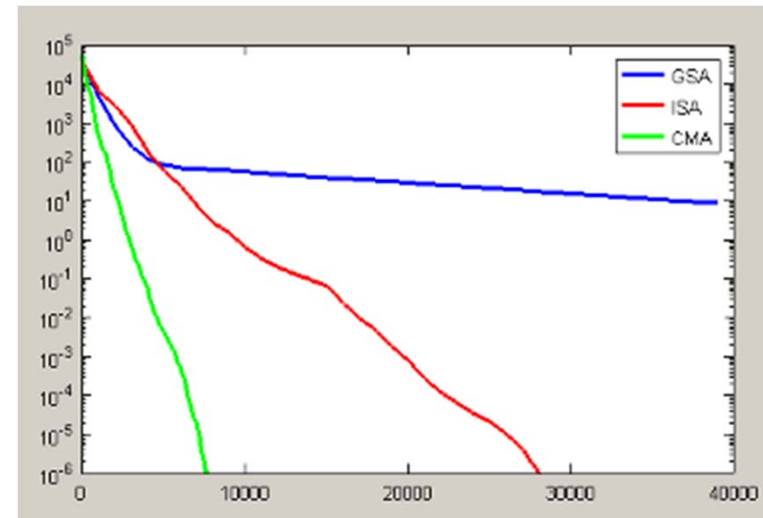
Mutation

1. Bestimmung der „selektierten“ Mutation
2. Berechnung des Pfades
3. Anpassung einer globalen Schrittweite
4. Anpassen der Kovarianzmatrix
5. Anpassen der Objektparameter

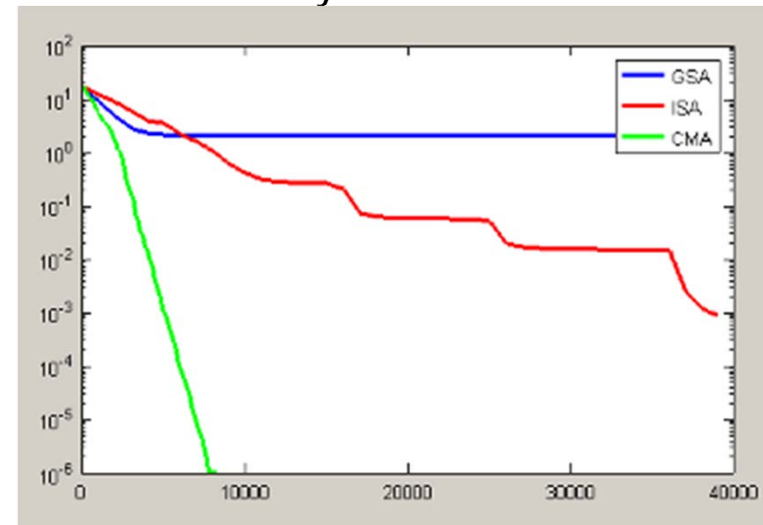
$$f(\vec{z}) = \frac{\sqrt{\det(\mathbf{C}^{-1})}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}(\vec{z}^T \mathbf{C}^{-1} \vec{z})\right)$$

korrelierte Mutationen

skaliertes Kugelmodell



Ackley Funktion



Zusammenfassung

- Evolutionsstrategien zur Optimierung kontinuierlicher Parameter
- Schrittweitenanpassung
 - globale mutative Schrittweitenanpassung
 - einfacher Algorithmus zur Anpassung einer Schrittweite
 - individuelle mutative Schrittweitenanpassung
 - Probleme bei mutativer Anpassung durch hohe Anzahl von Parametern (große Populationen nötig)
- CMA Strategie (Covarianz Matrix Adaptation)
 - „Derandomisierte“ Anpassung der Schrittweiten
 - Kummulation der Richtungen
 - korrelierte Mutationen

Teil II

Genetic Algorithm

Genetische Algorithmen (GA): Hintergrund und Historie (Holland, 1967)

Mikroskopische Modell: Populationsgenetik

Fokussierung auf den Genotyp Level: Gene sind Gegenstand der Selektion

Schema Processing and the K-Armed Bandit

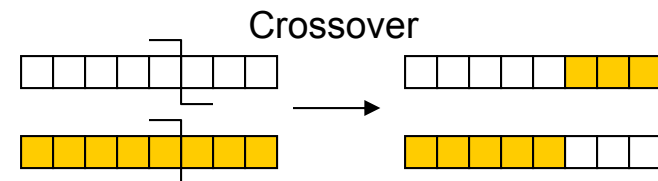
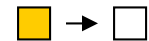
- Adaptation an eine gegebene Umwelt
- Adaptation ist eine Suchtrajektorie durch einen Zustandsraum mgl. Lösungen

Aufgabe: Beste Strategie für das Ziehen von n-Zufallsvariablen, die einen definierten aber unbekannten Mittelwert und eine unbekannte Varianz haben



Kanonischer Genetischer Algorithmus

- genetische Codierung (String, Alphabet, z.B. binär)
- Veränderungen auf dem Genotyplevel
 - Crossover und Inversion
 - Punkt-Mutation (sekundär)



- proportionale Selektion

die Wahrscheinlichkeit Nachkommen zu erzeugen ist proportional zur Fitness

Problem: Fitness muss immer positiv sein

$$p(x_i) = \frac{f(x_i)}{\sum_{i=1}^n f(x_i)}$$

Das *Travelling Salesman Problem*

Gesucht:

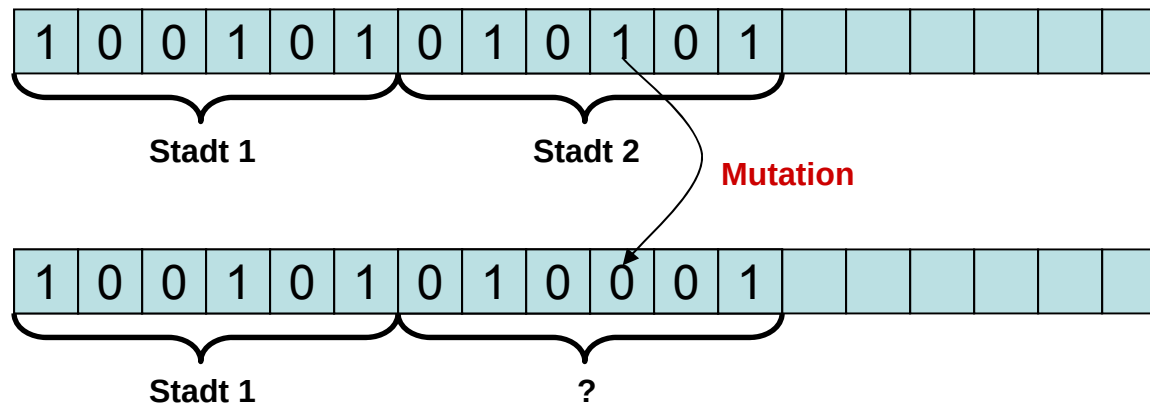
Route zwischen n Städten, die jede Stadt genau einmal besucht

- NP-vollständig
- unwahrscheinlich effiziente Algorithmen zur Lösung dieses Problems zu finden
- Der „primitive“ Algorithmus, der einfach alle Möglichkeiten durchprobiert hat eine Komplexität von $O(n!)$
- häufig „völlig ausreichend“ eine „gute“ Lösung zu finden

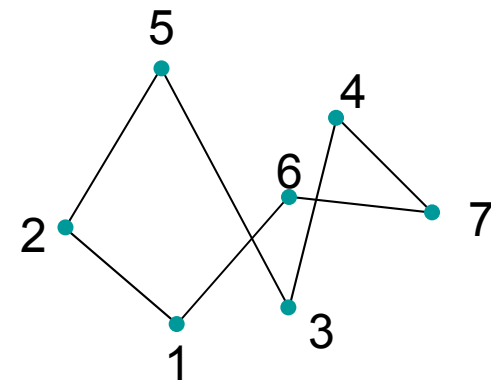


Repräsentation/Kodierung

Binärstrings:



- Bei Mutation und Rekombination entstehen ungültige Lösungen
- Stärke der Mutation abhängig vom der Position



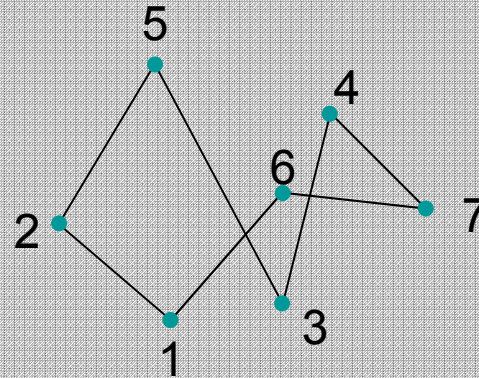
Pfad Repräsentation

Reihenfolge der Elemente in einer Liste wird als Weg interpretiert

Beispiel:

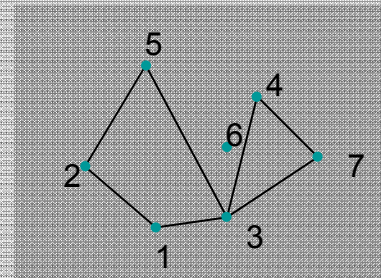
Genotyp: (3 4 7 6 1 2 5)

Tour 3-4-7-6-1-2-5



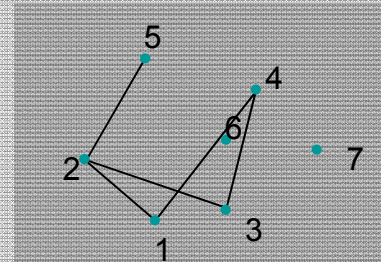
Mutation

(3 4 7 6 1 2 5) $\rightarrow$ (3 4 7 3 1 2 5)



Crossover

(3 4 7 6 | 1 2 5)
(1 2 3 4 | 5 6 7) $\rightarrow$ (3 4 7 6 | 5 6 7)
(1 2 3 4 | 1 2 5)



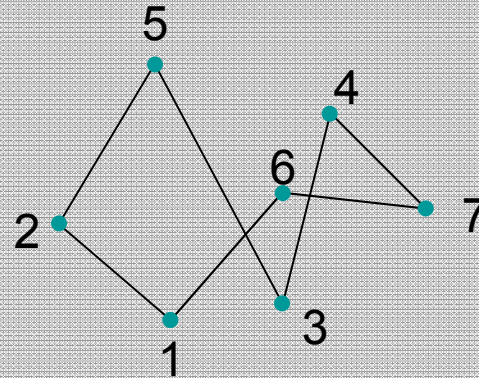
Pfad Repräsentation

Reihenfolge der Elemente in einer Liste wird als Weg interpretiert

Beispiel:

Genotyp: (3 4 7 6 1 2 5)

Tour 3-4-7-6-1-2-5



Problemangepaßte Operatoren:

Zufälliges Einfügen

Displacement / Insertion

(3 4 7 6 1 2 5) $\longrightarrow$ (3 4 6 1 2 **7** 5)

Zufälliges Vertauschen

Reciprocal Exchange

(3 4 7 6 1 2 5) $\longrightarrow$ (3 4 **2** 6 1 **7** 5)

Inversion

(3 4 |7 6 1| 2 5) $\longrightarrow$ (3 4 **1 6 7** 2 5)

Partially Matched Crossover (PMX)

Idee:

- Teiltour aus einem Elternteil übernommen
- versuchen die Reihenfolge und Position möglichst vieler Städte aus dem anderen Elternteil zu erhalten

Algorithmus

1. zwei zufällige Crossoverpunkte wählen
2. Werte zwischen den Punkten vertauschen
3. Städte links und rechts der Crossoverpunkte wieder an ihrer Originalposition einfügen, sofern sie die Gültigkeit der Darstellung nicht verletzen
4. Sollte dies aber der Fall sein: statt dessen den Wert einfügen, der vor der Crossover Operation an der Stelle stand, an der nun der gleiche Wert steht wie der, der eingetragen werden soll.

p1 = (1 4 |6 7 5 3| 2)

p2 = (5 2 |7 6 4 3| 1)

o1 = (x x |7 6 4 3| x)

o2 = (x x |6 7 5 3| x)

o1 = (1 x |7 6 4 3| 2)

o2 = (x 2 |6 7 5 3| 1)

o1 = (1 5 7 6 4 3 2)

o2 = (4 2 6 7 5 3 1)

Ordinal Representation

- Kodierung durch Liste natürlicher Zahlen
- Das Chromosom repräsentiert eine Tour unter Zunahme einer sortierten Referenzliste C

$C = (1\ 2\ 3\ 4\ 5\ 6\ 7)$

$(2\ 3\ 1\ 4\ 1\ 2\ 1)$ (Genotyp)  Tour 2-4-1-7-3-6-5.



Nehme das zweite Element
in C als ersten Wegpunkt und
lösche es aus C

(Die letzte Zahl ist immer 1, da C zu
diesem Zeitpunkt nur noch 1 Element
enthält.)

Ordinal Representation

Vorteil: klassische Crossover funktioniert!
keine ungültigen Routen

Beispiel für $n=7$:
 $C = (1\ 2\ 3\ 4\ 5\ 6\ 7)$

$p1 = (2\ 3\ 1\ \ 4\ 1\ 2\ 1)$ und	2-4-1-7-3-6-5
$p2 = (6\ 4\ 5\ \ 2\ 3\ 1\ 1)$	6-4-7-2-5-1-3
$o1 = (2\ 3\ 1\ 2\ 3\ 1\ 1)$	2-4-1-5-7-3-6
$o2 = (6\ 4\ 5\ 4\ 1\ 2\ 1)$	6-4-7-5-1-3-2

Nachteil: Teil-Weg linksseitig des Crossover Punktes
wird nicht verändert.
Die rechte Seite wird stark verändert.

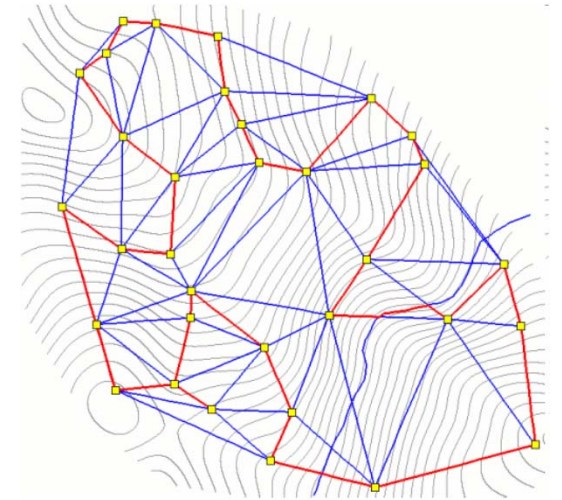
Mutation ist möglich, indem z.B. einzelne Werte im Rahmen des
für ihre Position gültigen Wertebereiches verändert werden.

Qualitätsfunktion

M_{ij} : Distanz von Stadt i zu Stadt j

$M_{ii} = 0$ für alle i

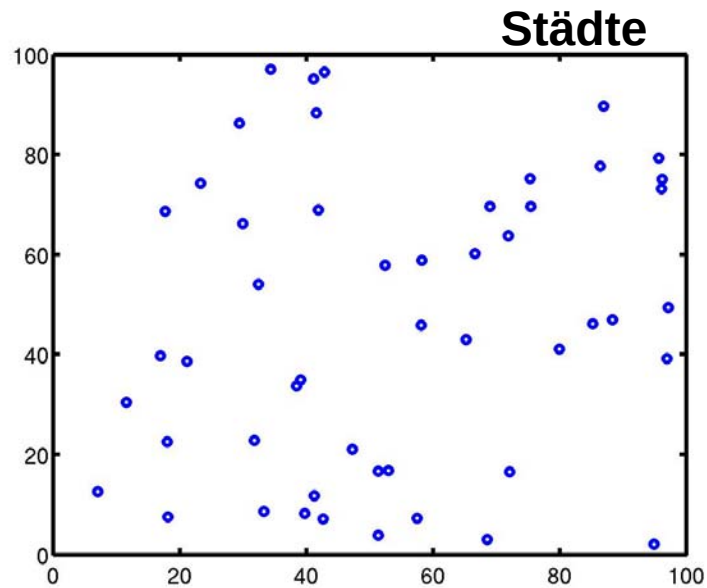
M symmetrisch wenn Weg nicht richtungsabhängig



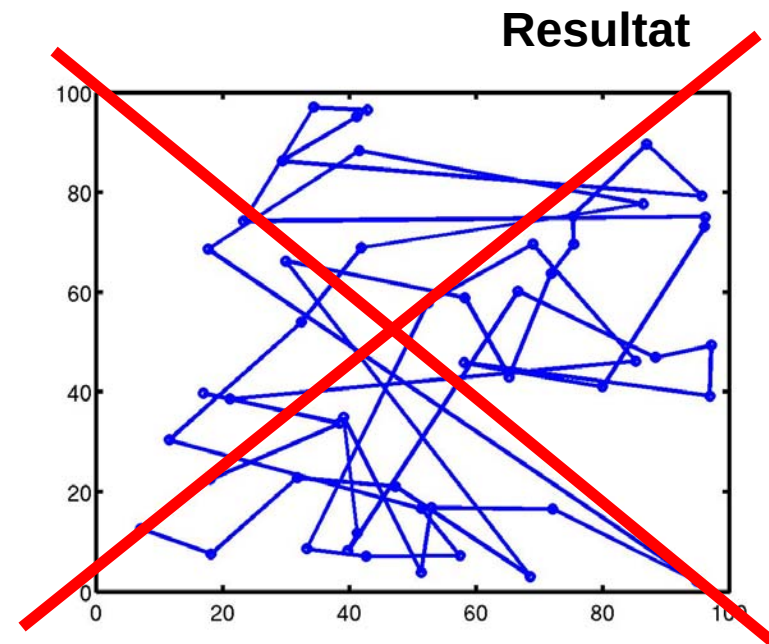
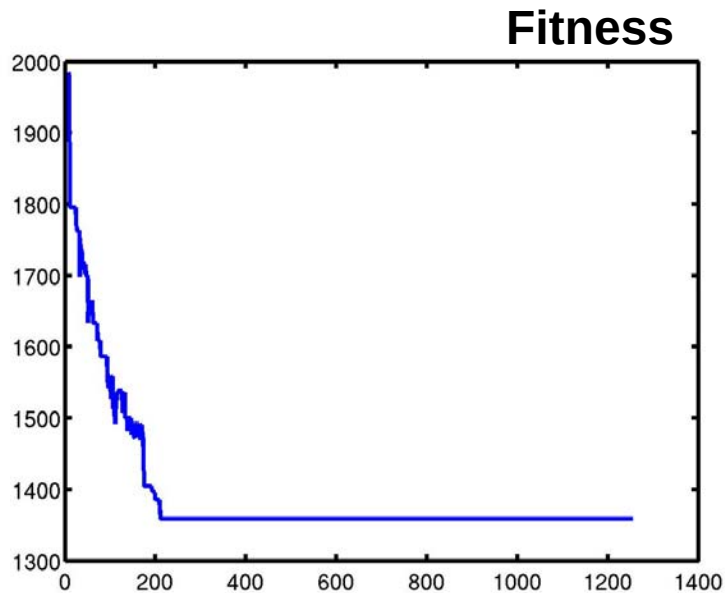
Qualitätsfunktion:

$$\phi(\pi) = \sum_{i=1}^{n-1} M_{\pi(i), \pi(i+1)} + M_{\pi(n), \pi(1)}$$

Ergebnis

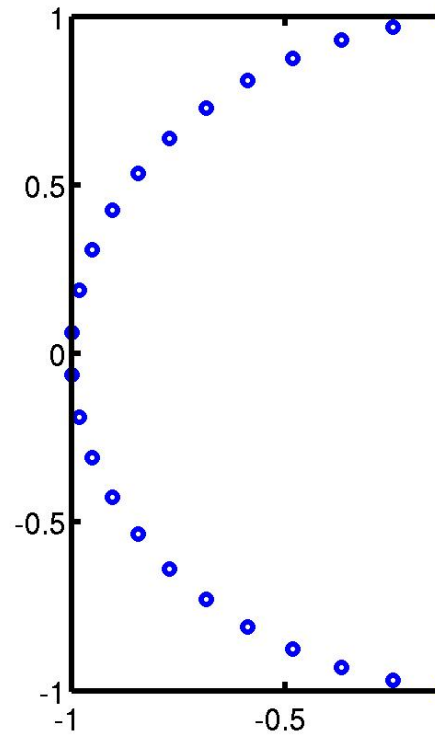


Population: 200- Individuen
Operator: PMX Crossover
Wk Crossover: 0.1

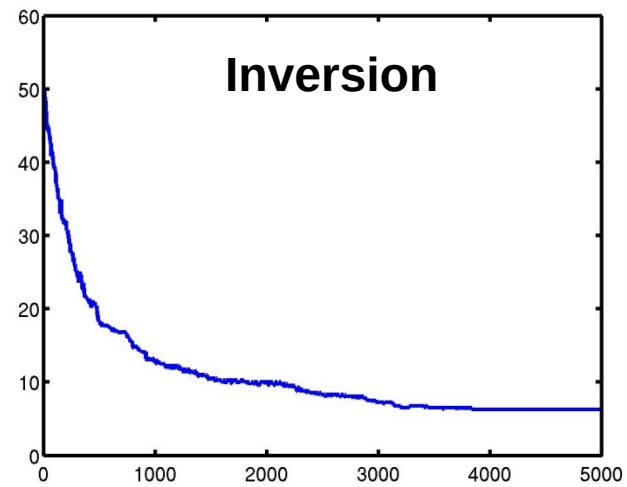


Testfunktion

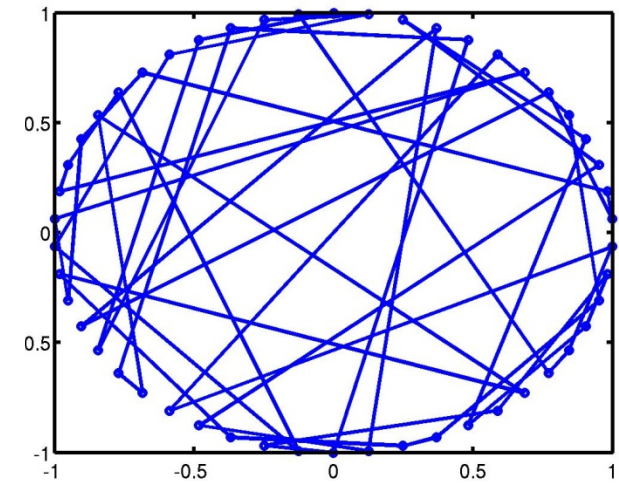
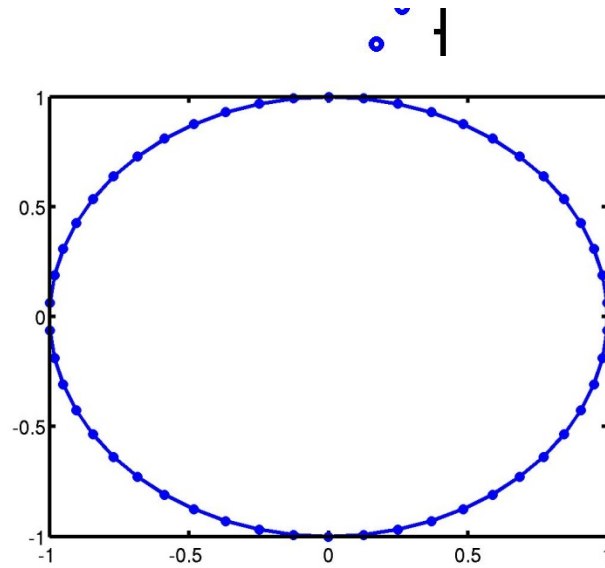
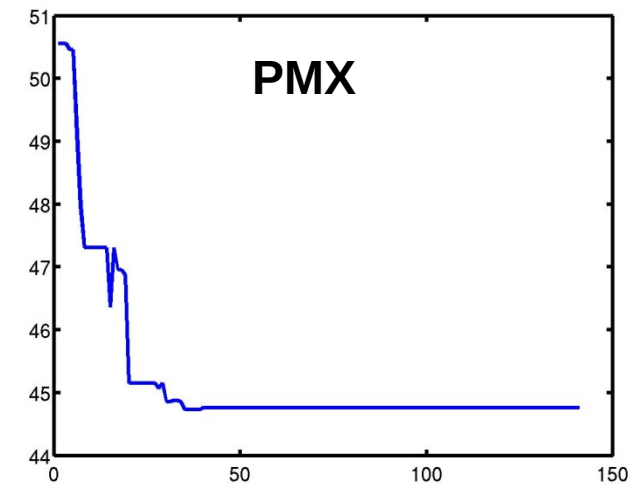
Städte



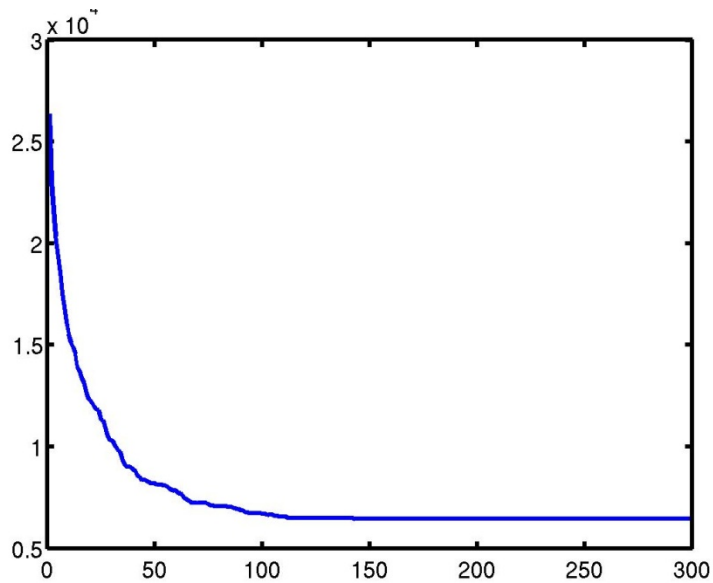
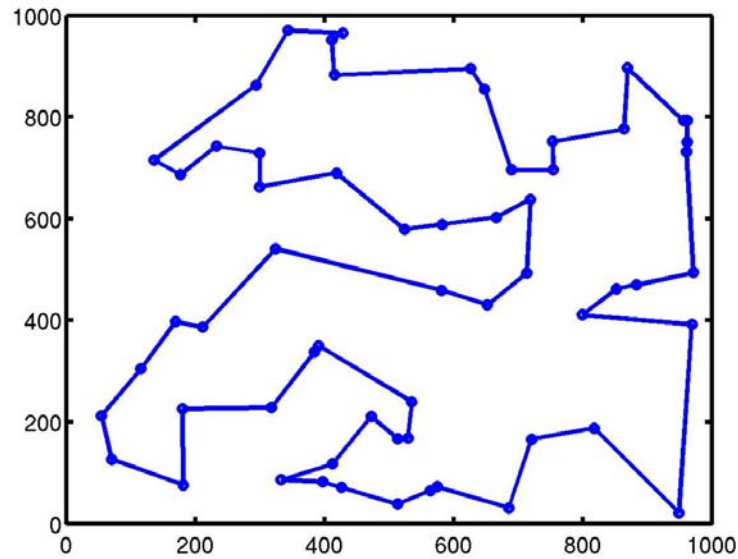
Inversion



PMX



Optimierung 60 Städte



- movie deleted

Zusammenfassung

- Problemangepaßte Kodierung
 - Binärstrings
 - diskrete Parameter
 - kontinuierliche Parameter
- Kodierung (Genotyp – Phänotyp Abbildung) und Operatoren
 - Erzeugen den Suchraum
 - Beeinflussen den Optimierungsprozess
- Operatoren
 - Cross-over: Erhalten von Informationen aus allen Eltern
 - Mutation: kleine Variationen um Informationen des Individuums zu erhalten
- Verbesserung der Lösung kann auch bei ungünstiger Kodierung erzielt werden – aber: Konvergenz zu suboptimalen Lösungen (lokale Optima)
 - Testfunktionen

Selektion I

Deterministische Selektion

Selektionsdruck $\sim \lambda / \mu$

- (μ, λ) Selektion: die Eltern werden nur aus den Nachkommen selektiert, lokale Minima können einfacher überwunden werden, gute Lösungen können vergessen werden
- $(\mu + \lambda)$ Selektion: Eltern der nächsten Generation werden aus der Nachkommen und der Elternpopulation (t-1) ausgewählt
- Elitist: der (die) besten Nachkommen werden behalten (mit bzw. ohne weitere Reproduktion)

Stochastische Selektion

Selektionsdruck ist über die *take-over time* definiert

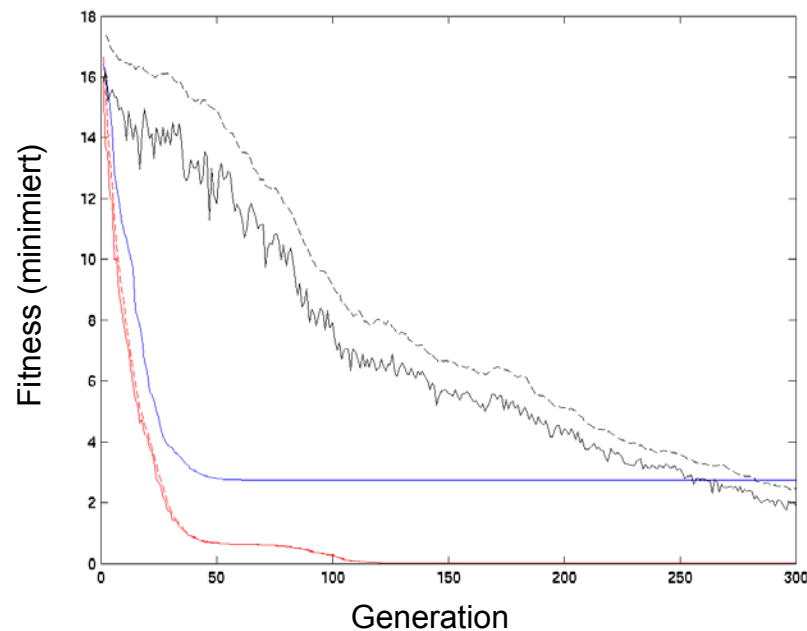
- Rouletterad Selektion (RS): Simulation eines gewichteten Roulettrades, Gewichtung ist eine Funktion der Fitneß der einzelnen Individuen
- proportionale Selektion:
Skalierung ist notwendig um einen hinreichenden Selektionsdruck aufrechtzuerhalten → RS
- Ranking: jedes Individuum erhält eine Selektionswahrscheinlichkeit die nur eine Funktion des Ranges des Individuums ist (Reihenfolge gemäß Fitneßwerte) → RS
- q -tournament Selektion: das beste Individuum aus einer zufällig gewählten Gruppe von q Individuen wird selektiert, der Prozeß wird n -mal wiederholt (n – Populationsgröße)
- EP-style tournament: alle Elter und Nachkommen werden mit einer zufälligen q -Gruppe verglichen. Alle gewonnenen Vergleiche werden aufsummiert, die Individuen mit dem höchsten Gewinn werden selektiert

Selektion II

Trade-off zwischen *Exploration* und *Exploitation*

- hoher Selektionsdruck (Exploitation):
 - + schnelle Qualitätssteigerung, fortschreiten entlang der *Gradientenlinie*
 - mangelnde Diversität in der Population → schwierig lokalen Minima zu entkommen
- niedriger Selektionsdruck (Exploration):
 - + hohe Diversität, globale Optimierung
 - Zerfließen der Information, Zufallssuche

Einfluss der Mutationstärke und der Selbstadaptation beachten



Evolutionsstrategie mit mutativer Selbstadaptation

- 30 dimensionale Ackley Funktion

$(\mu, \lambda) = (75, 100)$; Selektionsdruck ~ 1.3

$(\mu, \lambda) = (15, 100)$; Selektionsdruck $\sim 6,7$

$(\mu, \lambda) = (1, 100)$; Selektionsdruck ~ 100

— bestes Individuum
- - - Populationsmittel

Selektionsdruck

Wie stark wird selektiert?

μ - Größe der Elternpopulation

λ - Größe der Nachkommenpopulation

Notation:

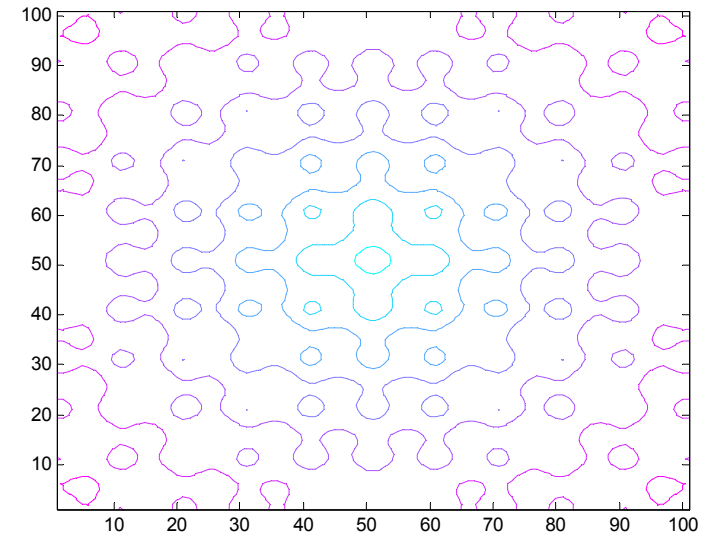
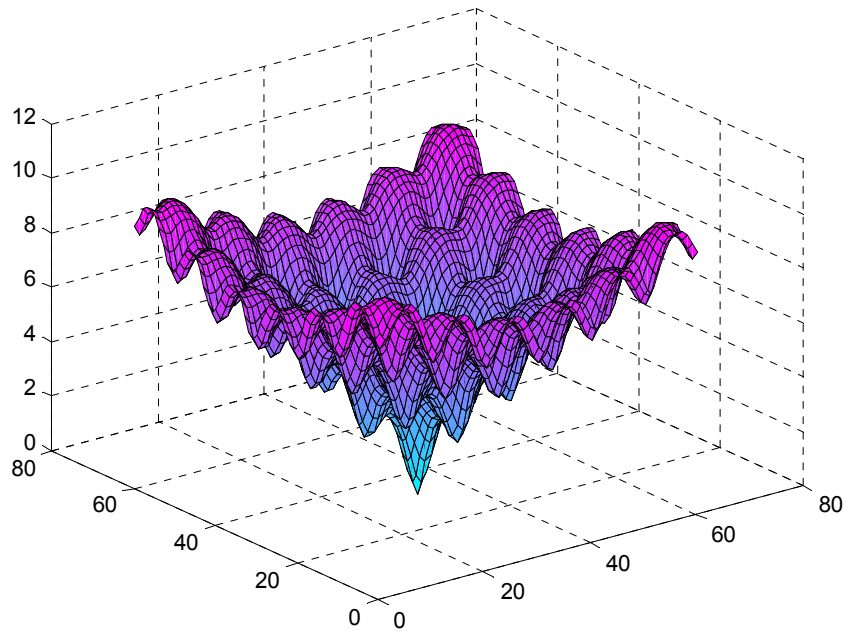
ES(10,100): $\mu = 10$, $\lambda = 100$

Selektionsdruck: $s = \mu / \lambda$

$s = 1$: keine Selektion -> kein Fortschritt
zufällige Suche bei der Kommastrategie

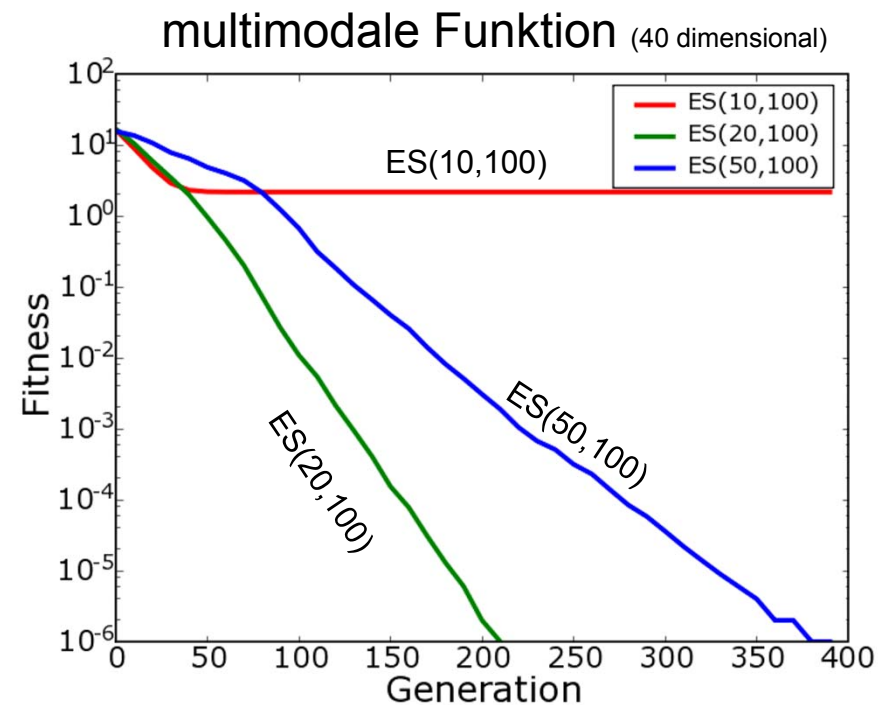
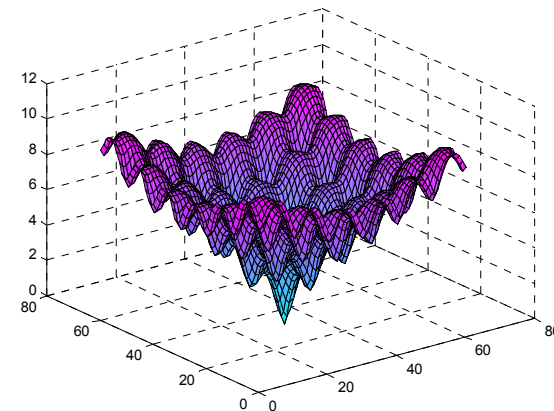
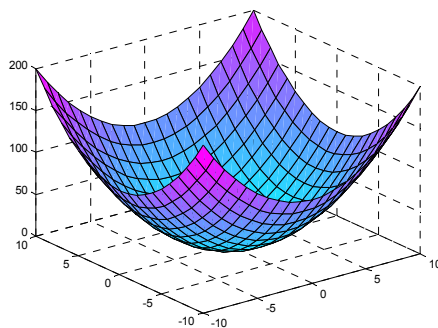
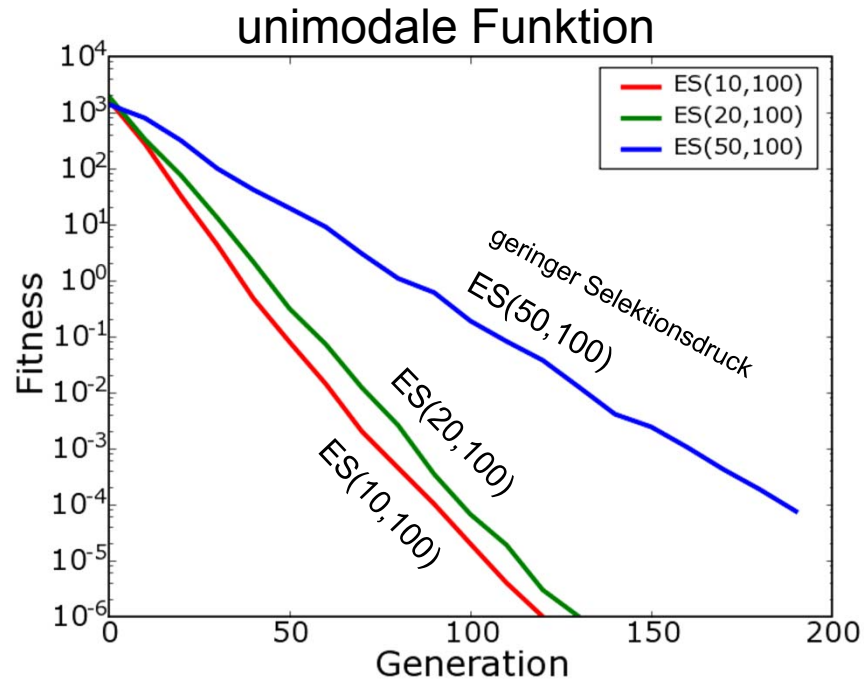
$s \rightarrow 0$: hoher Fortschritt
geringer Einfluß des Zufalls
z.B. geringe Wk. lokale Optima zu überwinden

Ackley's Testfunktion



$$f_6(x) = -20e^{(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^n x_i^2})} - e^{(\frac{1}{n}\sum_{i=1}^n \cos 2\pi x_i)}$$

Selektionsdruck - Beispiel



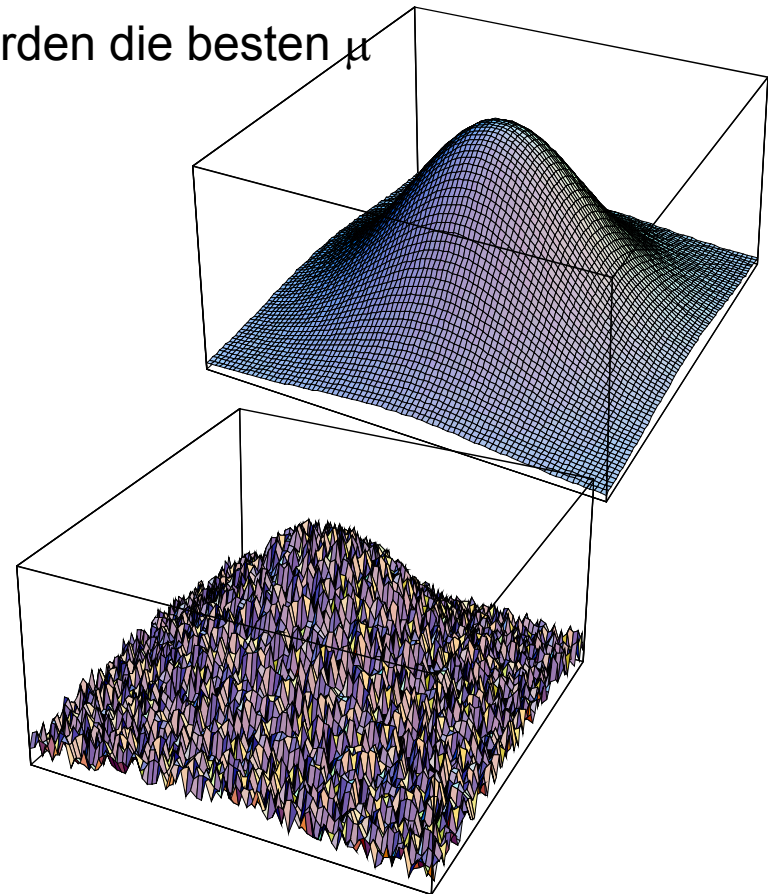
Selektionsstrategie

+ Strategie: Aus $(\mu + \lambda)$ Individuen werden die besten μ Individuen selektiert

, Strategie: Aus den λ Nachkommen werden die besten μ Individuen selektiert

plus Strategie:

- + nur Verbesserungen möglich
- + kein „Vergessen“ von guten Lösungen
- „Hängenbleiben“ in lokalen Minima
- zufällig gute Lösungen bleiben in der Population
(verrauschte Qualitätsfunktion)



EA –Library:

<http://shark-project.sourceforge.net/>



Vorlesung am 25. Juni fällt aus