

Untersuchen des Einflusses von Mutation und Populationsgröße auf die Fitness der entstehenden Individuen

Evolutionäre Algorithmen

Autoren: Paul Jähne
Matrikelnummer: 211801

Björn Schapitz
Matrikelnummer: 210812

Seminargruppe: IT2011

Datum: Leipzig, 03.03.2014

Inhalt

1 Einführung	3
1.1 Zielstellung	3
1.2 Umsetzung	3
1.3 Griewank-Funktion	3
1.4 Nullstellenberechnung.....	3
2. Umsetzung des Algorithmus	3
3. Untersuchen des Einflusses der Populationsgröße	4
4. Untersuchen des Einflusses der Mutation	5
5. Ergebnisse	5

1 Einführung

1.1 Zielstellung

Die Aufgabe umfasste das Entwickeln eines evolutionären Algorithmus zum Berechnen einer näherungsweise optimalen Lösung der Berechnung des globalen Minimums der Griewank-Funktion auf der einen und der Nullstellenberechnung einer gegebenen Funktion auf der anderen Seite. Die Güte der Lösung sollte dabei durch Variation der Parameter verbessert werden.

Insbesondere der Einfluss der Populationsgröße und der Einfluss der Mutation bei konstanten übrigen Parametern wurden in Rahmen dieser Aufgabe untersucht.

1.2 Umsetzung

Zur Entwicklung des Programms wurde die Programmiersprache Java in der Version 1.7 verwendet.

1.3 Griewank-Funktion

$$f_x(x_1, \dots, x_n) = 1 + \sum_{i=1}^n \frac{x_i^2}{400n} - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right)$$

Die Griewank-Funktion wird häufig zum Testen der Konvergenz von Optimierungsfunktionen verwendet. Sie besitzt ein globales Minimum von 0 bei $x = 0$. Das zu betrachtende Intervall beträgt $[-512; 511]$.

1.4 Nullstellenberechnung

$$f_i(x) = x_i + \sum_{j=1}^n x_j - (n + 1) = 0 \quad (i, j = 1, \dots, n - 1)$$

Die Funktion besitzt die Lösung $(a, a, \dots, a, a^{1-n})^T$, wobei a eine Lösung der Gleichung $na^x - (n + 1)a^{n-1} + 1 = 0$ ist. Das zu betrachtende Intervall beträgt $[-10; 10]$.

2. Umsetzung des Algorithmus

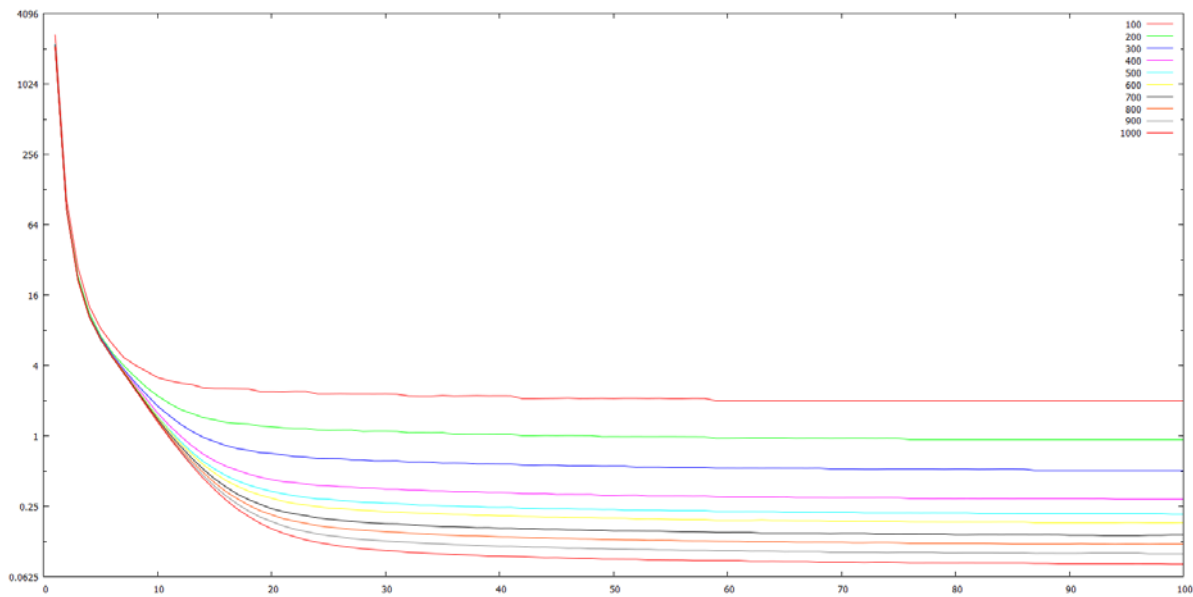
Um die Aufgabenstellung zu lösen, wurden Konzepte der evolutionären Programmierung verwendet. Als Selektionstyp wurde die Kommaselektion gewählt, da sie deterministisch ist und deshalb geringen zufälligen Einfluss besitzt. Als Rekombination kommt die arithmetische Rekombination zum Einsatz. Die Auswahl der zu rekombinierenden Individuen erfolgt dabei zufällig, eine Mehrfachauswahl desselben Individuums ist möglich. Auf Umweltselektion wurde verzichtet. Einzelne Individuen können ebenfalls mehrmals mutiert werden.

Der Algorithmus durchläuft folgende Schritte:

1. Erzeuge Startpopulation p
2. Berechne Fitnesswerte der Individuen der Startpopulation
3. Erzeuge neue Population p_{new} durch comma-Selektion der vorhandenen Population
4. Rekombiniere neue Population p_{new}
5. Mutiere p_{new}
6. Berechne Fitnesswerte der Individuen von p_{new}
7. Ersetze p durch p_{new}
8. Sortiere p nach Fitnesswerten
9. Wiederhole Schritte 3 bis 8 für die Anzahl der Generationen

3. Untersuchen des Einflusses der Populationsgröße

Zu beobachten war, dass die Ergebnisse generell besser wurden, wenn die Startpopulation größer war, unabhängig von Veränderungen während der Generationen.

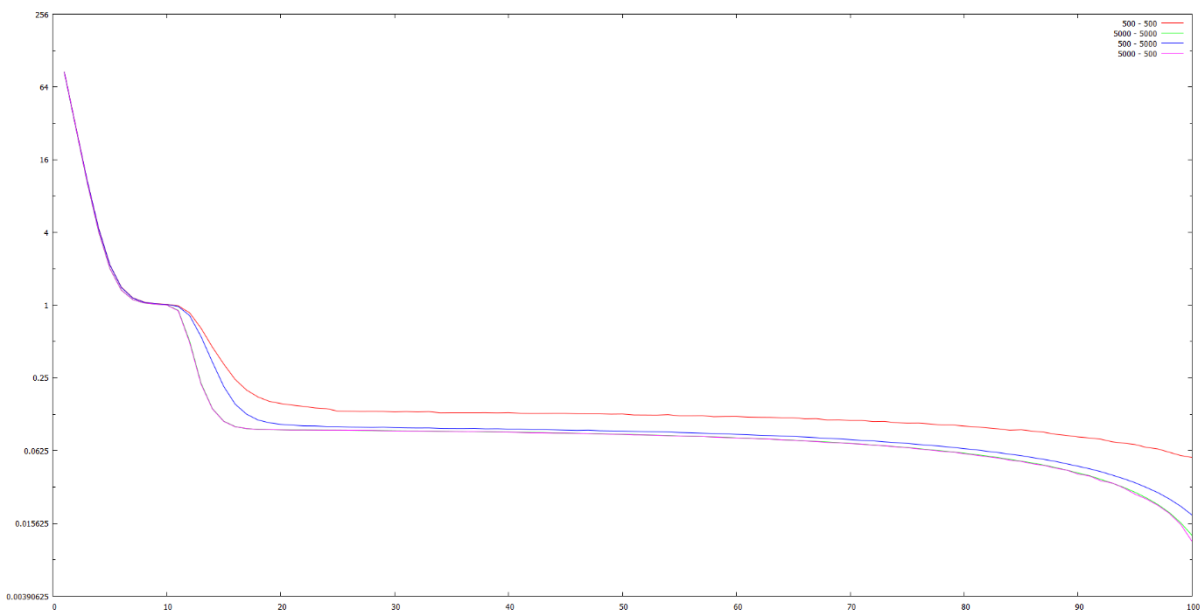


Einfluss Startpopulationsgröße (Griewank-Funktion)

Die Größe der Population kann sich zwischen den Generationen unterschiedlich verhalten:

1. Konstante Populationsgröße: in jeder Generation bleibt die Populationsgröße erhalten
2. Wachsende Population: mit jeder neuen Generation wächst die Populationsgröße an
3. Sinkende Populationsgröße: durch Selektion sinkt die Populationsgröße mit jeder Generation

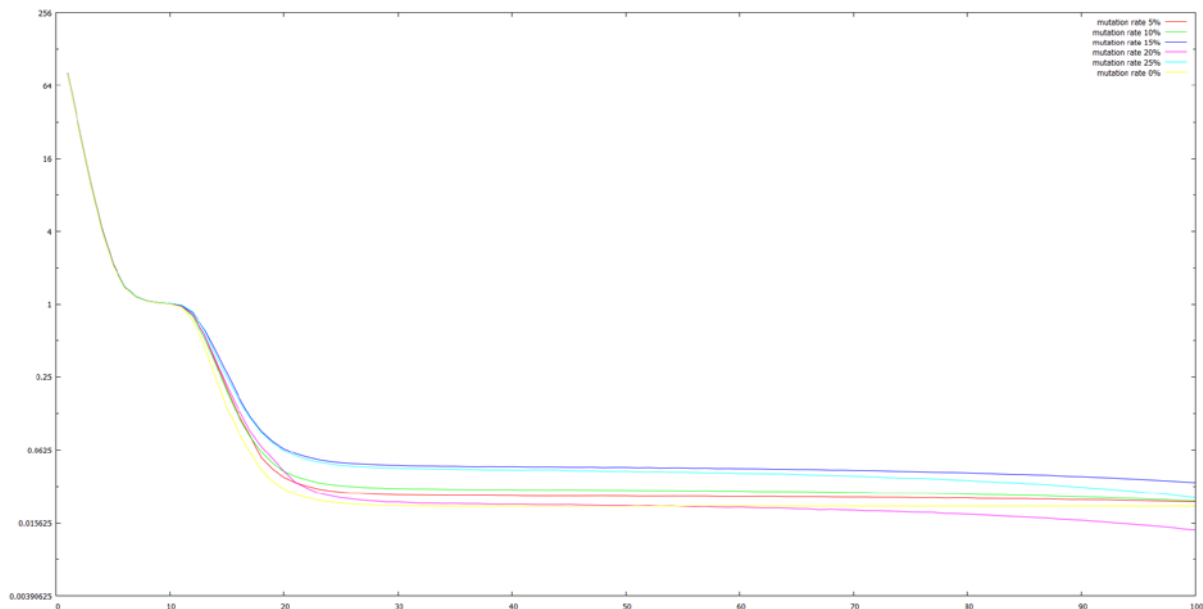
Im Vergleich zur konstanten Populationsgröße zeigt sich, dass bei wachsender Populationsgröße die Güte der Lösungen sinkt, wo hingegen bei sinkender Populationsgröße sich die Güte erhöht.



Einfluss Veränderung der Populationsgröße während der Generationen (Griewank-Funktion)

4. Untersuchen des Einflusses der Mutation

Der Einfluss der Mutation auf die Güte der Individuen kann durch verschiedene Parameter beeinflusst werden. Der Mutationsrate bestimmt, wie viele Individuen der Generation mutiert werden. Die Rate wurde zwischen 0% bis 25% in Schritten von 5% variiert und die Ergebnisse betrachtet.



Einfluss Mutationsrate (Griewank-Funktion)

Über die Mutationsstärke wird geregelt, in welchem Maß das einzelne Individuum mutiert wird. Die Mutationsstärke liegt in einem Intervall $[0,5; 5,0]$ und wird Schritten zu je 0,5 variiert. Außerdem kann festgelegt werden, ob der Verlauf der Mutation über alle Generationen hinweg konstant bleibt oder sich ändert. Bei konstantem Verlauf werden die Individuen der letzten Generation ebenso häufig mutiert wie die der ersten. Um den Verlauf zu beeinflussen, wurden drei unterschiedliche Funktionen verwendet: eine lineare Abnahme, eine exponentielle Abnahme und eine invertierte, exponentielle Abnahme der Mutation über die Generationen hinweg.

5. Ergebnisse

Um die Ergebnisse weniger anfällig für Ausreißer zu gestalten, wurden für jeden überprüften Parameter jeweils 100 Läufe durchgeführt und das Ergebnis mit dem geometrischen Mittelwert gemittelt. Mit zunehmender Allelgröße verschlechtert sich allerdings das Ergebnis erheblich. Ein Grund hierfür mag sein, dass die gewählten Parameter zu deterministisch sind.

Bei der Betrachtung der Populationsgröße ergab sich generell eine Verbesserung der Fitness der Individuen, wenn die Startpopulation vergrößert wurde, unabhängig davon, ob die Population abnahm, zunahm oder konstant blieb. Der Grund hierfür liegt in der größeren Diversität der Individuen der Startpopulation. Variiert man die Populationsgröße während der unterschiedlichen Generationen, dann ergibt sich das beste Ergebnis, wenn man die Größe in jeder Generation durch Selektion verringert. Dies gilt sowohl für die Griewank-Funktion als auch für die Nullstellenberechnung.

Für die Mutation ergeben sich für die beiden unterschiedlichen Funktionen unterschiedlich optimale Parameter. Für die Griewank-Funktion ergeben sich die besten Ergebnisse, wenn die Mutationsstärke bei 0,5 liegt, gefolgt von 3,0. Eine Mutationsrate von 10% zusammen mit einer linearen Abnahme des Mutationsverlaufs über die Generationen hinweg liefert die höchste Güte der Individuen.

Bei der Nullstellenberechnung müssen die Parameter anders gewählt werden. Die Mutationsstärke sollte im Intervall $[0,5; 1,5]$ liegen und die Mutationsrate niedrig sein. Über die Generationen hinweg sollte der Mutationsverlauf exponentiell abnehmen. Mit diesen Einstellungen liefert der Algorithmus die Individuen mit der höchsten Fitness.

Erwähnt werden muss noch, dass sich bei dem Algorithmus nach 30 bis 40 Generationen der beste Fitnesswert nicht mehr verbessert. Der Grund dafür ist, dass zu diesem Zeitpunkt der Genpool nur noch durch Mutation verändert werden kann, die stattfindenden Mutationen aber zu stark / groß sind, um die Individuen in einem sinnvollen Bereich zu mutieren.