Die approbierte Originalversion dieser Diplom-/Masterarbeit ist in der Hauptbibliothek der Technischen Universität Wien aufgestellt und zugänglich.
http://www.ub.tuwien.ac.at

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology.
http://www.ub.tuwien.ac.at/eng

FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Genetische Algorithmen im aktiven Portfoliomanagement

Entwicklung und Implementierung eines interaktiven Prototypen zur Handelsunterstützung im aktiven Portfoliomanagement auf Basis genetischer Algorithmen

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur/in

im Rahmen des Studiums

Masterstudium Wirtschaftsinformatik

eingereicht von

Jan El-Berry, B.Sc.
Matrikelnummer 0625756

an der
Fakultät für Informatik der Technischen Universität Wien

Betreuung
Betreuer/in: Prof. Mag.rer.soc.oec. Dr.rer.soc.oec Gerhard Hanappi

Wien, 06.01.2014

(Unterschrift Verfasser/in) (Unterschrift Betreuer/in)

Erklärung zur Verfassung der Arbeit

Jan El-Berry, BSc.
Margaretengürtel 122-124/5/14, 1050 Wien

„Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe."

(06.01.2014, Wien) (Unterschrift Verfasser)

Danksagung

Ich möchte mich vor allem bei Herrn Prof. Hanappi bedanken, welcher sich mit außergewöhnlicher Offenheit für die Betreuung meines Themas eingesetzt und mir alle Freiheiten zur Bearbeitung meiner Diplomarbeit gewährt hat. Weiters möchte ich mich bei meiner Frau, meiner Familie und meinen Freunden für die lange andauernde Unterstützung und Ausübung des für die Fertigstellung dieser Arbeit notwendigen Drucks bedanken.

Abstract

During the last decade fundamental changes within the real economy as well as the financial sector have occurred. Because of the time-varying number, dynamics and weights of external factors onto the different parts of financial markets, there is a rising need of dynamic analytic tools in the field of portfolio management. To comply with these demands a genetic algorithm has been developed which enables the generation of a dynamic asset allocation with the fitness-function and the underlying data being completely exchangeable. On the basis of this setup, a prototypic graphical user interface has been developed which allows systematic back-testing, as well as a direct access to the algorithms parameters and a possibility to compare the parameter-specific outcome. With the use of evolutionary algorithms - a well-known technique in the field of computer science - onto an application-oriented problem of asset management, the necessary dynamics for a useful optimization framework in active portfolio management can be sufficiently achieved.

Zusammenfassung

Die letzte Dekade ging mit einer Reihe von fundamentalen realwirtschaftlichen sowie finanzwirtschaftlichen Veränderungen einher. Durch die variierende Anzahl, Dynamik und Gewichtung externer Einflussfaktoren auf die verschiedenen Teilbereiche der Finanzmärkte, ergibt sich eine erhöhte Notwendigkeit dynamischer Analysemöglichkeiten im aktiven Portfoliomanagement. Um dies zu ermöglichen, wurde ein genetischer Algorithmus entwickelt mit dem eine Asset-Allokation für ein Portfolio auf Basis frei erweiterbarer Bewertungsfunktionen und Basisdaten möglich wird. Darauf aufbauend ist eine prototypische Testumgebung mit grafischer Benutzeroberfläche entstanden, welche systematisches Backtesting und eine direkte Eingriffsmöglichkeit in die Parametrisierung des Algorithmus sowie einen Vergleich der daraus entstandenen Ergebnisse ermöglicht. Durch die Anwendung evolutionärer Algorithmen aus dem Bereich der Informatik auf ein anwendungsorientiertes Optimierungsproblem der Vermögensverwaltung, soll das Rahmenwerk für ein im aktiven Portfoliomanagement verwertbares Werkzeug hinreichend abgebildet werden können.

Inhaltsverzeichnis

1. Einleitung 1
2. Finanzmärkte und Kapitalmarkttheorien 3
2.1. Einführende Erläuterungen 3
2.2. Asset-Klassen und Finanzinstrumente 4
2.2.1. Anleihen 4
2.2.2. Aktien 6
2.2.3. Rohstoffe 7
2.2.4. Währungen und liquide Mittel 7
2.2.5. Alternative Investments 8
2.3. Arten der Finanzmärkte 9
2.3.1. Kapitalmärkte 10
2.3.2. Geldmärkte 12
2.3.3. Rohstoffmärkte 12
2.3.4. FX-Märkte 13
2.3.5. Versicherungsmärkte 14
2.4. Markteffizienz und die moderne Portfoliotheorie 16
2.4.1. Markteffizienzhypothese 16
2.4.2. Moderne Portfoliotheorie (MPT) 18
2.4.3. Marktanomalien und Kritik 20
3. Portfolio- und Assetmanagement 23
3.1. Fundamentalanalyse 23
3.2. Technische Analyse 24
3.3. Asset-Allokation 25
3.4. Passives vs. aktives Management 25
3.5. Genetische Algorithmen und aktives Portfolio-Management 27
4. Theorie und Praxis genetischer Algorithmen 29
4.1. Prinzipien und Grundlagen 29
4.1.1. Evaluation 31
4.1.2. Selektion 32
4.1.3. Rekombination 33
4.1.4. Mutation 33
4.1.5. Abbruchbedingung 34
4.2. Formalismen, Umsetzungsvarianten und Beispiele 35

4.2.1. Formale Allgemeindefinition und Pseudocode 36
4.2.2. Individuum, Population und Kodierung 37
4.2.3. Selektionsoperator 39
4.2.4. Rekombinationsoperator 43
4.2.5. Bewertungs- und Fitnessfunktion 46
4.2.6. Mutationsoperator 47
4.2.7. Ersetzungs- und Ergänzungsschemata 49
4.2.8. Abbruchkriterium 49
5. Anwendung auf das Portfoliooptimierungsproblem 51
5.1. Rahmenbedingungen und Datenbasis 52
5.2. Experimentelle Definition und Implementierung 56
5.2.1. Individuenkodierung und Initialisierung 56
5.2.2. Fitnessfunktion und Bewertungsansätze 58
5.2.3. Selektions- und Rekombinationsmechanismus 62
5.2.4. Mutation und Abbruchkriterium 63
5.3. Backtesting- und Benchmarkingprozedur 65
5.3.1. Standard-Index 65
5.3.2. 1/N generischer Index 66
5.4. Ergebnisse, Erkenntnisse und Kritik 66
6. Implementierung einer dynamischen Testumgebung 79
6.1. Eingabedaten und Schnittstelle 79
6.1.1. Struktur 79
6.1.2. Synchronisation 80
6.1.3. Vollständigkeit 80
6.2. Benutzeroberfläche 80
6.2.1. Auswahl der Basisdaten 81
6.2.2. Auswahl der Training und Backtesting-Bereiche 82
6.2.3. Konfiguration des genetischen Algorithmus 83
6.2.4. Visualisierung der Optimierungsanalyse 84
6.2.5. Visualisierung der Backtesting-Ergebnisse 85
6.2.6. Sonstige Einstellungsmöglichkeiten 86
7. Conclusio und Ausblick 87

Abbildungsverzeichnis

Abbildung 1 - Primärmart und Sekundärmarkt 10
Abbildung 2 - invertierte vs. normale "Yield - Curve" 12
Abbildung 3 - Aktienpreisentwicklung vor und nach CNBC-Reports [9] 17
Abbildung 4 - Risiko/Rendite Profil eines simplen Portfolios [15 S. 11] 20
Abbildung 5 - Unterstützungslinie und Widerstandslinie 24
Abbildung 6 - Population, Chromosom, Gen 30
Abbildung 7 - Zyklus eines genetischen Algorithmus 31
Abbildung 8 - Ausgangspopulation vs. evaluierte Population 32
Abbildung 9 - Selektion der Individuen 32
Abbildung 10 - Rekombination der Elternpaare 33
Abbildung 11 - Mutation der Individuen 34
Abbildung 12 - Fitnessproportionale Selektion 41
Abbildung 13 - Stochastic Universal Sampling 42
Abbildung 14 - One-point-crossover 44
Abbildung 15 - Multi-point-crossover 44
Abbildung 16 - Uniform-crossover 45
Abbildung 17 - Shuffle-crossover 45
Abbildung 18 - Dow Jones Index 2000 bis 2012 53
Abbildung 20 - Beispiel leeres Portfolio 57
Abbildung 21 - Erstellung eines Zufallsportfolios 57
Abbildung 19 - Anwendungsbeispiel Individuum / Portfolio 57
Abbildung 22 - Hohes Sharpe Ratio vs. niedriges Sharpe Ratio 61
Abbildung 23 - Hohes Calmar Ratio vs. niedriges Calmar Ratio 61
Abbildung 24 - Auswahl der besten Individuen 62
Abbildung 25 - Sortieren der Gene (Zusammenstellen gleicher Gene) 63
Abbildung 26 - Multi-Point-Crossover der Individuen 63
Abbildung 27 - Mutation einzelner Gene 64
Abbildung 28 - Optimierungsgraphen mit unterschiedlichen Mutationsraten 64
Abbildung 29 - Beispielhafter Trainings- und Backtestingzeitraum 67
Abbildung 31 - Beispielhaft Max / Avg / Min Fitness innerhalb der Population 68
Abbildung 30 - Konvergenzgraph des genetischen Algorithmus 68
Abbildung 32 - Portfolioallokation über die Zeit 70
Abbildung 33 - Entwicklung der Vergleichsindizes 2004-2007 71
Abbildung 34 - Backtesting im Zeitraum der Finanzkrise 2008 / 2009 72
Abbildung 35 - Entwicklung der Vergleichsindizes 2008 / 2009 72
Abbildung 36 - Backtesting im Zeitraum der Erholung 2009-2012 73
Abbildung 37 - Entwicklung der Vergleichsindizes 2009-2012 73
Abbildung 38 - Optimierung mittels Calmar Ratio 75
Abbildung 39 - Konvergenzgraph und Portfolio-Allokation mit Calmar Ratio 75
Abbildung 40 - Optimierung über systemkritische Phase 2008/2009 76
Abbildung 41 - Konvergenzgraph und Portfolio-Allokation 76
Abbildung 42 - Optimiertes Portfolio vs. Benchmarks 77
Abbildung 43 - Schema und Beispiel Importdatei 79

Abbildung 44 - Benutzeroberfläche des Prototyps 81
Abbildung 45 - Importieren eines Datensatzes 82
Abbildung 46 - Fixierung der Zeitbereiche 82
Abbildung 47 - Fixierter Trainings- und Backtestingzeitraum 83
Abbildung 48 - Konfiguration der Parameter 83
Abbildung 49 - Optimierungsanalyse 84
Abbildung 50 - Visualisierung des Backtesting 85
Abbildung 52 - Speichern den der generierten Grafiken 86
Abbildung 51 - Zooming in der Anzeige des Optimierungsgraphen 86

Tabellenverzeichnis

Tabelle 1 - Binärkodierung vs. Gray-Kodierung 39
Tabelle 2 - Verteilung der Auswahlswahrscheinlichkeiten im Stochastic Universal Sampling 43
Tabelle 3 - Vergleiche zwischen den Crossover-Typen 46
Tabelle 4 - Zusammensetzung Dow Jones Index zum Jahr 2000 55
Tabelle 5 - Beispielparameter des genetischen Algorithmus 68
Tabelle 6 - Resultierende Gewichtung der Assets 70
Tabelle 7 - Messzahlen der Indizes von 2004-2007 71
Tabelle 8 - Messzahlen der Indizes von 2008 / 2009 72
Tabelle 9 - Messzahlen der Indizes von 2009-2012 73
Tabelle 10 - Optimiertes Portfolio vs. Benchmarks über gesamten Zeitraum 74
Tabelle 11 - Messzahlen der Indizes von 2004-2012 74
Tabelle 12 - Kennzahlen der Portfolios im Vergleich 77

1. Einleitung

Die zunehmende Dynamik der weltweiten Finanzmärkte während der letzten Dekade stellt die klassischen analytischen Portfoliooptimierungsverfahren zunehmend in Frage. Durch die daraus resultierende erhöhte Variabilität externer Markteinflussfaktoren sowie deren über die Zeit veränderte Gewichtung, lassen sich nur schwierig zuverlässige Modelle auf Basis fixierter Einflussgrößen ermitteln. Im Speziellen im Bereich des aktiven Portfoliomanagement muss ein Weg gefunden werden, ein möglichst optimales Portfolio in Bezug auf Rendite und Volatilität unter den vorgegebenen Rahmenbedingungen zu kreieren. Die Thematik an sich steht oftmals zwischen den Fronten von Wissenschaft und Praxis, sowie Rationalität und Psychologie.

Die Basis der Arbeit stellt daher eine Auseinandersetzung mit den Grundlagen moderner Kapitalmärkte sowie den klassischen Methoden der Portfoliooptimierung dar. Die daran anschließenden Ausführungen erläutern genetische Algorithmen als alternatives heuristisches Lösungsverfahren und präsentieren eine mögliche Überführung des Portfoliooptimierungsproblems in eine für genetische Algorithmen lösbare Form. Dabei werden sowohl die besonderen Möglichkeiten als auch die Grenzen des erläuterten Verfahrens und insbesondere deren Einsatz im aktiven Portfoliomanagement in der Finanzwirtschaft aufgezeigt. Auf dieser Basis wird in weiterer Folge ein experimenteller genetischer Algorithmus entwickelt, welcher für einen vorgegebenen Testzeitraum ein global optimales Portfolio, bestehend aus den gewichteten Bestandteilen einzelner Assets, unter der Voraussetzung zuvor spezifizierter Bewertungskriterien (Rendite und Volatilität), generiert. Weiters soll die Möglichkeit einer dynamischen Backtesting-Umgebung geschaffen werden - das über einen Trainingszeitraum optimierte Portfolio soll also auf beliebigen Zeiträumen auf die vom Portofolio-Manager gewünschten Eigenschaften hin überprüft werden können.

Der auf diesem Weg erzeugte Algorithmus soll im Rahmen einer eigenständig lauffähigen Applikation zur Verfügung gestellt werden. Die Anwendung muss dabei jene Grundanforderungen bezüglich Dynamik und Adaptierbarkeit erfüllen, welche für einen möglichen Portfoliomanager zur Portfoliooptimierung unabdingbar sind. Wesentliche Bestandteile sind daher die Austauschbarkeit der Basisdaten, die freie Wahl der Trainings- und Backtesting Zeiträume sowie die dynamische Parametrisierung des genetischen

Algorithmus an sich und ein möglicher Vergleich zu Passivportfolios. Um den Zugang so einfach wie möglich zu gestalten, wurde eine grafische Benutzeroberfläche implementiert über den der gesamte Optimierungsprozess beobachtet und gesteuert werden kann.

2. Finanzmärkte und Kapitalmarkttheorien

2.1. Einführende Erläuterungen

Der globale Finanzmarkt kann als das Nervensystem jedes marktwirtschaftlichen Kreislaufs bezeichnet werden. Er beschreibt im Allgemeinen ein weltumspannendes Netzwerk, welches Käufer und Verkäufer miteinander verbindet und Transaktionen vereinfacht und unterstützt. Daraus folgend, ist die effiziente und möglichst reibungslose Allokation von Kapital eine primäre Funktion der Kapitalmärkte. Die Entscheidung darüber welche Märkte oder Unternehmen aktuell, sowie für die vermeintlich vorhersehbare Zukunft Wert sind, mit einer bestimmten Summe an Kapital versorgt zu werden, obliegt den Investoren. Der Preis eines Investitionsobjektes (z.B. einer Aktie) ergibt sich also indirekt aus der Summe der Kapitalzuflüsse und Abflüsse zu einem bestimmten Zeitpunkt und spiegelt daher die Einschätzung eines Kollektivs wieder. Wesentlich dabei ist, dass jede Investition unter Unsicherheit geschieht. Oberflächlich betrachtet erscheint der Prozess der Kapitalallokation daher oftmals exzessiv und irrational, er ist jedoch oftmals eine unausweichliche Implikation aus potentiell großen Möglichkeiten bei zugleich hoher Unsicherheit. Finanzmärkte agieren daher als Proxy eines institutionellen oder privaten Anlegers zur (un)mittelbaren und dezentralen Steuerung zukünftiger wirtschaftlicher Entwicklungen.

Ein wesentlicher Vorteil globaler Finanzmärkte ist einerseits die meist hohe Liquidität und einheitliche Sprache bzw. Schnittstelle welche über die verschiedensten Intermediäre von jedermann in Anspruch genommen werden kann. Hieraus ergibt sich die Möglichkeit erwirtschaftetes Vermögen tendenziell ertragsbringend (daher oftmals zumeist rudimentär inflationsgeschützt) in beispielsweise Aktien oder Anleihen "einzufrieren" und diese bei Bedarf zu einem späteren Zeitpunkt wieder zu liquidieren. Kapital welches im Zeitraum der Erwerbstätigkeit nicht verkonsumiert werden muss, kann so also für spätere Perioden aufbewahrt werden. In Abhängigkeit des Alters und der Höhe des momentanen Einkommens kann eine Allokation des Risikos mit Hilfe der verschiedenen Investmentvehikel durchgeführt werden - während in jüngeren Jahren Investments auf Grund der langen Zeiträume bis zur ersten Liquidierung der Vermögenswerte mit tendenziell hoher Volatilität getätigt werden sollten, muss bei fortgeschrittenem Alter die Risikokomponente stärker in Betracht gezogen werden.

In Summe stellt der Finanzmarkt also ein essenzielles Konstrukt dar, durch welches das fundamentale Bedürfnis gestillt werden kann, notwendige und vermutlich höhere Konsumausgaben auf einkommensschwächere Zeiträume zu verschieben. Ein sich aus den vorherigen Erläuterungen ergebendes Anwendungsbeispiel, stellen staatliche und private Pensionsfonds dar. Diese sind entgegen allgemeiner Meinung die bei weitem größten und einflussreichsten Akteure an den internationalen Finanzmärkten. Sie agieren gebündelt für ein Kollektiv an Privatpersonen und genießen oftmals Steuervergünstigungen bei der Kapitalveranlagung.

Trotz der dezentralen Funktionsweise der Finanzmärkte, existieren eine Vielzahl großer, neuralgischer Punkte die auf Grund einer enormen Anzahl von Interaktion in der Vergangenheit, wechselseitig hohe und dauerhafte Abhängigkeiten für die Zukunft entwickelt haben.

Zu den am häufigsten gehandelten Waren gehören Aktien, Anleihen, Rohstoffe, Währungen sowie Finanzinstrumente zur Absicherung jeglicher Art von Risiko. In weiterer Folge soll eine kurze Einführung zu den wichtigsten Asset-Klassen mit erläuternden Referenzen zur globalen Finanzkrise das Verständnis für die weiteren Kapitel erleichtern.

2.2. Asset-Klassen und Finanzinstrumente

Eine Unterteilung der typischerweise gehandelten Vermögensgegenstände in Anlageklassen, wird je nach Methodik zur Erläuterung des Sachverhalts höchst unterschiedlich gewählt. Da im Rahmen dieser Arbeit sämtliche Erläuterungen im Kontext des klassischen Portfoliomanagements und der Portfoliooptimierung erläutert werden sollen, wird die oftmals erwähnte Asset-Klasse "Immobilien" auf Grund geringerer Relevanz nicht behandelt. Der manchmal in den anderen Klassen subsumierte Punkt "Alternative Investments" (beinhaltet z.B. Derivate) wird dafür explizit erläutert - Fondsanbieter beschreiben diesen Punkt häufig als Teil der Vermögensallokation und renditeträchtige Anlagestrategie. In weiterer Folge sollen jene Asset-Klassen mit standardisierter Handelbarkeit wie in [1] strukturiert und in [2] jeweils näher beschrieben, erläutert werden.

2.2.1. Anleihen

Anleihen sind Schuldverschreibungen über welche die genauen Konditionen festgelegt werden können, zu denen ein Gläubiger (Anleihekäufer) einem Schuldner (Anleiheemittent)

Kapitalmittel für einen definierten Zeitraum zur Verfügung stellt. Im Normalfall wird der Gläubiger über eine regelmäßige Zahlung des Schuldners über den Zeitraum der Schuld vergütet - diese Bezahlung wird als Kupon bezeichnet (d.h. "Kupon-Anleihe"). Manchmal erfolgt die Vergütung auch erst gemeinsam mit der gesamten Rückzahlung am Ende des Zeitraums (d.h. "Zero-Kupon-Anleihe"). Um die Höhe des Ertrags unabhängig von der Art der Anleihe angeben zu können, wird auch allgemein von Rendite (engl. "Yield") gesprochen. Diese bezeichnet die zum aktuellen Zeitpunkt erwarteten Gesamteinkünfte eines Anleihe-Investments. Sobald eine Anleihe begeben wurde, besitzt diese einen Kurs auf deren Basis sie gehandelt werden kann. Je nachdem wie sich die Bonität eines Schuldners über die Zeit entwickelt, kann sich der Kurs dieser Anleihe zwischenzeitlich ganz massiv zum Nachteil bzw. zum Vorteil des Schuldners entwickeln. Grundsätzlich gilt jedoch, solange die Schuld beglichen werden kann und der Investor zum Ausgabepreis gekauft hat, hat dieser keine Verluste zu befürchten - selbst wenn der theoretische Marktwert des Wertpapiers noch am Tag vor der Auszahlung nahe Null liegt. Der Kurswert spiegelt veränderte Rahmenbedingungen wieder und kann die potentielle Rendite für einen Zweitbesitzer der Anleihe auf Grund gestiegenen Ausfallsrisikos erhöhen (Kurs liegt "unter Par", d.h. < 100) oder senken (Kurst liegt "über Par", d.h. > 100). Ein möglicherweise vorhandener, laufend bezahlter Kupon wird daher letztlich durch eine Begleichung auf exakt 100 geschmälert bzw. erhöht.

Der Bereich der Anleihen ist durch eine besonders hohe Anzahl an Finanzprodukten und Investmentvehikeln geprägt. Zu den wichtigsten handelbaren, übergeordneten Typen gehören:

a. Staatsanleihen: Die mit Abstand höchste Anzahl an Anleihen-Varianten (z.B. inflationsgeschützte Anleihen) sowie gehandelten Volumina befindet sich im Bereich der Staatsanleihen. Diese werden von Staaten begeben um z.B. Infrastrukturprojekte und andere Investitionen zu finanzieren sowie um auslaufende Schulden zu refinanzieren. Die gängigsten Laufzeiten betragen 1-10 sowie 20 und 30 Jahre - wobei 10-Jährige die am häufigsten gehandelten Wertpapiere sind. Auf Grund der weltweiten Vormachtstellung des US-Dollars gehören die U.S. Treasury Bonds ("T-Bonds") zu den bedeutendsten Anleihen.

b. Unternehmensanleihen: Eine der günstigsten Arten der Refinanzierung für Unternehmen besteht in der Ausgabe von Anleihen an das öffentliche Börsenpublikum. Der wesentlichste Unterschied zu den meisten Staatsanleihen besteht im Ausfallsrisiko, welches üblicherweise durch eine höhere Gesamtrendite kompensiert wird. Im Falle eines Zahlungsausfalls werden bei der Unternehmensabwicklung sogenannte gesicherte Anleihen von ungesicherten Obligationen unterschieden. Halter der Ersteren werden beim Insolvenzverfahren vorrangig bedient.

c. Hypothekenanleihen: Haben sich in den letzten Jahrzehnten zu einer häufig gewählten Anlageform und daher einer großen Komponente des Anleihenmarktes entwickelt. Auf Grund des geringen Volumens der Einzeltitel und der hohen Varianz der Schuldnerbonität, war es bis zu den 1970er Jahren nur einer sehr geringen Anzahl an Investoren möglich in hypothekenbesicherte Schuldverschreibungen zu investieren. Dies änderte sich mit der Bündelung der Kredite gleichartiger Schuldner durch Institutionen wie der ehemalige Federal National Mortgage Association (Fannie Mac) und der teilweisen Besicherung durch den Staat zur Förderung des Eigenheimbaus. Diese weitergereichten und gebündelten Wertpapiere (pass-through securites) konnten von einem breiten Investorenpublikum gekauft werden. Im Zuge der Finanzkrise 2008 welche durch die US-Immobilienkrise ausgelöst wurde, trocknete der Markt vollkommen aus. Fannie Mac und ähnliche Institute wurden insolvent und brachten weltweit Banken und Versicherungen zum Schwanken. Im Zuge der Erholung aller Asset-Preise von 2009 bis 2013 besserte sich auch der Markt für Hypothekenanleihen wieder.

2.2.2. Aktien

Unter Aktien werden üblicherweise Anteile eines Unternehmens verstanden. Mit dem Besitz eines solchen Anteilsscheines erwirbt man neben dem Recht auf finanzielle Beteiligung am Unternehmenserfolg, oftmals auch ein Stimmrecht für die mindestens einmal jährlich stattfindenden Hauptversammlungen. Im Zuge der Versammlungen wird unter Anderem Kapitalerhöhungen, Restrukturierungen und Satzungsänderungen entschieden. Grundsätzlich wird zwischen zwei Formen von Aktientypen unterschieden:

a. Stammaktie: Gewährt dem Besitzer einer solchen Aktie exakt ein Stimmrecht auf den Hauptversammlungen. Üblicherweise sind Stammaktien bei den Auszahlung von Dividenden (Gewinnausschüttungen) benachteiligt, besitzen also eine deutlich geringere

Dividende als eine Vorzugsaktie. Da diese aus fundamentaler Sicht eine höhere Bedeutung hat (z.B. bei Übernahmen) liegt sie preislich häufig über der Vorzugsaktie.

b. Vorzugsaktie: Besitzern von Vorzugsaktien wird, wie der Name bereits vermuten lässt, eine bevorzugte Behandlung bei der Ausschüttung von Dividenden zuteil. Des weiteren werden diese im Fall einer Insolvenz ebenfalls vorranging behandelt, im Gegenzug verzichtet der Aktionär auf sein Stimmrecht bei Versammlungen.

2.2.3. Rohstoffe

Rohstoffinvestments sind seit den 2000er Jahren im Zuge des Platzens der New-Economy zunehmend beliebter geworden. Die wichtigsten Kategorien sind Öl, Edelmetalle (Gold, Silber, Platin), Industriemetalle (Kupfer) und Agrarrohstoffe. Ein Investment erfolgt aus Gründen der Kosteneffizienz mit Hilfe von Warentermingeschäften (auch "Futures" genannt). Inzwischen existiert eine Vielzahl von Rohstofffonds mit deren Hilfe ein indirektes Rohstoffinvestment möglich ist. Häufig werden Rohstoffe im Rahmen eines breit über alle Anlageklassen gestreuten Portfolios genutzt um die Schwankungsbreite des Portfoliowertes in Summe zu verringern.

2.2.4. Währungen und liquide Mittel

Bei einem Investment in jeder der zuvor beschrieben Asset-Klassen wird zugleich in die unterliegende Handelswährung investiert. Kauft man eine US-Staatsanleihe, ist dieses Wertpapier mit US-Dollar unterlegt, bei der Anschaffung einer brasilianischen Aktie besitzt man zugleich Geldeinheiten brasilianischen Real. Sowohl Kupon- als auch Dividendenzahlungen werden im Regelfall in der jeweiligen Währung ausgeschüttet. Der Wert einer Währungseinheit wird stets im Verhältnis zu einer beliebigen anderen Währungseinheit bemessen. Die Notation erfolgt daher relativ wie unten stehende Beispiele zeigen:

EUR/USD = 1,2801 USD/JPY = 92,8775 EUR/CHF = 1,2141

Neben der indirekten Anschaffung von Devisen, kann auch ein Direktinvestment in eine Währung mittels Währungstausch und Anlage auf einem Festgeldkonto erfolgen. Im Rahmen des aktiven Portfoliomanagements wird in Portfolioberichten explizit das "Währungsexposure", also der Anteil der jeweiligen Währungen am Portfolio, die sich aus der Summe der impliziten und expliziten Währungsanlagen ergeben, ausgewiesen.

Grundsätzlich gilt, dass sich das Risiko/Ertrags Profil eines Portfolios deutlich verbessert, wenn eine ausgewogener Mischung der größten, nicht regulierten Währungsräume vertreten ist - also japanischer Yen, US-Dollar sowie Euro.

2.2.5. Alternative Investments

Der zurechenbare Umfang von Anlagegegenständen zum Bereich alternative Investments ist in der Literatur oftmals strittig und besitzt daher eine hohe Bandbreite an möglichen Definitionen. Im Zuge aktiven Portfoliomanagements werden diesem Bereich vor allem Investitionen zugeordnet, hinter welchen ein Konstrukt aus oftmals mehreren Derivaten steckt, welche im Zusammenspiel aus einer zukünftig erwarteten Marktentwicklung bzw. Konstellation Profit schlagen sollen. Vor der Darlegung eines greifbaren Beispiels, werden die für diese Art der Spekulation primär verwendeten Finanzprodukte in Anlehnung an die thematisch relevanten Ausführungen in [3] überblicksmäßig erläutert:

a. Optionen: Neben Aktien und Anleihen gehören Optionen zu den ältesten Finanzinstrumenten. Wie der Name bereits impliziert, erkauft man sich ein Ausübungsrecht einer Transaktion (Kauf oder Verkauf) einer Ware innerhalb einer festgelegten Frist (amerikanische Option) oder zu einem bestimmten Zeitpunkt (europäische Option). Beispielsweise kann der Kauf einer Aktie in 4 Monaten zum Preis X bereits in der Gegenwart fixiert werden. Der Preis der für diesen Vertrag zu bezahlen ist, ergibt sich unter anderem aus der Höhe der Preisdifferenz der Ware aus heutiger Sicht und dem vereinbarten Lieferpreis, sowie der durchschnittlichen Schwankungsbreite (Volatilität) der Ware und der zeitlichen Distanz bis zum Verfall der Option. Tendenziell geringere Einflussgrößen sind der kurzfristige, risikofreie Zins und zu erwartende Dividendenzahlungen (bei Aktien als Basiswert). Hinter jedem Optionskäufer steht ein Stillhalter welcher die Gegenseite des Geschäftes akzeptiert und dafür mit einer Optionsprämie vergütet wird, unabhängig davon ob der Käufer die Option letztendlich ausübt oder verfallen lässt.

b. Futures: Im Gegensatz zur Option ist ein Future durch seinen verpflichtenden Charakter gekennzeichnet. Der Zeitpunkt sowie die Menge und der Preis bei Lieferung bzw. Entgegennahme des Basiswerts wird vertraglich festgelegt. Da die letztendliche Ausübung nicht optional ist, fließen keinerlei Prämien an den Vertragspartner. Der Future kann jedoch auch vor Ablauf der Lieferfrist weiterverkauft werden - in

Abhängigkeit von der zwischenzeitlichen Entwicklung des Basiswertes können dabei entsprechend beträchtliche Verluste oder Gewinne fällig werden.

c. Swaps: Allgemein gesprochen werden bei einem Swap Zahlungsströme getauscht. Wenn bspw. die Zahlungen einer zinsvariablen Anleihe gegen Fixzinszahlungen eingetauscht werden, kann damit ein zuvor eingegangenes Zinsänderungsrisiko abgesichert werden.

Zwei typische Handelsbeispiele die allgemeinhin als alternative Investments bezeichnet werden und mit Optionen oder Swaps realisiert werden können, sollen ebenfalls kurz erläutert werden. Nähere Informationen hierüber finden sich in [4].

a) Handeln von Volatilitäten: Volatilität bezeichnet die historische Schwankungsbreite eines Wertpapieres über einen definierten Zeitraum. Üblicherweise steigt die Volatilität explosionsartig sobald ein plötzlicher Abverkauf erfolgt und sinkt kontinuierlich sobald sich der Preis des Wertpapieres stabilisiert - ein Anstieg der Volatilität kann auch bei diametral entgegengesetztem Verhalten stattfinden, dies geschieht jedoch äußerst selten. Mit Hilfe von Optionen kann nun ein Konstrukt (der sogenannte "Straddle") erzeugt werden, welches umso profitabler ist je stärker sich die Schwankungsbreite eines Assets erhöht.

b) Handeln von Korrelationen: Korrelationen beschreiben die Stärke der Beziehung zwischen mindestens zwei Variablen bzw. Ereignissen - diese kann sowohl stochastisch als auch kausal sein. Bricht die Korrelation zwischen zwei auf Grund von fundamentalen Gründen üblicherweise hochgradig miteinander korrelierten Rohstoffen auf Grund eines externen Ereignisses, ist es nun möglich auf die Wiederherstellung der üblichen Korrelation zu spekulieren. Konkrete Beispiele dafür wären die Preisentwicklung zwischen Mais und Weizen sowie den Ölsorten WTI und BRENT.

2.3. Arten der Finanzmärkte

Die Gründe zur Partizipation an den internationalen Finanzmärkten sind so unterschiedlich wie die darüber gehandelten Waren und Verpflichtungen. Die Unterscheidung der im Folgenden gelisteten Markttypen ist vor allem logischer Natur, um einen einfach zugänglichen Einblick in eine über die Jahrhunderte gewachsene und komplexe

Marktstruktur geben zu können. In [5] wird zwischen folgenden, gut beschreib- und abgrenzbaren Arten von Finanzmärkten unterschieden:

2.3.1. Kapitalmärkte

Dieser Typus repräsentiert eine Kategorie der Finanzmärkte, welche tendenziell zur langfristigen Kredit- und Geldanlage genutzt wird. Der Kapitalmarkt wird logisch in den sogenannten *Primär-* und den *Sekundärmarkt* unterteilt. Am *Primärmarkt* werden neu emittierte Aktien oder Anleihen üblicherweise von einer oder mehreren Investmentbanken an interessierte Investoren direkt verkauft. Der Sekundärmarkt fungiert der Logik folgend im übertragenen Sinne als "Gebrauchtwarenmarkt". Sobald ein Vermögenswert an einer oder mehreren Börsen (z.B. der New York Stock Exchange) offiziell gelistet ist, können die Anteilhaber ihre Werte an den entsprechenden Umschlagplätzen zum Verkauf anbieten. Sämtliche allgemein zugänglichen Börsencharts repräsentieren daher die Kursentwicklung eines Vermögenswertes am Sekundärmarkt. Dieser ergibt sich letztendlich ausschließlich aus Angebot und Nachfrage der handelnden Personen. Nachstehende schematische Abbildung veranschaulicht die erläuterte Struktur.

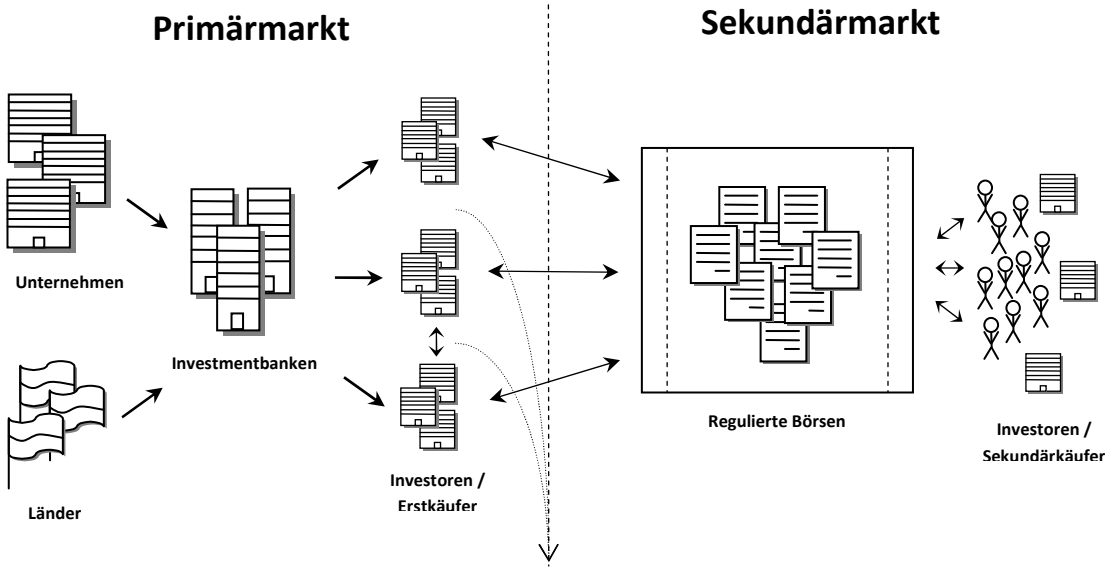

Abbildung 1 - Primärmart und Sekundärmarkt

Die in der Grafik als "OTC" markierten Verbindungen beschreiben "over-the-counter" Geschäfte. Diese finden fernab jedes Marktes und direkter Marktaufsicht dezentral zwischen zwei Entitäten statt. Der gehandelte Preis wird dabei im Normalfall nicht öffentlich gemacht. Durch diese bilateralen Verträge setzen sich die handelnden Subjekte erhöhten Risiken aus.

Ein vor allem im Zuge der globalen Finanzkrise 2008 auf breiter Linie schlagend gewordenes Risiko war jenes des im Allgemeinen als "counter-party risk" spezifizierten Ausfallrisikos des Geschäftspartners mit welchem der Kaufvertrag geschlossen wurde. Das abgekürzt als OTC-Markt bezeichnete Sammelsurium sämtlicher bilateral eingegangen Geschäfte (über sämtliche denkbaren Vermögensobjekte), steht seit der Finanzkrise vor allem wegen seiner Intransparenz und Unkontrollierbarkeit in Kritik.

Die eigentlichen Unterkategorien werden durch die folgenden beiden Märkte repräsentiert:

a. Aktienmärkte: Stellen eine der wichtigsten Quellen von Aktiengesellschaften dar um an Kapital gegen Unternehmensanteile im Rahmen von Kapitalerhöhungen zu tauschen. Weiters bietet dieser eine öffentlich zugängliche Plattform zur transparenten Preisbildung eines Anteils. Über den Wert der gehandelten Einheit kann im Groben der Unternehmenswert zu einem Zeitpunkt ermittelt werden. Umso höher die Anzahl der Teilnehmer, umso fairer der Preis und höher die Geschwindigkeit zu der ein Vermögenswert veräußert werden kann - diese Regel gilt für sämtliche beschriebenen Finanzmärkte. Die NYSE ist eine der größten und liquidesten Aktienmärkte der Welt.

b. Anleihenmärkte: Bezeichnet sämtliche Märkte an welchen Schulden begeben oder gehandelt werden können. Er dient also primär zur langfristigen Finanzierung öffentlicher und privater Körperschaften. Die Struktur des Anleihenmarktes entspricht jener in Abbildung XY gezeigten, jedoch werden die bei weitem größten Volumina im Primärmarkt sowie dezentralisiert "over the counter" gehandelt. Lediglich ein relativ geringer Anteil der emittierten Schuldverschreibungen ist an einer Börse gelistet und daher am Sekundärmarkt frei handelbar. Die Preise und Preisentwicklungen für Anleihen am Sekundärmarkt sind oftmals ein wichtiger Hinweis auf die Entwicklung der Bonität eines Schuldners über die Zeit. Sinken die Anleihepreise (daher: steigen die Zinsen) eines Schuldners (z.B. Italien, Spanien, Griechenland) stellt dies im Normalfall kein sofortiges und akutes Problem dar, es zeigt allerdings die von potentiellen Investoren aktuell geforderte Risikoprämie für ein Investment in zukünftig emittierte Schuldverschreibungen. Der US-Anleihenmarkt ist mit Ende 2012 nach wie vor der größte der Welt, mit einem Gesamtvolumen an ausständigen Schulden von ca. 38,7 Billionen (=38.700 Milliarden) [6].

2.3.2. Geldmärkte

Im Unterschied zu Kapitalmärkten sind Geldmärkte vor allem zur kurzfristigen Finanzierung und Anlage von bis zu 12 Monaten immens wichtig. Handel am Geldmarkt wird vor allem OTC betrieben und dient als rasche Liquiditätsquelle des globalen Finanzsystems. Kurzfristiger Handel von überschüssigen Geldmitteln erhöht die wirtschaftliche Effizienz enorm und bietet rasche, temporäre Geldanlage für den Kreditor und unkomplizierten Überbrückungskredit für den Debitor. Naturgemäß agieren am Geldmarkt vor allem Staaten, Banken und große Unternehmen - auf Basis der gehandelten Zinssätze lässt sich die Entwicklung kurzfristiger Zinsen über die Zeit erkennen, welche gemeinsam mit den längerfristigen Zinsen als wichtiger wirtschaftlicher Indikator genutzt wird. Trägt man die Höhe der Zinsen über verschiedene Verschuldungszeiträume hinweg ein, ergibt das die sogenannte "Yield-Curve". Sind kurzfristige Zinsen bei Staaten höher als Langfristige, erwartet der Markt eine moderate Rezession.

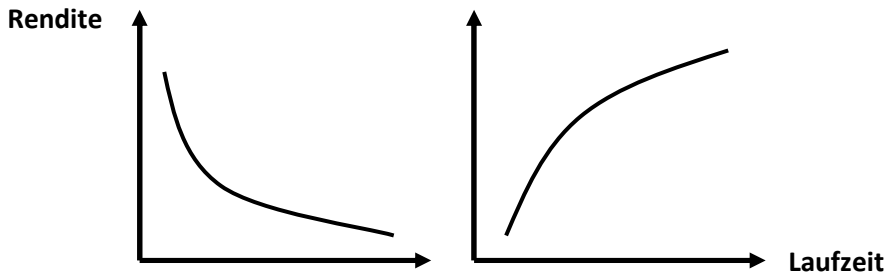

Abbildung 2 - invertierte vs. normale "Yield - Curve"

Eine der katastrophalsten Auswirkungen der Finanzkrise bestand in der völligen Austrocknung des globalen Geldmarktes. Da sich Banken untereinander misstrauten und zudem auf Grund massiv fallender sämtlicher Vermögenswerte (vor allem jener der sog. Asset Backed Securities) gezwungen waren, sämtliche vorhandenen Geldmittel zu horten. Als Intermediär sind weltweit Zentralbanken eingesprungen um einen Systemkollaps und massive Deflation zu verhindern.

2.3.3. Rohstoffmärkte

Diese zählen zu den ältesten und vom realwirtschaftlichen Standpunkt sicherlich wichtigsten Märkten der Welt. Hier werden hauptsächlich Rohstoffe in standardisierter Form und Mengeneinheiten (Kontrakten) sowie erste verarbeitete Produkte gehandelt. Den Ursprung der Handelstätigkeit bilden landwirtschaftliche Produkte, vornehmlich Weizen und Mais. Im

Laufe der Zeit sind eine Reihe von zusätzlichen Rohstoffen wie Soja, Zucker, Kakao und Baumwolle hinzugekommen. Man unterscheidet vor allem zwischen dem "Spot"- und dem "Futures"-Trading. Kauft man eine Ware "Spot" dann geschieht die Lieferung der Ware zum nächstmöglichen Zeitpunkt. Mittels Futures-Kauf wird ein bestimmter Zeitpunkt vereinbart zu welchem die Ware geliefert werden muss. Futures-Preise können sich je nach erwarteter Preisentwicklung zum Teil massiv vom aktuellen Spot-Preis unterscheiden. Das Prinzip des Futures ist eines der ältesten finanzwirtschaftlichen Konzepte (bzw. Derivat) und bedient ein fundamentales Sicherheitsbedürfnis der verkaufenden Partei. So kann unabhängig von der tatsächlichen Preisentwicklung ein Preis fixiert werden und für die noch nicht erzeugte Ware bereits vorab ein Kaufvertrag geschlossen werden der für den Erzeuger entsprechende Planungssicherheit bringt. Heutzutage wird die Majorität der Kontrakte ohne physisches Settlement gehandelt - üblicherweise wird nur der Geldwert zum Auslauf des Kontraktes bedient (Cash-Settlement). Einer der Hauptgründe dafür ist, dass Rohstoffe in den letzten 10 Jahren von einer Vielzahl von Investmentfonds als weitere Asset Klasse zur Kapitalanlage und Portfolio-Diversifikation entdeckt wurden. Der weltweit größte Warenterminmarkt ist die New York Mercantile Exchange (NYMEX).

2.3.4. FX-Märkte

Der FX-Markt ist sicherlich der wichtigste Markt zur Unterstützung des globalen Handels und von dezentralisierter Natur. Es existiert daher nicht eine zentrale Handelsstelle sondern eine große Zahl von miteinander vernetzten Marktplätzen auf denen verschiedenste Finanzinstrumente gekauft und verkauft werden können. Zu den wichtigsten Handelsplätzen gehören London und New York, eine Reihe großer Banken bietet jedoch ebenfalls Umschlagplätze und Direkthandelsmöglichkeiten. Referenzkurse für Währungspaare existieren dadurch nicht, diese ergeben sich auf Basis von Angebot und Nachfrage auf der jeweiligen Handelsplattform. Auf Grund von ausgeprägter Arbitragetätigkeit sind Preisdifferenzen zwischen verschiedenen Handelsplätzen heutzutage jedoch de facto ausgeschlossen.

Diese vielen, teilweise direkt miteinander verbundenen Marktplätze erlauben es Währungen in einem bestimmten Verhältnis zueinander zu tauschen und dadurch Auslandsinvestments und Handel stark zu vereinfachen. Wird also ein US-Investor in indischen Rupien bezahlt, kann er auf Grund der am FX-Markt allgemein sehr hohen Liquidität die Fremdwährung mit

geringstem Reibungsverlust in seine Heimatwährung umtauschen. Weiters ist die effiziente Absicherung von Risiken durch Leerverkäufe jener Währung in welcher das aus unternehmerischen Gründen eingegangene Risikodurchgeführt wurde, eine weitere wichtige Funktion (auch "Hedging" genannt).

Neben dem mikroökonomischen Aspekt, gibt es jedoch auch den Makroökonomischen. Da der Handel mit den meisten Währungen keinen Einschränkungen unterliegt, wird mitunter auch offensiv auf Kursentwicklungen zwischen zwei Währungspaaren spekuliert. Häufig werden auch besonders eklatante Zinsdifferenzen ausgenutzt und diese in Folge der Kapitalströme über die Zeit auch oftmals vermindert. In Niedrigzinswährungen wie beispielsweise der japanische Yen werden daher häufig Kredite aufgenommen und mit dem Gegenwert in Hochzinswährungen wie der türkischen Lira investiert. Bleibt der Wechselkurs über den Rückzahlungszeitraum unverändert ergibt das ein äußerst lohnendes Geschäft, wertet die Kreditwährung auf Grund eines sich verändernden Marktumfeldes auf, endet dies für den Investor häufig problematisch (siehe Schweizer Franken Kreditnehmer).

Der internationale Währungsmarkt ist durch die Anzahl und das Volumen der Transaktionen in Summe der bei weitem größte und liquideste Markt der Welt. Das wiederum erlaubt Transaktionen mit einer bis zu 100-fachen Menge an Kapital, gemessen an den Eigenmitteln auf einem Broker-Konto. Laut Bank für internationalen Zahlungsausgleich beträgt das weltweite Handelsvolumen zum April 2013 5,3 Billionen pro Tag [7] - Der Handel findet mit Ausnahme von Wochenenden rund um die Uhr statt.

2.3.5. Versicherungsmärkte

Bereits vor der globalen Finanzkrise war das Management von Risiko ein zentrales Thema bei sämtlichen großen Kapitalmarktakteuren. Versicherungsmärkte dienen als Proxy zum Transfer des Risikos eines möglichen Kapitalverlusts gegen eine Gebühr zwischen zwei Handelspartnern. Die Absicherung von Risiken geschieht zwar auch im Zuge der Handelstätigkeit auf anderen Märkten, da es jedoch eine extrem große Bandbreite an Finanzinstrumenten und Möglichkeiten zur maßgeschneiderten Absicherung von Risiken aller Art gibt, lässt sich diese Variation des zumeist dezentralen Handels ebenfalls als eigene Kategorie führen.

Ein besonders häufig gehandeltes und auch oftmals in der Öffentlichkeit kritisiertes Instrument ist jenes des "Credit-Default-Swaps" - einer Kreditausfallsversicherung. Gegen eine Initialprämie und oftmals auch eine laufende Gebühr kann die Insolvenz eines Kreditnehmers abgesichert werden. Die Höhe der Prämie des Kreditversicherers hängt direkt von der vermuteten Ausfallswahrscheinlichkeit des Kreditnehmers ab. Ändert sich dessen Bonität, erhöht sich auch der Wert der Versicherung welche sich wiederum weiterverkaufen lässt. Ein wichtiger Kritikpunkt ist der, dass auch jene Versicherungsgeschäfte eingehen können, welche das zu Grunde liegende Risiko gar nicht tragen - daher ohne weiteren Konnex eine Wette auf den Kreditausfall einer beliebigen Entität eingehen können. Das Volumen von im Umlauf befindlichen Kreditausfallsversicherungen übersteigt dadurch oftmals das zugrunde liegende Volumen der Basisgeschäfte. Dadurch kann wiederum die Pleite einer relativ unbedeutenden Entität Auswirkungen und Vermögensverluste nach sich ziehen, welche den eigentlich durch die Insolvenz erzeugten Schäden bei direkten und indirekten Geschäftspartnern um ein hundertfaches übersteigt.

Kann also beispielsweise durch einen Kreditausfall ein Volumen von 100 Millionen Euro nicht mehr bedient werden, tritt ein Kreditereignis ein welches die Besitzer der Ausfallsversicherungen zu Ersatzzahlungen verpflichtet. Das Volumen von weltweit schlagend werdenden Versicherungen welche aus Gründen der Spekulation eingegangen worden sind, kann dabei 100 Milliarden Euro ausmachen. Rasch wird dabei aus einem kleinen Ausfall ein systembedrohendes Ereignis welches die Versicherer ebenfalls in die Insolvenz ziehen würde. Das letzte in dieser Hinsicht bedrohliche Ereignis war Griechenland im Jahr 2011 - auch hier wurden exzessiv CDS, vor allem auch von europäischen Banken, abgeschlossen. Durch diese Art des "Derivatemissbrauchs" zu rein spekulativen Zwecken werden Versicherungen also oftmals auch als "Brandbeschleuniger" oder "finanzielle Massenvernichtungswaffen" tituliert.

Mit Hilfe der beschriebenen Märkte kann nun ein breit diversifiziertes Portfolio an Anlagewerten zusammengestellt werden. Die optimale Zusammenstellung und Gewichtung dieser Werte über die Zeit, ist sicherlich eines der am intensivsten diskutierten und erforschten Gebiete der Finanzwirtschaft. In weiterer Folge soll auf die grundlegendsten klassischen Kapitalmarkttheorien näher eingegangen werden und einige Grundbegriffe der

Kapitalanlage geklärt werden die für das Verständnis der weiterführenden Kapitel jedenfalls notwendig sind.

2.4. Markteffizienz und die moderne Portfoliotheorie

Mit Hilfe der klassischen Kapitalmarkttheorien konnten erstmals seit Beginn der Aufzeichnung von Wertschwankungen von Vermögensgütern, grundsätzliche Aussagen über die Marktdynamik und seine Beschaffenheit getätigt werden. Neben den wichtigsten Thesen sollen einige wenige Grundbegriffe erläutert werden, sowie im wissenschaftlichen Diskurs geäußerte Kritik der gängigen Theorien.

2.4.1. Markteffizienzhypothese

Eine der ersten Anwendungsgebiete von Computern war die Analyse ökonomischer Zeitreihen zur Entdeckung von Regelmäßigkeiten und Erhöhung der Voraussagekraft wirtschaftlicher Modelle. Auf Grund der besonders hohen Datenqualität und deren vollständigen Verfügbarkeit war der naheliegendste Analysegegenstand die Bewegung der Aktienmärkte. Maurice Kandall hat 1953 in einer der ersten Analysen dieser Art festgestellt, dass es keinerlei vorhersagbare Muster innerhalb der Preisbewegungen gibt. Diese scheinen also vollkommen zufällig zu sein, da sie zu jedem Zeitpunkt mit der gleichen Wahrscheinlichkeit unabhängig von ihrer historischen Rendite, nach oben wie nach unten bewegen können (auch "random walk" genannt). Die ersten Interpretationen dieser Erkenntnis legten daher auch eine vollkommene Irrationalität und eine rein massenpsychologische Steuerung der Märkte nahe. In weiterer Folge wurden die Ergebnisse jedoch völlig uminterpretiert - zufällige Preisbewegungen müssten einen nahezu perfekt effizienten Markt andeuten, da es keine systematische Möglichkeit gibt aus ihnen Profit zu schlagen. Sämtliche Informationen über die nächste Zukunft und die Gegenwart sind derart eingepreist, dass keine Möglichkeit besteht überproportionalen Profit zu generieren. Würde eine bestimmte Systematik das zulassen, würde diese Möglichkeit innerhalb kürzester Zeit ebenso eingepreist sein. Eine immerzu richtige Prognose über eine in T+1 eintretende Preissteigerung X würde die Preissteigerung bei vollkommener Informationstransparenz daher sofort in T eintreten lassen - da auf Basis der 100%-Prognosefähigkeit eines Modells kein Verkäufer bereits zum Zeitpunkt T einen geringeren Preis akzeptieren. Ein populärer Verfechter dieser Theorie ist der Nobelpreisträger Eugene Fama, welcher sich in

anerkannten Arbeiten (siehe [8] und 31 Jahre später [9]) erst streng und später relativierend für die Markteffizienzhypothese ausspricht.

Preisbewegungen entstehen dieser Theorie nach also ausschließlich durch vollkommen neu verfügbare Informationen die wiederum durch unvorhersehbare Ereignisse eintreten. Da die Basis der Preisbewegung also unvorhersehbar ist, muss es die Preisentwicklung selbst folgerichtig auch sein. [10]

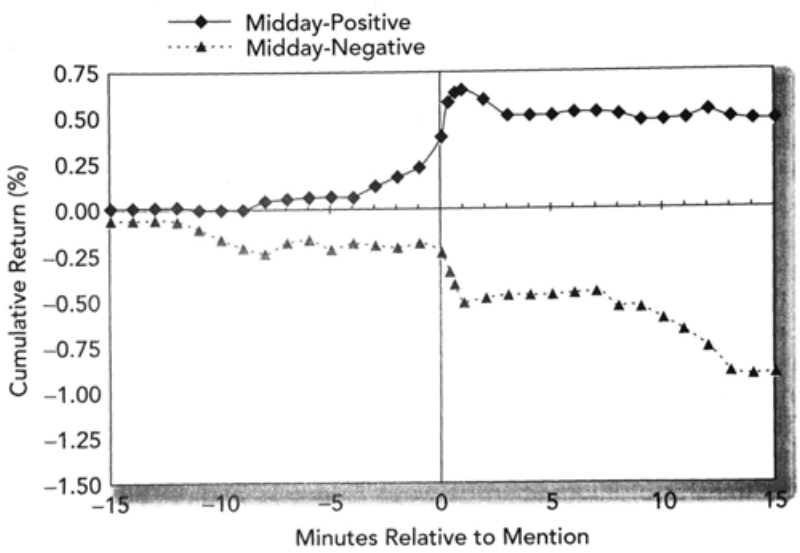

Abbildung 3 - Aktienpreisentwicklung vor und nach CNBC-Reports [9]

Obenstehende Grafik zeigt die kumulative Rendite eines Wertpapiers 15 Minuten vor bis 15 Minuten nachdem auf CNBC positive bzw. negative Nachrichten zu einem Unternehmen veröffentlicht wurden. Positive Nachrichten sind im Schnitt innerhalb der ersten 5 Minuten verarbeitet, negative innerhalb der ersten 12 Minuten. Zu bedenken ist jedenfalls, dass die Studie 2002 durchgeführt wurde - die Geschwindigkeit der Marktadaption wird auch auf Grund automatischer Handelssysteme bis dato weiter vervielfacht haben. [11]

Die Literatur unterscheidet zwischen verschiedenen Stärkegraden der Markteffizienzhypothese, welche sich primär durch die Definition der vollkommenen Informationstransparenz unterscheiden [2 S. 348-349]:

a. Schwache Form: Sämtliche vergangenen Informationen sind bereits im aktuellen Preis enthalten - eine Analyse historischer Preisbewegungen besitzt daher keinerlei Voraussagekraft für die zukünftige Preisentwicklung. Jegliche technische Analyse (siehe nächstes Kapitel) wäre daher vollkommen unergiebig.

b. Mittelstarke Form: Neben den in der schwachen Form erwähnten Voraussetzungen gilt zusätzlich, dass jede allgemein verfügbare Information zum finanziellen Status der Firma sowie sämtliche Gewinn- und Umsatzprognosen ihrer Produktlinie und der Qualität des Managements im aktuellen Preis enthalten ist. Das würde neben der technischen Analyse auch jegliche Fundamentalanalyse (Erläuterungen dazu im nächsten Kapitel) sinnlos machen.

c. Starke Form: Diese Variante geht davon aus, dass zusätzlich auch sämtliche Informationen die nur Firmeninsidern zur Verfügung stehen in aktuellen Aktienpreisen enthalten sind. In der Literatur wird diese Variante der Markteffizienzhypothese aber weitflächig abgelehnt, da die ständige Aktivität von "Insider-Trading" welches trotz des hohen Risikos durchgeführt wird, auf ein lukratives Geschäft hindeutet.

2.4.2. Moderne Portfoliotheorie (MPT)

Aus dem Geiste der EMH lässt sich die sogenannte „Modern Portfolio Theory“ (MPT) ableiten. Die MPT zeigt im Allgemeinen mathematisch den Wert des Konzepts der Diversifikation und gibt dem Investor die Möglichkeit per Asset-Allocation bei gegebenem Risiko (=Volatilität) eine maximal mögliche Rendite (und umgekehrt) zu erreichen.

Die MPT geht dabei von folgenden Annahmen aus:

- Die Asset-Rendite folgt einer Normalverteilung
- Das Risiko der Assets entspricht der Standardabweichung der Rendite
- Ein Portfolio entspricht einer gewichteten Zusammenstellung aus Einzelwerten – die Portfoliorendite ist daher auch die Summe der gewichteten Einzelwertrenditen

Jedes Asset welches historisch nicht vollkommen positiv korreliert mit jenem Portfolio ist zu dem es hinzugefügt wird, senkt das Gesamtrisiko des Portfolios. Weiters wird das Risiko ausschließlich über die Volatilität eines Wertes bestimmt, jeglicher Zugewinn bei der erwarteten Rendite wird also durch eine entsprechend höhere Schwankungsbreite des Assets „erkauft“.

In Summe wird durch die getroffenen Annahmen und die Einfachheit des Modells impliziert, dass Investoren vollkommen rational und Märkte effizient sind. Das zugehörige

mathematische Modell soll in weiterer Folge kurz erläutert werden, da es für das Verständnis der weiteren Arbeit jedenfalls notwendig ist.

Die erwartete Rendite ist definiert als

$$\mathrm{E}(R_p) = \sum_i w_i \, \mathrm{E}(R_i)$$

Wobei $R_p$der Rendite des Portfolios entspricht, $R_i$der Rendite des Assets i und $w_i$die Gewichtung des Assets i entspricht.

Die Portfoliovarianz der Gesamtrendite lautet allgemein

$$\sigma_p^2 = \sum_i \sum_j w_i w_j \sigma_i \sigma_j \rho_{ij}$$

Wobei ρ_{ij} der Korrelationskoeffizient zwischen den Returns von Asset i und j ist. Ist $i = j$ gilt daher $\rho_{ij} = 1$.

Die im Finanzbereich allgemein gängigere Portfoliovolatilität ist lediglich die Wurzelder Varianz.

$$\sigma_p = \sqrt{\sigma_p^2}$$

Hält ein Investor nun Werte welche nicht perfekt korreliert sind, wird dadurch automatisch die Volatilität und damit laut Theorie auch das Risiko reduziert. Für den Korrelationskoeffizienten gilt daher $-1 \leq \rho_{ij} < 1$ und für die Summe der Gewichte aller Assets im Portfolio $\sum_{i=1}^{m} w_i = 1$.

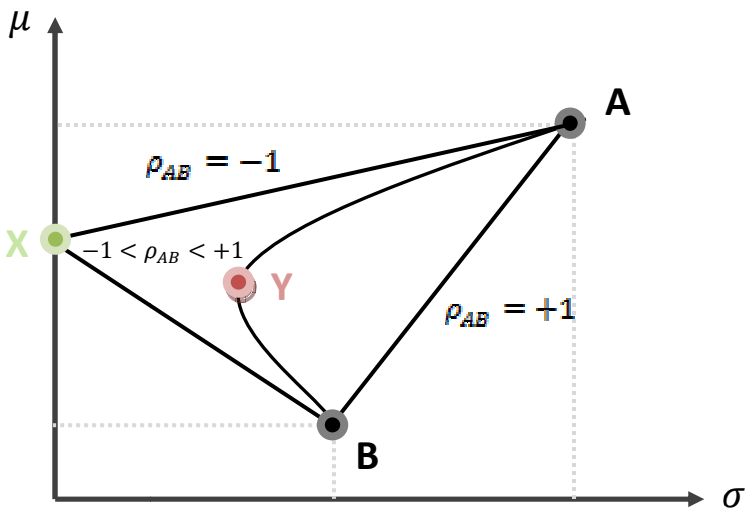

Abbildung 4 - Risiko/Rendite Profil eines simplen Portfolios [15 S. 11]

Oben stehende Grafik zeigt beispielhaft mögliche Portfoliokombinationen aus lediglich zwei verschiedenen Assets (A, B). Die X-Achse bildet die Volatilität ab, die Y-Achse den erwarteten Return. Beide Assets besitzen für sich genommen individuelle Risk/Return-Profile. Die abgebildete Hyperbel zeigt das Profil sämtlicher möglichen Portfolios bei variierender Gewichtung der Einzelwerte. Die Kanten des Dreiecks stellen jeweils das Ergebnis von verschiedenen Kombinationen dar, unter der Annahme die Werte wären jeweils perfekt positiv oder perfekt negativ korreliert – unter reellen Bedingungen, ergibt sich allerdings eine „efficient frontier“ auf welcher ein Punkt (Y) existiert, der rein rechnerisch (daher historisch) das optimale Verhältnis zwischen Risiko und Rendite unter den im Modell gegebenen Investitionsmöglichkeiten (A, B) ergibt.

Auf dieser Basis wird in der einfachsten Form der MPT ein effizientes Portfolio errechnet bzw. die Gewichtung für ein neues Asset zu einem bestehenden Portfolio ermittelt. Dabei kann auch B für ein gesamtes Portfolio stehen und A für einen neuen Einzelwert. [12]

2.4.3. Marktanomalien und Kritik

Auf Grund der Vielzahl an recht systematisch auftretenden Marktanomalien sowie massive Über- oder Untertreibungen die in den letzten Jahrzehnten beobachtet werden konnten, gerät die EMH (efficient-market hypothesis) zunehmend in Kritik. In weiterer Folge sollen einige intensiv diskutierte Marktanomalien kurz beschrieben werden, deren Existenz laut Markteffizienzhypothese und den daraus resultierenden Bedingungen nicht denkbar wäre.

a. „Low-Volatility“-Anomalie: Eine seit bereits 85 Jahren existierende Anomalie in US-Märkten, welche eine signifikant höhere Rendite für Portfolios aus Einzeltiteln mit besonders geringer Volatilität ausweist. Die risikoadjustierte Rendite liegt also deutlich über jener von Werten mit hoher Volatilität was der Annahme effizienter Märkte vollkommen entgegenläuft. Das würde bedeuten, dass die langfristige Rendite umso höher ausfällt, umso geringer das eingegangene Risiko liegt. Nichtsdestotrotz handelt es sich hierbei um einen von wenigen Effekten, welcher mehrfach in den verschiedensten Märkten bestätigt wurde. [13] [14]

b. „Small-Cap“-Effekt: Es zeigt sich, dass der breite Markt von Nebenwerten und Unternehmen mit geringer Marktkapitalisierung konsequent, risikobereinigt outperformed wird. Als fundamentale Gründe hierfür können die höhere Innovationskraft (im Speziellen jene die auch am Markt ankommt bzw. verwertet wird) und Dynamik kleinerer Unternehmen dienen. Denkbar ist zudem, dass eine weniger effiziente Preisbildung auf Grund des geringen Interesses durch Banken und Analysehäuser stattfindet. [15]

c. „Value“-Effekt: Unternehmen mit besonders niedrigem P/E-Ratio (=Verhältnis von Preis der Aktie zu Gewinn pro Aktie) und/oder hohem P/B-Ratio (= Verhältnis von Preis der Aktie zu Buchwerten pro Aktie) erscheinen im Verhältnis zum Marktpreis als besonders Werthaltig. In Summe erbringen diese risikobereinigt ebenso höhere Returns als der Gesamtmarkt. [16]

d. "Momentum"-Effekt: Momentum beschreibt die fortgeführte Eigendynamik einer Preisbewegung die sich auf die relative Stärke bzw. Schwäche von Aktien bezieht. Es zeigt sich, dass Aktien mit historisch hoher Performance konsequent jene mit historisch schwacher Entwicklung "outperformen". Die durchschnittliche Überrendite im Vergleich zu einem gleichgewichteten Index aus Einzeltiteln welche durch den Kauf der starken und den (Leer)-Verkauf der schwachen Werten generiert werden kann, liegt empirisch betrachtet bei 1% pro Monat. Diese enorm hohe Überrendite die langfristig bereits durch eine verhältnismäßig triviale Strategie gesichert werden kann, widerspricht der Markteffizienztheorie vollkommen. Tatsächlich ist der "Momentum"-Effekt die wahrscheinlich anerkannteste Marktanomalie und wird von institutionellen Investoren auch aktiv genutzt. [17] [18]

e. Kalender-Effekte: Beschreibt sämtliche zeitlich getriebenen, wiederkehrenden Marktanomalien. Hierbei kann es sich um Tagesereignisse oder aber auch ganze Zyklen wie den 4-jährigen Präsidentschaftszyklus handeln. Anomalien dieser Art gäben einem Investor die Gelegenheit sich zeitlich vor dem Eintritt des Effekts zu positionieren. Kalendereffekte sind in der Literatur jedoch äußerst strittig –in einem 2001 publizierten Paper spricht Sullivan dieser Art der Anomalien ihre statistische Signifikanz ab. [19]

Einige der genannten Effekte sind zwar durchaus strittig, andere jedoch zumindest nicht widerlegbar. Doch bereits die Menge unterschiedlichster Effekte deutet darauf hin, dass die MPT in ihrer ursprünglichen Darstellung nicht haltbar ist. Insbesondere die Annahme, dass Renditen normalverteilt oder auch nur symmetrisch verteilt seien erweist sich als nicht haltbar. Weiters basiert die Errechnung historischer Korrelationen auf der Prämisse, dass sich diese niemals gemeinschaftlich derart verändern können, sodass sich sämtliche Werte hochgradig korreliert bewegen. Jedoch brechen im Speziellen in Krisenzeiten Zusammenhänge welche möglicherweise über Jahrzehnte hinweg bestanden haben zumeist konsequent weg.

Neben den zuvor beschriebenen, diffizileren Anomalien, gab es stets medienpräsente und offensichtliche Übertreibungen (2000: Dotcom-Blase, 2007: Subprime-Blase) die nur von uninformierten Marktteilnehmern getrieben werden konnten. Es liegt also nahe zu vermuten dass die Markteffizienz mit der Qualität der Marktteilnehmer steht und fällt. Seit dem ersten Aufkommen der Theorie wurde eine große Anzahl empirischer Studien durchgeführt, welche in Summe durchwachsene Ergebnisse lieferten – die starke Form der EMH konnte jedoch nie mit Sicherheit bestätigt werden.

Im Zuge der immer lauter werdenden Kritik hat sich ein neuer Wissenschaftszweig formiert, nämlich jener der „Behavioural Finance“. Dabei wird versucht klassische Anomalien mittels bekannter menschlicher kognitiver Befangenheiten bzw. Tendenzen zu erklären. Beispiele dafür sind klassische Überreaktionen optimistischer und pessimistischer Art sowie Fehler in der Schlussfolgerung und Informationsverarbeitung des menschlichen Gehirns.

3. Portfolio- und Assetmanagement

Die aktive Verwaltung von Vermögen beinhaltet zumeist den Aufbau eines ausgewogenen Portfolios aus diversen Vermögensgegenständen zur Erreichung eines spezifizierten Investitionsziels. Dieses Ziel sieht immer eine zu erwartende Rendite, manchmal auch einen Richtwert zur maximal tolerierbaren Volatilität unter der die Renditevorgabe erzielt werden kann, vor. Der Leistungsumfang eines typischen aktiven Vermögensverwalters beinhaltet neben der Renditegenerierung ebenfalls die Erstellung von Finanzberichten zur aktuellen finanziellen Situation des Kunden, Marktanalysen und Monitorings, sowie die Planung und Kommunikation der zukünftigen Portfolioausrichtung und „Asset-Allocation".

In weiterer Folge sollen einige zentrale Begriffe erläutert werden, welche bei der Entscheidung von Investitionen je nach Ausrichtung und Investitionsprinzipien der Vermögensverwaltung eine Rolle spielen können. Abschließend soll die Möglichkeit technischer Hilfsmittel, insbesondere mittels genetischer Algorithmen, unter dem Aspekt der Markteffizienz erläutert werden.

3.1. Fundamentalanalyse

Die Fundamentalanalyse beschäftigt sich mit den einem Unternehmen zu Grund liegenden Unternehmensdaten und deren Prognose. Durch die Untersuchung vergangener Gewinne und Bilanzen, sowie einer detaillierten Auseinandersetzung des aktuellen und zukünftigen ökonomischen Umfelds soll der aktuelle Firmenwert und das Entwicklungspotential approximiert werden. Weitere wichtige Bestandteile dieser Art der Analyse sind die qualitative Bewertung des Managements sowie die Zukunftsaussichten der Branche und die Verortung des Unternehmens innerhalb dieser Branche.

Der wichtigste Aspekt dabei ist es, jenen Mehrwert eines Unternehmens zu erkennen dem der Rest des Marktes verborgen ist. Es ist auch bei der Fundamentalanalyse lediglich möglich an der Differenz zwischen eigener Meinung und der Marktmeinung überproportionale Gewinne zu erzielen. Eine ausgezeichnete Analyse auf Basis der öffentlich zugänglichen Zahlen ist üblicherweise also nicht ausreichend - ein wesentlicher Mehrwert könnte daher beispielsweise eine gute persönliche Kenntnis und korrekte menschliche Beurteilung der größten Anteilseigner und wichtigsten Manager darstellen. [1 S. 231 ff.]

Ein äußerst bekannter Vertreter der Fundamentalanalyse ist Warren Buffet, einer der erfolgreichsten "Value"-Investoren der Welt mit einer erzielten langfristigen Durchschnittsrendite von fast 20% pro Jahr von 1965 bis 2012 [20].

3.2. Technische Analyse

Mit Hilfe der technischen Analyse werden Preiskurven historisch untersucht. Aus dieser Analyse sollen Rückschlüsse für zukünftige Entwicklungen gezogen werden. Betrachtet werden oftmals wiederkehrende „Chartmuster“ die sich sowohl über Einzelwerte als auch Indizes der unterschiedlichen Asset-Klassen gleichen. Eine der bekanntesten Komponenten der technischen Analyse ist das Konzept von Widerständen und Unterstützungen. Dabei geht man davon aus, das ein weiterer Preisverfall bei Widerstandslinien unwahrscheinlich und ein Preisanstieg bei Unterstützungslinien wahrscheinlich ist. Unten stehende Abbildung soll das Konzept etwas verdeutlichen.

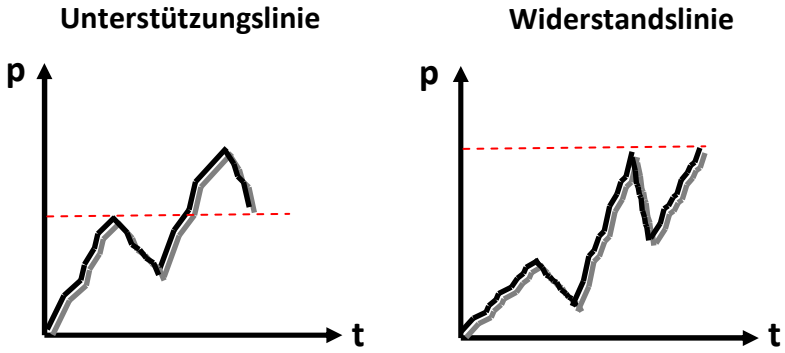

Abbildung 5 - Unterstützungslinie und Widerstandslinie

Eine interessante Fragestellung ist jene ob sich jede konsequent ausnutzbare Strategie nicht automatisch von selbst überholt. Sobald eine nutzbare Regelmäßigkeit von zu vielen Marktteilnehmern entdeckt und gehandelt wird, löst sich diese von selbst wieder auf, da zu niedrige oder zu hohe Preise eines Moments durch vermehrte Käufe oder Verkäufe wieder geglättet werden.

Im weiteren Sinne wird aber auch eine Art der markttechnischen Analyse verstanden, welche über die der reinen Chartanalyse hinausgeht. Hierbei fließen weitere Indikatoren wie die Volatilität und auch der Verlauf der gehandelten Volumina in die Analyse mit ein. Ein weiterer Versuch der systematischen Erweiterung des technischen Analyseframeworks stellen zudem Indikatoren dar, welche versuchen die aktuelle Stimmungslage verschiedener Anlegertypen (z.B. Privatinvestoren, institutionelle Anleger) durch Befragungen einzufangen

und zu bewerten. Im Bereich der „Markttechnik“ wird also durch die Kombination der einzelnen Werkzeuge versucht, ein Stimmungsbild des Marktes zu zeichnen und zukünftige Entwicklungen möglicherweise abschätzen zu können. [1 S. 272 ff.]

3.3. Asset-Allokation

Die Entscheidungsfindung darüber wie Kapital in Zukunft allokiert werden soll, kann mitunter recht komplex werden – dabei werden klassischerweise sowohl makroökonomische und qualitative Aspekte wie auch eine quantitative Analysen in Betracht gezogen.

Im vorherigen Kapitel wurden die einzelnen Asset-Klassen kurz erläutert – im Rahmen aktiver Vermögensverwaltung wird die Gewichtung der Klassen im Portfolio jeweils dynamisch an die eigene Markterwartung angepasst. Durch den Versuch der Höhergewichtung von „überverkauften“ Asset-Klassen, kann die langfristige Portfolioperformance signifikant beeinflusst werden. Überrendite im Vergleich zum Gesamtmarkt, wird fast ausschließlich durch eine korrekte Umgewichtung der Asset-Klassen erzeugt – die gezielte Auswahl von Einzelwerten innerhalb der Asset-Klassen trägt zumeist kaum zur Überrendite bei.

Ein Beispiel hierfür wäre eine Umschichtung erstklassiger Staatsanleihen 2008/2009 in Aktien, riskante High-Yield Anleihen sowie „Mortgage Backed Securities“. Überall wo historisch gesehen extrem hohe Volatilität zu finden ist sowie eine extrem geringe Bewertung der dahinterliegenden Fundamentalwerte vermutet werden kann, sollte durch gezielte Investition langfristig Überrendite erzielbar sein. Dies gilt auch, wie am eben genannten Beispiel aufgezeigt, innerhalb von Asset-Klassen (d.h. erstklassige Staatsanleihen vs. Unternehmensanleihen bzw. High-Yield Anleihen).

3.4. Passives vs. aktives Management

Nach bisherigen Erkenntnissen hängt die Frage nach Sinn oder Unsinn aktiven Portfoliomanagements stark von der individuellen Situation des jeweiligen Anlegers ab. Die eigenständige Selektion von Einzeltiteln oder einer Verschiebung seines Vermögens zwischen den Asset-Klassen sollte sich nach der Größe des verwalteten Kapitals richten. Im Zuge des vorangegangenen Kapitels wurde eine Reihe von Anomalien beschrieben, welche durchaus auf gewisse systematische Marktineffizienzen hindeuten. Ebenso in Zweifel

gezogen werden kann die vollkommene Informationstransparenz – die Kosten des Aufwands welcher durch die Beschaffung der Rendite zuträglicher Informationen entsteht, müssen stets im Verhältnis zum eingesetzten Kapital stehen. Sind also beispielsweise 5 Mrd. Euro unter Verwaltung und lässt sich mit hohem Kostenaufwand relativ zuverlässig eine Zusatzrendite von 0,1% erwirtschaften, so sind Kosten von ca. 5.000.000 Euro gerechtfertigt. Die erworbenen Informationen werden zumeist exklusiv genutzt oder zu hohen Kosten weiterverkauft – die Grenze zwischen Insider Know-How und legalem Wissen fließen dabei zuweilen ineinander über. Dies veranschaulicht auch deutlich einen Teil des kompetitiven Nachteils von Kleinanlegern. Ein weiterer Vorteil aktiven Managements kann in einer komplexen Steuergesetzgebung liegen. In Abhängigkeit nationaler Gesetze kann es sein, dass bestimmte Investitionen steuerlich massiv benachteiligt werden. Durch die Höhe der verwalteten Volumina können Steuerrückforderungen jedoch effizient und daher ausreichend lukrativ durchgeführt werden (z.B. Rückholung von Quellensteuern). Durch dauerhaft steueroptimales Handeln, werden durch einen sich über die Jahre verstärkenden Zinseszinseffekt signifikante Überschüsse erzielt. Eine der größten und bekanntesten aktiven Vermögensverwaltungen ist Black Rock.

Passives Portfoliomanagement ist vor allem seit dem Jahr 2000 immer stärker in den Fokus gerückt. Das Handelsvolumina mit sogenannten „Exchange Traded Funds“ (ETFs) hat sich im letzten Jahrzehnt vervielfacht. Hinter den meisten ETFs steht zumeist die Replikation eines allgemein bekannten Index. Hierbei werden Investitionsentscheidungen nicht auf Basis von individuellen Analysen getroffen, sondern aus rein technischen Gründen. Das Ziel ist eine möglichst detailgetreue Abbildung z.B. des S&P500 – dafür müssen in regelmäßigen Abständen Wertpapiere verkauft und angekauft werden um die vorgegebene Gewichtung der Einzelwerte möglichst gut zu treffen. Die Kosten solcher Fonds sind auf Grund des fehlenden Managements besonders gering und erlauben auch Kleininvestoren breit diversifizierte Geldanlage. Zu den weltweit größten Anbietern zählen unter anderem IShares mit Sitz in den USA und Lyxor mit Sitz in Frankreich.

Generell kann angenommen werden, dass die meisten Fondsmanager passive Fonds im gleichen Segment nach Kosten dauerhaft nicht schlagen können – jedoch existieren eine Reihe bekannter Größen die vor allem ein Faktum von anderen institutionellen Anlegern unterscheidet: vollkommene Unabhängigkeit. Letztlich unterstehen die meisten

Fondsmanager dem Druck die Performance einer konkreten „Benchmark“ schlagen und mindestens quartalsweise Rechenschaft über die in den letzten Monaten erzielte Rendite ablegen zu müssen. Jede größere Kontraktion des Marktes, zwingt Manager zu zyklischem Handeln – macht dieser den Versuch von antizyklischen Käufen die nicht innerhalb weniger Monate erfolgreich verlaufen, geht er das Risiko ein über zu lange Zeit Boden zur Konkurrenz oder Benchmark zu verlieren und damit seinen Job zu riskieren. Illiquiditätsprämien oder Risikoprämien zu lukrieren, welche sich möglicherweise erst im Laufe mehrerer Jahre manifestieren, ist in der Fondsindustrie nur schwer möglich. Die schwache Performance der meisten Fonds liegt daher eher an der Notwendigkeit rascher Ergebnisse und externem Druck.

Der Sinn oder Unsinn professionellen Portfolio-Managements, lässt sich aber nicht ausschließlich auf Basis roher Fakten und Zahlen entscheiden. Lässt man Vermögen verwalten, erhält man bei professioneller Verwaltung in regelmäßigen Sitzungen Einblick in Strategien und Einschätzung der aktuellen Marktsituation. Unbestritten ist, dass Exchange Traded Funds der einschlägigen Anbieter die kostengünstigste Variante darstellen – schafft man es allerdings nicht, selbstständig durch manchmal langjährige Tiefs hindurch zu tauchen ohne durch unüberlegtes Handeln Verluste zu realisieren, kann diese Art der „Bevormundung“ durch einen Fondsmanager hilfreich sein. Entfernt man den Faktor Mensch welcher hinter einer Investitionsentscheidung steht nicht aus der Gleichung, kann das Vorhandensein konkreter Ansprechpersonen individuell positive Effekte erzeugen. [2 S. 251 ff.]

3.5. Genetische Algorithmen und aktives Portfolio-Management

Angesichts des heutigen wissenschaftlichen Standes, kann nicht von einem effizienten Markt ausgegangen werden. Die Frage nach der Höhe der Markteffizienz, lässt sich jedoch nicht pauschal beantworten. Vor allem in Marktphasen mit Liquiditätsengpässen können Situationen in bestimmten Asset-Klassen entstehen, welche eine Unterbewertung einzelner Werte quasi erzwingen. Getrieben durch gesetzliche Rahmenbedingungen zur Liquiditätsquote (Stichwort BASEL), hohen Abflüssen von Privatpersonen aus Fonds und gesteigertem Misstrauen zwischen Finanzteilnehmern, bestehen nicht genügend liquide

Mittel um einen korrekten Preis zustande kommen zu lassen. Diese Strukturbrüche können durch quantitative Analysen gut auch als solche erkannt werden.

Die zunehmende Dynamik der weltweiten Finanzmärkte während der letzten Dekade stellt jedoch auch die klassischen analytischen Portfoliooptimierungsverfahren zunehmend in Frage. Durch die daraus resultierende erhöhte Variabilität externer Markteinflussfaktoren sowie deren über die Zeit veränderte Gewichtung, lassen sich nur schwierig zuverlässige Modelle auf Basis fixierter Einflussgrößen ermitteln. Im weiterer Folge soll eine Möglichkeit aufgezeigt werden, wie mit Hilfe innovativer Ansätze aus dem Bereich der Informatik ein erweiterbares Framework kreiert werden kann, welches bei der Konstruktion möglichst optimaler Portfolios sowie der Asset-Allokation hilfreich ist.

Insbesondere mittels eines aus fachlicher Sicht einfach erweiterbaren heuristischen Optimierungsverfahrens wie jenem genetischer Algorithmen, lässt sich eine Umgebung erzeugen mit welcher sich Thesen innerhalb eines abgeschlossenen Anlageuniversums einfach überprüfen und erweitern lassen. Entscheidend sind stets das Zusammenspiel und die Unterstützung technischer Werkzeuge und fundamentaler Aspekte um unterstützend auf die Entscheidungen des Asset-Managers zu wirken und sein "Bauchgefühl" zu objektivieren. Durch die quantitative Analyse historischer Daten können wertvolle Einblicke in die Dynamik der Zusammenhänge zwischen einzelnen Werten, Branchen oder ganzen Asset-Klassen gewonnen werden. Wesentlich ist allerdings immer das Wissen um die makroökonomischen und politischen Geschehnisse der quantitativ analysierten Zeiträume, da die Ergebnisse ohne den entsprechenden Kontext nicht richtig interpretiert werden können.

Im nächsten Kapitel sollen genetische Algorithmen als heuristisches Optimierungsverfahren vorgestellt und im Detail erläutert werden. Weiters soll ein Weg aufgezeigt werden, wie die Anwendung dieser Art des Optimierungsverfahrens in ein erweiterbares Analysetool integriert werden kann, welches für die Zwecke aktiver Vermögensverwaltung hilfreich ist.

4. Theorie und Praxis genetischer Algorithmen

4.1. Prinzipien und Grundlagen

Genetische Algorithmen sind der Klasse der evolutionären Algorithmen zuzuordnen welche als heuristisches Optimierungsverfahren bezeichnet werden. Diese werden vorzugsweise natürlichen Vorbildern nachempfunden und vor allem in jenen Bereichen verwendet in denen keine exakte analytische Lösung eines Problems möglich ist. Eine große Menge dieser Methoden stammt aus dem inzwischen populär gewordenen Forschungsbereich der künstlichen Intelligenz. Zu den Bekanntesten zählen hierbei die neuronalen Netze, Fuzzy Logic und evolutionäre Algorithmen. Letztere sind in ihrer Funktionsweise an den wesentlichsten Prinzipien der von Charles Darwin begründeten Theorie der evolutionären Biologie angelehnt. Der Gedanke des "Survival oft he fittest" liegt hier zu Grunde. Das technische Sprachgebären orientiert sich daher auch stark an den ursprünglichen Begrifflichkeiten.

Einen wichtigen Beitrag zur Popularität von genetischen Algorithmen im Speziellen lieferte John H. Holland mit [21], in welchem er versuchte Modelle zu entwickeln und darzulegen wie das Zusammenspiel von miteinander nichtlinear in Beziehung stehenden Faktoren einen Anpassungsprozess der umgebenden Elemente bewirken. Eine absolut exakte Definition eines genetischen Algorithmus kann nicht gegeben werden, lediglich die Funktionsweise und die darin typischerweise enthaltenen **Bausteine** definieren einen solchen:

Die Grundbausteine und möglichen Eigenschaften und Ausprägungen eines **Individuums** werden durch seine **DNS** definiert, welche wiederum aus den zur Verfügung stehenden **Genen** zusammengesetzt ist. Der **Phänotyp** entspricht dem äußeren Erscheinungsbild des Individuums welches durch seine DNS bestimmt wird. Die Menge aller vorhandenen Gene einer Spezies entspricht dem **Genom**. Der **Genotyp** ist die für jedes Individuum einzigartige Untermenge und Zusammensetzung des Genoms. Ein Allel entspricht einer konkreten Ausprägung eines Gens, welches sich wiederum an einer Stelle eines **Chromosoms** befindet. Das Chromosom (Genotyp) wird bei der modellhaften technischen Umsetzung oftmals auch direkt als das Individuum (Phänotyp) bezeichnet. Eine Menge von Individuen ist als **Population** definiert.

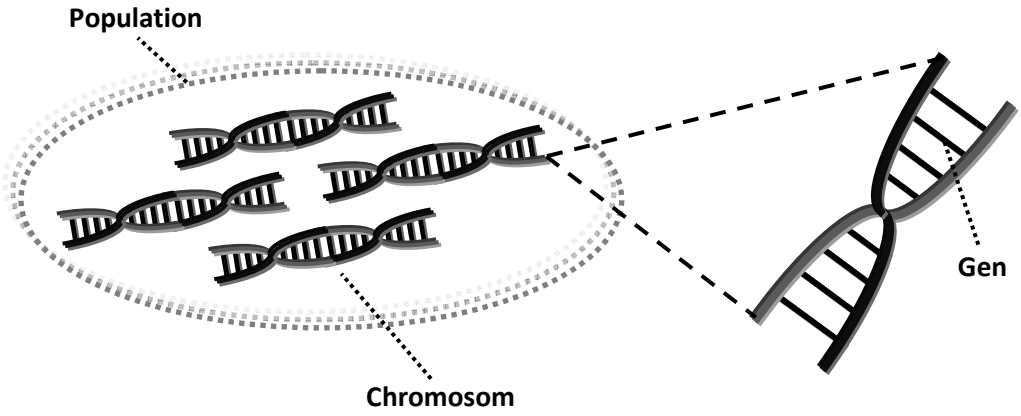

Abbildung 6 - Population, Chromosom, Gen

Eine beliebige Population muss sich ihrer laufend verändernden Umwelt schnellstmöglich anpassen um den wechselnden Bedingungen gerecht zu werden und überlebensfähig zu bleiben. Durch den Vorzug des Genoms von Individuen mit hoher Wettbewerbsfähigkeit unter gegebenen Parametern, überleben Primär jene Individuen welche die sogenannte "Fitness" der Gesamtpopulation erhöhen. Auf natürlichem Wege wird hier ein komplexes Optimierungsproblem unter veränderlichen Zelvorgaben gelöst. Durch den modellhaften Übertrag einer gegebenen Problemstellung in ein Basisformat, welches für klassische evolutionäre Mechanismen greifbar ist - ergeben sich oftmals interessante Möglichkeiten der Problemlösung und Optimierung.

Ein wesentlicher Vorteil genetischer Algorithmen ist, dass die Determinanten eines vorliegenden Problems nicht detailliert bekannt und ausspezifiziert sein müssen um zu einer zufriedenstellenden Lösung zu gelangen. Potentielle Lösungen werden über die sogenannte Fitnessfunktion ermittelt, welche jenen Teilbereich des Problems repräsentiert, der bestmöglich heuristisch gelöst werden soll. Im anschließend zu durchlaufenden, evolutionären Prozess sollte die Kombination guter Teillösungen - daher jene mit den besten Werten laut Fitnessfunktion - eine effiziente Gesamtlösung ergeben. Durch die Verwendung von Mutationen, also zufälligen Veränderungen des Lösungskandidaten in bestimmten Teilbereichen, soll die Wahrscheinlichkeit für das Auffinden eines globalen Optimums erhöht und daher eine bessere Gesamtlösung erzeugt werden. Eine vorzeitige Konvergenz der Population hin auf suboptimale Lösungen wird dadurch tendenziell vermieden. Eine Population besteht aus einer Menge von Chromosomen. Jedes Chromosom entspricht einer

Lösung. Ein Chromosom besteht wiederum aus einer Reihe von Allelen, welche die Eigenschaften des Chromosoms darstellen.

Der dahinterliegende Ablauf zur Lösung eines Problems mittels genetischer Algorithmen sieht unabhängig von der Aufgabenstellung immer grundsätzlich äquivalente Schritte vor:

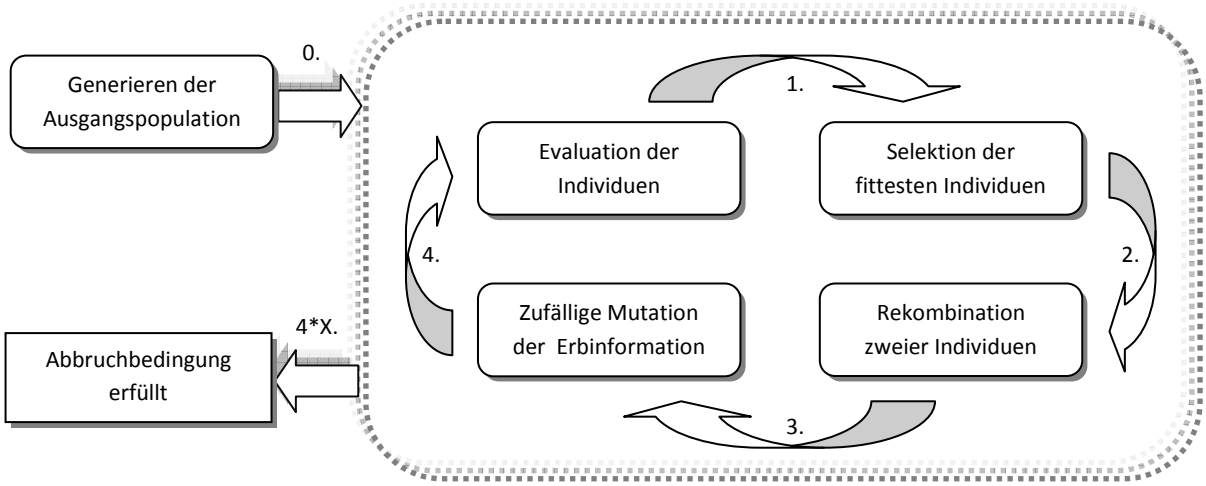

Abbildung 7 - Zyklus eines genetischen Algorithmus

Die Kernelemente des genetischen Algorithmus sind daher auf die Evaluation, Selektion, Mutation und Rekombination zu reduzieren. Diese Elemente werden solange ausgeführt bis ein vordefiniertes Abbruchkriterium erreich wurde, welches die aktuellste Lösung als vermutlich ausreichend optimal kennzeichnet.

In weiterer Folge sollen die wichtigsten Schritte zum besseren Verständnis grob schematisch beschrieben werden. Im folgenden Kapitel wird im Detail mit Hilfe konkreter Beispiele auf die dahinterliegenden Prozesse eingegangen. Noch detailliertere Beschreibungen zu den Erläuterungen in Unterkapitel 4.1 können sowohl aus [22] als auch aus [23] entnommen werden.

4.1.1. Evaluation

Entspricht der Bewertung der in der jeweils letzten Iteration generierten Population. Diese erfolgt anhand der vordefinierten Fitnessfunktion mit welcher die Wertigkeit des betrachteten Elements ermittelt werden soll. Beim Auffinden der kürzesten Routen in einem komplexen Routensystem, repräsentiert die Fitnessfunktion also lediglich die Summe der

benötigten Wegzeit der jeweiligen Teilstrecken. Jedes Allel entspräche einem Wegteil, das gesamte Chromosom ist daher eine mehr oder weniger optimale mögliche Lösung des Problems.

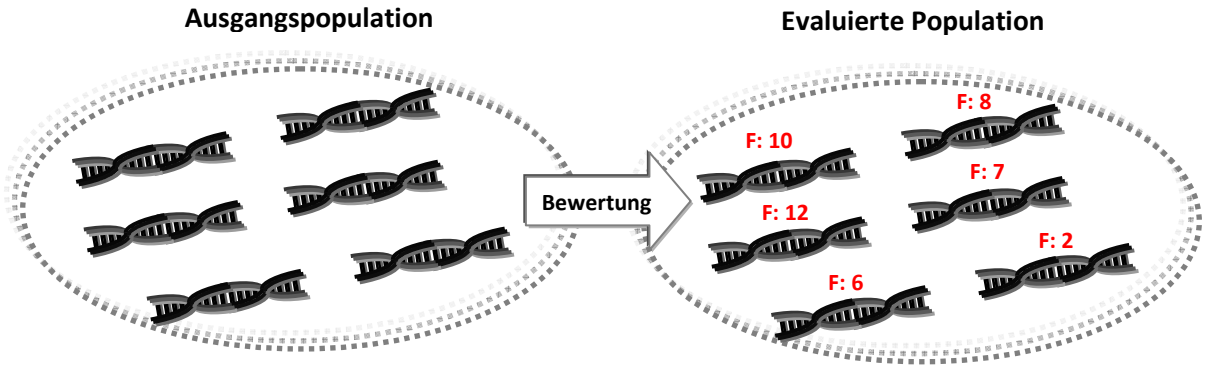

Abbildung 8 - Ausgangspopulation vs. evaluierte Population

Im einfachsten Fall ist die aus der Bewertung entsprungene Fitness eine natürliche Zahl mit Hilfe derer man die möglichen Lösungen reihen kann.

4.1.2. Selektion

Hier werden zwei geeignete Individuen (Vater- und Mutterindividuum) gewählt, aus denen zwei Kinderindividuen erzeugt werden sollen. Die konkrete Auswahl der Elternindividuen kann auf unterschiedliche Weise von Statten gehen. Eine Zufallsauswahl von Elternteilen innerhalb einer Subgruppe von Individuen mit hoher Fitness (bspw. die obersten 70% der Gesamtpopulation) repräsentiert eine valide Möglichkeit. Um die genetische Vielfalt zu bewahren und ähnlich wie bei der Mutation eine zu frühe Optimierung auf ein lokales Optimum zu verhindern, ist es legitim Chromosomen mit geringer Fitness mit vordefinierter Wahrscheinlichkeit ebenfalls in die Neupopulation aufzunehmen.

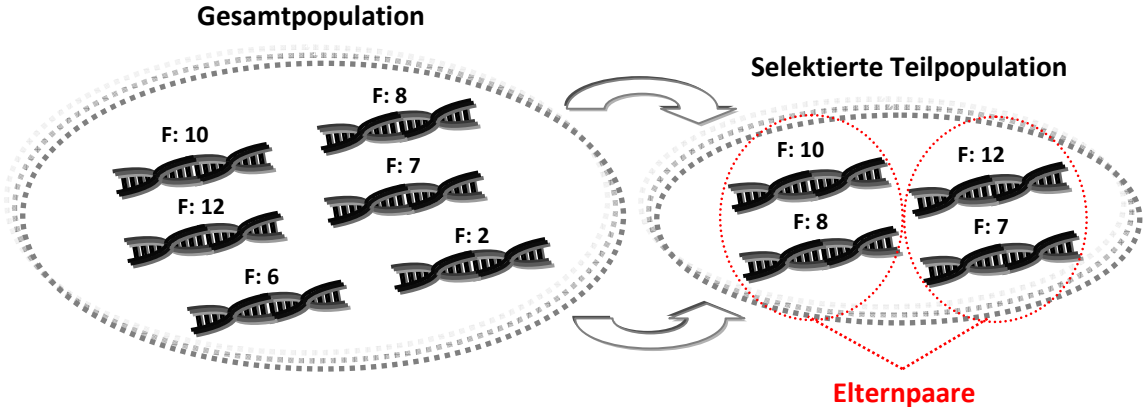

Abbildung 9 - Selektion der Individuen

Die Auswahl der Elternpaare innerhalb der vorselektierten Teilpopulation kann zufällig oder aber auch über deren Ähnlichkeit erfolgen.

4.1.3. Rekombination

Dieser Schritt wird auch als „Crossover“ bezeichnet und ist ein zentraler Baustein bei der Entwicklung des genetischen Algorithmus. Hierbei können völlig neue Merkmalskombinationen durch den Austausch von Allelen zweier Elternchromosomen entstehen. Die konkrete Implementierung des Austausches kann je nach Problemstellung auf unterschiedliche Weise erfolgen. Sowohl Ein- als auch Mehrpunktrekombinationen, das heißt der Austausch von Allelen oder Gruppen von Allelen an einer oder mehreren Stellen der Chromosomen. Zu beachten gilt, dass durch die Implementierung der Rekombination wiederum valide Kinderindividuen entstehen müssen, was durch entsprechende "Constraints" abgesichert werden muss.

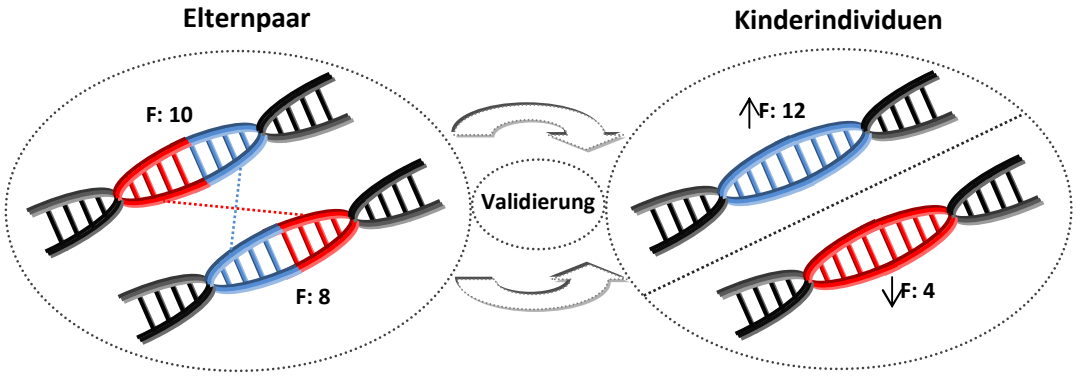

Abbildung 10 - Rekombination der Elternpaare

4.1.4. Mutation

Um eine zu frühe Optimierung hin auf ein lokales Optimum zu vermeiden, ist eine nicht zu unterschätzende Komponente jene der Mutation. Auf Basis einer vordefinierten Mutationswahrscheinlichkeit (P_m; 0 <= P_m<= 1.0) welche in den Steuerungsparametern des genetischen Algorithmus festgelegt wird, sollen einzelne Allele des Chromosoms zufällig verändert werden. Wie auch in der Natur können Mutationen Individuen schädigen oder aber auch deren Überlebensfähigkeit erhöhen. Bei der Implementierung muss wie auch bei der Rekombination darauf geachtet werden, dass ein für die weiteren Iterationen des Algorithmus gültiges Chromosom hinterlassen wird.

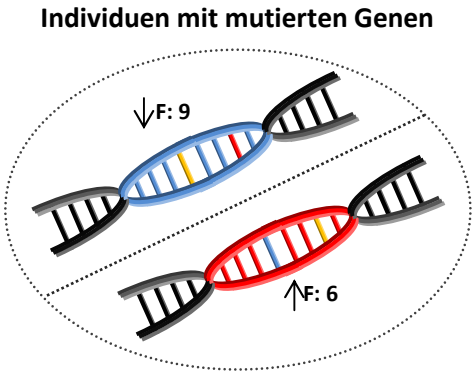

Abbildung 11 - Mutation der Individuen

Gemeinsam mit der Rekombination trägt die Mutation zur genetischen Vielfalt der neuen Population bei.

Die eigentliche Schwierigkeit bei der Lösung eines Problems mit genetischen Algorithmen besteht im Übertrag der Aufgabenstellung in eine für den Algorithmus verwertbare Form und der Bestimmung der Datenbasis anhand derer eine annähernd optimale Lösung ermittelt werden soll. Die Kodierung der Individuen sowie die Rekombination der Individuen stellen einen zentralen Problempunkt dar der überwunden werden muss. Liegt ein konkludentes Basiskonzept vor, ist der übrige Implementierungsaufwand verhältnismäßig gering. Da die Fitnessfunktion, anhand derer die Überlebensfähigkeit eines Individuums bewertet werden soll im Rahmen der Prozedur unproblematisch als eigenes „Modul“ erweitert werden kann, ist auch eine komplexe Adaption des Näherungsverfahrens sehr einfach möglich.

Aus diesem Grund werden oftmals auch deterministisch lösbare Probleme mittels heuristischer Verfahren gelöst. Im Speziellen beim Problem der Portfoliooptimierung auf Basis historischer Daten ist die einfache Möglichkeit der Erweiterung der Bewertungsfunktion in Kombination mit der externen Anpassung des genetischen Algorithmus mittels Startparameter besonders vorteilhaft. Durch die einfach handhabbare Dynamik des Systems können die sich in regelmäßigen Abständen verändernden externen Einflüsse auf den Finanzmarkt modellhaft abgebildet und über die Fitnessfunktion der Individuen eingebunden werden.

4.1.5. Abbruchbedingung

Da ein heuristisches Verfahren keine finale Lösung generiert und das exakte Optimum unbekannt ist, muss der Algorithmus durch vorher spezifizierte Abbruchkriterien beendet

werden. Hierfür gibt es unterschiedliche Ansätze welche in Abhängigkeit vom Optimierungsverhalten des Algorithmus jeweils unterschiedlich gewählt werden müssen. Ist die Abbruchbedingung erfüllt, ist das zuletzt erzeugte Ergebnis gültig und das stärkste Individuum der aktuellen Population die beste erzeugte Lösung.

In den folgenden Kapiteln soll näher auf die hier überblicksmäßig vorgestellten Operatoren an Hand kleiner Beispiele und konkreter formaler Definitionen der Operatoren eingegangen werden. Für jeden der grob beschriebenen Abläufe gibt es unterschiedlichste Möglichkeiten der Implementierung und Kodierung.

4.2. Formalismen, Umsetzungsvarianten und Beispiele

Der Einsatz heuristischer Lösungsverfahren erfolgt im akademischen Regelfall ausschließlich bei Anwendungsproblemen welche nachweislich nicht berechenbar sind oder zumindest zu sein scheinen. Probleme dieser Art werden als **NP („Nondeterministicpolynomial") vollständig** beschrieben. Gängige Beispiele für diese Art von Fragestellungen wären:

- Wie kann vorhandener Platz bei gegebenen Gegenständen optimal genutzt werden wenn nicht alle Gegenstände Platz finden
- Wie können die Stundenpläne aller Schulklassen bei gegebenen Beschränkungen so gestaltet werden, dass möglichst wenige Lücken entstehen?
- Was ist die optimalste Strecke bei einer Rundreise durch mehrere verschiedene Städte?

Die Anzahl der möglichen Lösungen steigt exponentiell mit der Anzahl der Gegenstände, Schulklassen und Städte die Teil der Parameter des Problems sind.

Ein plakatives Beispiel stellt die erstgenannte Fragestellung dar, welche in der Informatik auch als sogenanntes „Rucksackproblem" bekannt ist. Für jeden Gegenstand muss entschieden werden ob dieser eingepackt oder draußen gelassen werden soll. Schon bei 100 Gegenständen unterschiedlicher Größe erhalten wir ein Entscheidungsproblem mit 2^{100} Möglichkeiten. Selbst wenn die offensichtlich unsinnigsten Kombinationen bereits vorab vermieden werden, kann das lediglich zu einer polynominalen Reduzierung führen deren Größenordnung bei n Gegenständen mit $n > 0$ im Bereich von 2^n bleibt.

Trotzdem die Berechnung einer Milliarde Lösungen innerhalb einer Sekunde erfolgen könnte, würde die Ermittlung des Endergebnisses mehrere Milliarden Jahre in Anspruch nehmen. Spielt die Dauer der Berechnung eines Ergebnisses also eine Rolle, gelangt man rasch zu heuristischen Lösungsansätzen. Dies ist oftmals nicht nur bei NP-vollständigen Problemen der Fall, sondern auch bei regulär errechenbaren Problemen wenn eine exakte Lösung deutlich mehr Zeit in Anspruch nimmt als eine heuristische Lösung und die Exakte auch nicht zwingend benötigt wird. Siehe [24] für eine ausführliche Beschreibung des Problems und eine komparative Analyse der verschiedenen algorithmischen Lösungsansätze.

4.2.1. Formale Allgemeindefinition und Pseudocode

Das Optimierungsproblem

$$\mathfrak{B} := (\boldsymbol{\Omega}, \boldsymbol{\mathcal{F}}, \preccurlyeq)$$

ist gegeben durch einen Suchraum,

$$\boldsymbol{\Omega}$$

einer Zielfunktion die jedem $x \in \Omega$ einen Fitnesswert $\mathcal{F}(x)$ zuweist,

$$\boldsymbol{\mathcal{F}}: \boldsymbol{\Omega} \rightarrow \mathbb{R}$$

sowie einem Vergleichsoperator

$$\preccurlyeq \in \{\leq, \geq\}.$$

Die Menge $\mathcal{M}$ der globalen Optima ist wie folgt definiert

$$\boldsymbol{\mathcal{M}} := \{\boldsymbol{x} \in \boldsymbol{\Omega} \mid \boldsymbol{\mathcal{F}}(\boldsymbol{x}) \preccurlyeq \boldsymbol{\mathcal{F}}(\boldsymbol{y}) \qquad \forall \boldsymbol{y} \in \boldsymbol{\Omega}\}.$$

Ein Phänotyp entspricht einem x aus dem Suchraum $\boldsymbol{\Omega}$. Der Genotyp sei definiertüber die bijektive Abbildung vom Suchraum $\boldsymbol{\Omega}$ in die Menge aller möglichen Genotypen $\boldsymbol{\mathcal{G}} := \boldsymbol{\mathcal{G}}_{\boldsymbol{\Omega}}$.

Eine simple Version des zugrundeliegenden Prozesses kann auch anschaulich als **Pseudocode** dargestellt werden:

`Problem:`

```
Optimierungsproblem mit Suchdomäne S und Fitness-Funktion f()
```

`Eingabe:`

```
Zielfunktion f, Populationsgröße p, Zahl der Kinder z,
Mutationswahrscheinlichkeit m, Abbruchbedingung c
```

Lösung:

```
Initialisiere zufällige startpopulation P
solange(abbruchbedingung c nicht erfüllt)
{
     errechne fitness aller chromosomen von P
     selektiere beste individuenin abhängigkeit vom fitness wert
     generiere z Kinder durch Anwenden der genetischen operatoren
     speichere Kinder und Eltern in P‘
     mutiere individuenaus P‘
}
```

Ausgabe:

```
Bestes Chromosom der Population
```

4.2.2. Individuum, Population und Kodierung

Sei das wie bereits in letztem Kapitel beschriebene Konstrukt

$$\mathfrak{B} := (\boldsymbol{\Omega}, \boldsymbol{\mathcal{F}}, \preccurlyeq)$$

ein Optimierungsproblem und der entsprechende Genotyp daher

$$\boldsymbol{G := dec^{-1}(\Omega).}$$

Das Tupel

$$\left(\mathbf{x}, \mathbf{g} := \mathbf{dec^{-1}(x)}, \mathbf{F(x)}\right) =: \mathbf{I(x)}$$

heißt **Individuum**, $F(x)$ wird die **Fitness** des Individuums $I(x)$ genannt. Weiter sei $\mathcal{P}(\mathbb{P})$ die Potenzmenge von

$$\mathbb{P} := \{\mathbf{I(x)} | \, \mathbf{x} \in \boldsymbol{\Omega}\} \subset (\boldsymbol{\Omega} \times \mathbf{dec^{-1}}(\boldsymbol{\Omega}) \times \mathbb{R})$$

der Menge aller zu $\mathfrak{B}$ gehörenden Individuen. Ein

$$\mathbf{P} \in \boldsymbol{\mathcal{P}}(\mathbb{P})$$

nennt man **Population**.

Je nach Problemstellung und den zugrunde liegenden Basisdaten eignen sich unterschiedliche Kodierungsformen für das Individuum bzw. damit auch indirekt für die Population. [25]

4.2.2.1. Standard-Binärcodierung

Diese Art der Kodierung ist eine der bekanntesten genotypischen Individuums-Repräsentationen und stellt eine dezimale Zahl im Binärformat dar. Jedes Gen entspricht dann einem Bit im „Bit-String". Ein Beispiel dafür ist die binäre Darstellung der Dezimalzahlen 3 (011) und 4 (100). Ein wesentlicher Nachteil dieser simplen Technik zeigt sich in weiterer Folge bei der Anwendung der genetischen Operatoren. Wird ein einziges Bit invertiert oder eine Bit-Kette beim Crossover ausgetauscht, kann das zu einer überproportionalen Änderung im Phänotyp führen. Umgekehrt können sich Genotypen benachbarter Phänotypen in jedem Bit unterscheiden was sowohl intuitiv nicht sinnvoll erscheint als auch bei der Umsetzung zu Problemen führt, da sich auch die kodierten Repräsentationen ähnlicher Individuen ebenfalls ähneln sollten. Ist das nicht der Fall, kann das bei der Kreuzung zweier Individuen zu beliebigen (also zufälligen) Ergebnissen führen wodurch der genetische Algorithmus letztendlich in einer nicht gewünschten Zufallssuche mündet.

4.2.2.2. Gray-Code

Der sogenannte Gray-Code bietet eine alternative binäre Darstellung die dieses Problem zumindest verringert. Ursprünglich verwendet wurde das Kodierungsverfahren zur robusten Übertragung digitaler Größen über analoge Signalwege – im Laufe der Zeit fand es allerdings in verschiedenen artfremden Bereichen wie jenem der Informatik Anwendung. Folgende Arbeitsschritte sind notwendig um eine Dezimalzahl in einen Gray-Code zu überführen:

(1) Codieren der Dezimalzahl im Standard-Binärcode als Dualzahl

(2) Rechts-Shift der Dual-Zahl um 1 Bit

(3) XOR-Verknüpfung von Ergebnis in (1) und (2)

Aus der Zahl 1 wird 010, nach Kopie und Rechts-Shift des Binärstrings ergibt sich 001 als zweites Ergebnis. Schließlich werden die Strings XOR-Verknüpft, dh. falls auf gleicher Position bei *einem einzigen der Beiden* eine Binäre 1 vorkommt, erscheint diese im aus der Verknüpfung resultierenden Ergebnis.

$$
\begin{array}{rll}
 & 010_2 & \text{Binärkodierung von Zahl 2} \\
\text{XOR} & 001_2 & \text{Ergebnis nach Right-Shift} \\
\hline
 & 011_2 & \text{Gray-Kodierung von Zahl 2}
\end{array}
$$

Im Gegensatz zur Standard-Binärcodierung ist die aus dieser Art der Kodierung resultierende Hamming-Distanz von k-1 und k immer 1. Die Hamming-Distanz wird als Maß für die Unterschiedlichkeit von Zeichenketten definiert [26]. Die Distanz zweier Blöcke mit fester Länge ist dabei die Anzahl der unterschiedlichen Stellen. In unten stehender Tabelle wird deutlich dass sich jene Zahlen die direkt untereinander stehen auch in ihrer Binärdarstellung mittels Gray-Kodierung lediglich geringfügig und vor allem konstant voneinander unterscheiden. Die Hamming-Distanz jener Zahlen die der Standard-Binärkodierung unterliegen, schwanken jedoch beliebig und hochgradig im Verhältnis zur Bit-Länge.

Zahl	Binärkodierung	Hamming-Dist.	Gray-Kodierung	Hamming-Dist.
0	000	-	000	-
1	001	1	001	1
2	010	2	011	1
3	011	1	010	1
4	100	3	110	1
5	101	2	111	1

Tabelle 1 - Binärkodierung vs. Gray-Kodierung

Dadurch wird ein wesentliches Kriterium erfüllt um in weiterer Folge die Operabilität der Kodierung im genetischen Algorithmus zu ermöglichen.

4.2.2.3. Nichtbinäre Repräsentation

Die nichtbinäre Repräsentation von Individuen kann für spezielle Problemstellungen manchmal zu wesentlich natürlicheren Lösungswegen als bei der Binärkodierung führen. Ein Vektor aus reellen Zahlen sowie eine Kodierung durch Matrizen, Permutationen oder auch Baumstrukturen sind hier denkbar.

4.2.3. Selektionsoperator

Die inhaltlichen Hintergründe der Individuum-Selektion wurden bereits im vorherigen Kapitel erläutert. Formal repräsentiert der Selektionsoperator also eine Abbildung

$$\boldsymbol{S^{e,k}}: \boldsymbol{\mathcal{P}}(\mathbb{P}) \rightarrow \boldsymbol{\mathcal{P}}(\mathbb{P}), \boldsymbol{E} \mapsto \boldsymbol{K} \subset \boldsymbol{E}$$

mit

$$|\boldsymbol{E}| = \boldsymbol{e} > k = |\boldsymbol{K}|.$$

wobei aus $\boldsymbol{e}$ Individuen $\boldsymbol{k}$ ausgewählt werden.

Die Selektion gibt dem GA also eine Suchrichtung vor deren Grundlage die Zielfunktion ist und erfolgt unter den Annahmen:

(1) die Individuen seien nach Fitness geordnet, d.h.

$$\mathbf{P(t)} = \{\boldsymbol{x_1}, \ldots, \boldsymbol{x_2}\}$$

dann gilt

$$\boldsymbol{\mathcal{F}(x_i)} \geq \boldsymbol{\mathcal{F}(x_j)} \textbf{ für i} \leq \textbf{j}$$

(2) wenn $Prob(x_1)$ als die Wahrscheinlichkeit gilt das i-te Individuum zu wählen, dann gilt

$$\sum_{\boldsymbol{i=1}}^{\boldsymbol{m}} \boldsymbol{Prob(x_i)} = \mathbf{1}$$

Je nach gewählter Art der Selektionsmethode, stehen die Individuen unter verschieden hohem Selektionsdruck bzw. Selektionsintensität in_s. Wenn $\boldsymbol{G}$ die durchschnittliche Güte und $\boldsymbol{\sigma}$die Standardabweichung der Güte der Population vor der Selektion ist, sowie $\boldsymbol{G_{sel}}$ die durchschnittliche Güte nach der Selektion darstellt dann gilt für die Selektionsintensität

$$\boldsymbol{in_s} = \frac{(\boldsymbol{G_{sel}} - \boldsymbol{G})}{\boldsymbol{\sigma}}$$

In weiterer Folge sollen die in [23] erläuterten Möglichkeiten der Selektion vorgestellt werden.

4.2.3.1. Fitnessproportionale Selektion

Hier wird die Wahrscheinlichkeit der Auswahl eines Individuums von der Höhe des Fitnesswertes relativ zu den übrigen Individuen einer Population errechnet.

$$p_i = \frac{f_i}{\sum_{j-1}^{N} f_j}$$

f_i ist dabei die Bewertungsfunktion, i ein Individuum und j der Index der Individuen bei insgesamt N Individuen - p_i steht daher direktproportional zur Fitness. Die Methode wird auch als "Roulette Wheel Selection" bezeichnet und kann gut durch eine Drehscheibe veranschaulicht werden. Die Größe jedes Abschnittes der Drehscheibe stellt ein Individuum mit bestimmter Fitness dar. Unten stehende Abbildung verdeutlicht die Funktionsweise.

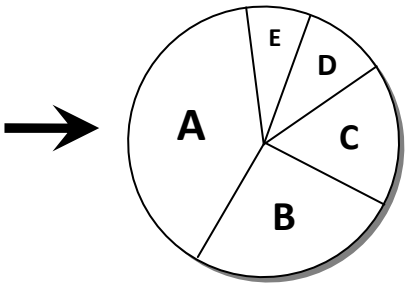

Abbildung 12 - Fitnessproportionale Selektion

Durch symbolisches Drehen der Scheibe wird ein Element selektiert - in ungünstigen Fällen von besonders hohen Ungleichheiten der Fitnesswerte einer Population, kann es passieren dass über die Iterationen hinweg fast ausschließlich ein einziges Individuum gewählt wird. Eine Lösung für das Problem bietet "Stochastic universal sampling".

Die Notwendigkeit anderer Selektionsmechanismen ergibt sich daraus, dass der Selektionsdruck bei dieser Variante in Abhängigkeit der Parameter verhältnismäßig hoch ausfallen kann. Dies erhöht die Wahrscheinlichkeit einer zu frühen Konvergenz in ein lokales Optimum, wodurch die Arbeitsgeschwindigkeit des Algorithmus und die Effizienz in der Auffindung einer zumindest akzeptablen Lösung allerdings verhältnismäßig hoch ist (ein Ausgleich kann vor allem durch die sogenannte Mutationswahrscheinlichkeit erreicht werden).

4.2.3.2. Rangbasierte Selektion

Bei dieser Form der Selektion, werden die Individuen nach der Bewertung durch die Fitnessfunktion anhand der Höhe der Fitness sortiert. Die Wahrscheinlichkeit hängt dann nur noch indirekt von der konkreten Fitness ab, sondern ist proportional zum Rang. Der Selektionsdruck kann dabei über die Einstellung der maximalen Auswahlwahrscheinlichkeit des ranghöchsten Elementes gesteuert werden.

4.2.3.3. Wettkampfselektion

Aus dem Gesamtbestand aller Individuen werden jeweils x Individuen gewählt, wobei gilt

$$2 \leq x \leq z, mit\ z = |\mathbf{P}|.$$

Die Individuen werden dabei mit gleicher Selektionswahrscheinlichkeit gewählt und anschließend jenes Individuum mit der höchsten Fitness innerhalb der gezogenen Untermenge mit Größe x für die weitere Kreuzung ausgewählt. Dieser Vorgang wird solange wiederholt bis eine genügend große Anzahl an Individuen für die nächste Iteration des genetischen Algorithmus gewählt wurde. Der Selektionsdruck kann über die Größe der gewählten Untermenge bestimmt werden, wobei gilt: umso größer die Gruppe umso höher der Selektionsdruck. Zu bedenken ist jedenfalls, dass eine Mehrfachselektion ein und desselben Individuums erfolgen kann, wodurch sich die Gefahr einer zu frühen Konvergenz erhöht. Diese Art der Selektion ist im Allgemeinen unabhängig von der Verteilung der Fitnesswerte.

4.2.3.4. Stochastic Universal Sampling

Eine bekannte Möglichkeit der probabilistischen Selektion ist die Auswahl durch das sogenannte „stochastic universal sampling“ welche 1987 von James Baker eingeführt wurde. Wie auch bei der "Roulette WheelSelection" stellt eine Drehscheibe die Basis dar, die Anzahl der Auswahlpfeile entspricht der Anzahl der in der Population enthaltenen Individuen - es erfolgt daher nur ein einziger Selektionsdurchgang. Die Auswahlpfeile sind im gleichmäßigen Abstand um die Scheibe angeordnet.

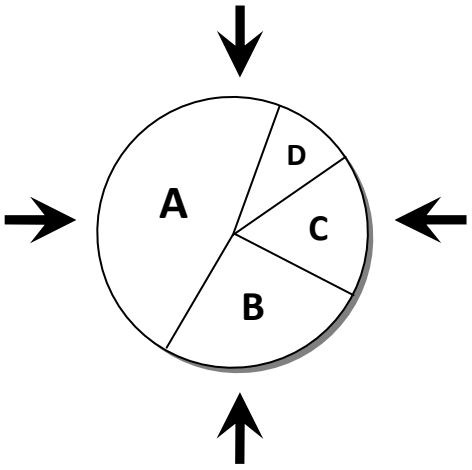

Abbildung 13 - Stochastic Universal Sampling

Unten stehende Tabelle zeigt die daraus resultierende Anzahl von Individuen im finalen Pool.

Individuum	Erwartungswert	Auswahlhäufigkeit
A	0,45	2
B	0,30	1
C	0,15	1
D	0,10	0

Tabelle 2 - Verteilung der Auswahlswahrscheinlichkeiten im Stochastic Universal Sampling

Möchte man die Selektionsmechanismen nach ihrem Selektionsdruck reihen, würde sich im Groben tendenziell folgende Reihung ergeben:

1. Fitnessproportionale Selektion
2. Rangbasierte Selektion
3. Turnierselektion
4. Selektion mit Stochastic Universal Sampling

4.2.4. Rekombinationsoperator

Ein Rekombinationsoperator mit $\boldsymbol{e}$ Eltern und $\boldsymbol{k}$ Kindern entspricht der Abbildung

$$\boldsymbol{R}^{e,k}: \boldsymbol{\mathcal{P}}(\mathbb{P}) \rightarrow \boldsymbol{\mathcal{P}}(\mathbb{P}), \boldsymbol{E} \mapsto \boldsymbol{K} \supset \boldsymbol{E}$$

mit

$$|\boldsymbol{E}| = \boldsymbol{e}, \boldsymbol{e} + \boldsymbol{k} = |\boldsymbol{K}|.$$

Wesentlich ist hier, dass die Rekombination die Population um neue Individuen erweitert, die Genvielfalt der Population dabei jedoch tendenziell einschränkt. Die Anzahl der in der Population enthaltenen Individuen bleibt dabei gleich – jene Individuen die zuvor durch den Selektionsoperator nicht zur Rekombination ausgewählt worden sind, werden durch Kinderindividuen der ausgewählten Paare ersetzt wodurch vorteilhafte Eigenschaften zusammengeführt werden sollen

Es gibt unterschiedliche Grundsatzvarianten von Rekombinationsoperatoren:

- **Kombinierender Operator:** Erbinformationen werden ausschließlich aus dem Genotyp der Eltern extrahiert und neu angeordnet.
- **Expandierender Operator:** Die Kombination von Individuen erfolgt bei dieser Variante nicht "blind" bzw. ohne zusätzliche Information. Eine Suchrichtung die auf der Vergabe von Gütewerten von Genkombinationen der Elternindividuen kann dabei vorgegeben werden. Die Rekombination kann dadurch in eine

problemspezifischere Richtung gelenkt werden, was zu optimaleren Ergebnissen führen kann.

In weiterer Folge sollen die wichtigsten technischen Verfahrensmöglichkeiten der Rekombination, teils angelehnt an [23], näher beschrieben werden.

4.2.4.1. One-point-crossover

Bei der Kreuzung sind jeweils ausgehend von einer beliebig zu wählenden Position zwei Elternindividuen zu kombinieren. Gegeben sind also 2 Bit-Strings x, y der Länge l=8 wobei für den Schnittpunkt gilt: $s \in \mathbb{N}: 0 \leq s \leq l$.

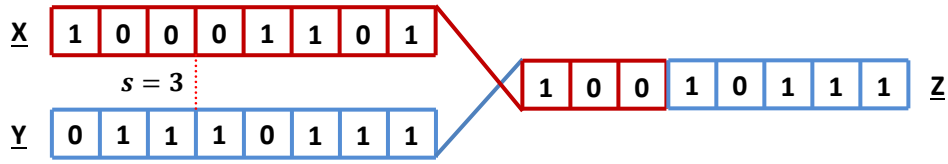

Abbildung 14 - One-point-crossover

Die oben stehende Grafik zeigt ein einfaches Crossover zwischen zwei binär kodierten Individuen an bei Position 3. Im Regelfall werden zwei Elternindividuen durch zwei neue Kinderindividuen ersetzt um die Population konstant zu halten, aus Gründen der Übersicht wird lediglich die Erstellung eines Nachkommens gezeigt.

4.2.4.2. Multi-point-crossover

Die Kreuzung beim multi-point-crossover erfolgt wie der Name bereits suggeriert an N Punkten $s_1, \ldots, s_N$ mit $s_1 < s_2 < \cdots < s_N$ wobei $N \leq l$.

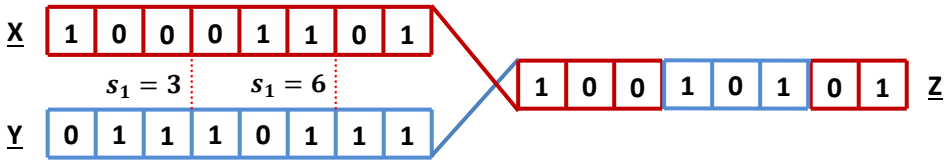

Abbildung 15 - Multi-point-crossover

Der Vorteil dieses Kreuzungstypen liegt in den vielfältigeren Resultaten und der besseren Durchmischung der genetischen Bausteine. Dieses trägt tendenziell wie auch die Mutation von Einzelgenen dazu bei zu frühe Konvergenz in ein lokales Optimum zu verhindern.

4.2.4.3. Uniform-crossover

Bei der Variante des uniform-crossover wird für jedes einzelne Gen an der jeweiligen Stelle eine Zufallsauswahl mit gleicher (also uniformer) Wahrscheinlichkeit getroffen. Die unten stehende Abbildung soll diese Auswahl verdeutlichen.

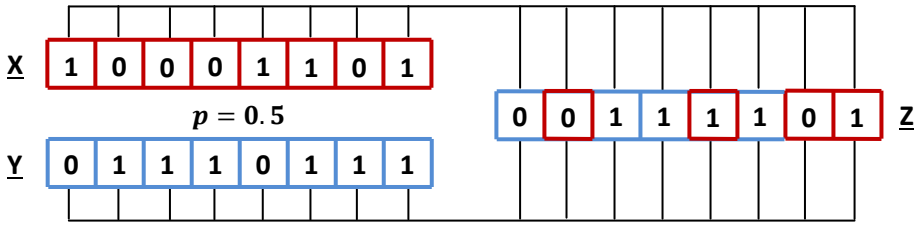

Abbildung 16 - Uniform-crossover

Das Ergebnis Z ergibt sich also aus einer Kombination der Einzelelemente von X, Y auf Basis von $p = 0.5$. [27]

4.2.4.4. Shuffle-crossover

Der Shuffle-crossover ist eine simple Erweiterung des One- bzw. Multi-point-crossovers um eine Durchmischung der Genposition vor der eigentlichen Rekombination. Nach der Operation werden die Genpositionen wieder in ihre Ursprungsposition zurückversetzt. Dies führt zu einer Eliminierung des sogenannten Distributional Bias. Nachfolgende Grafik zeigt den Vorgang an Hand eines One-point-crossovers in einzelnen Schritten.

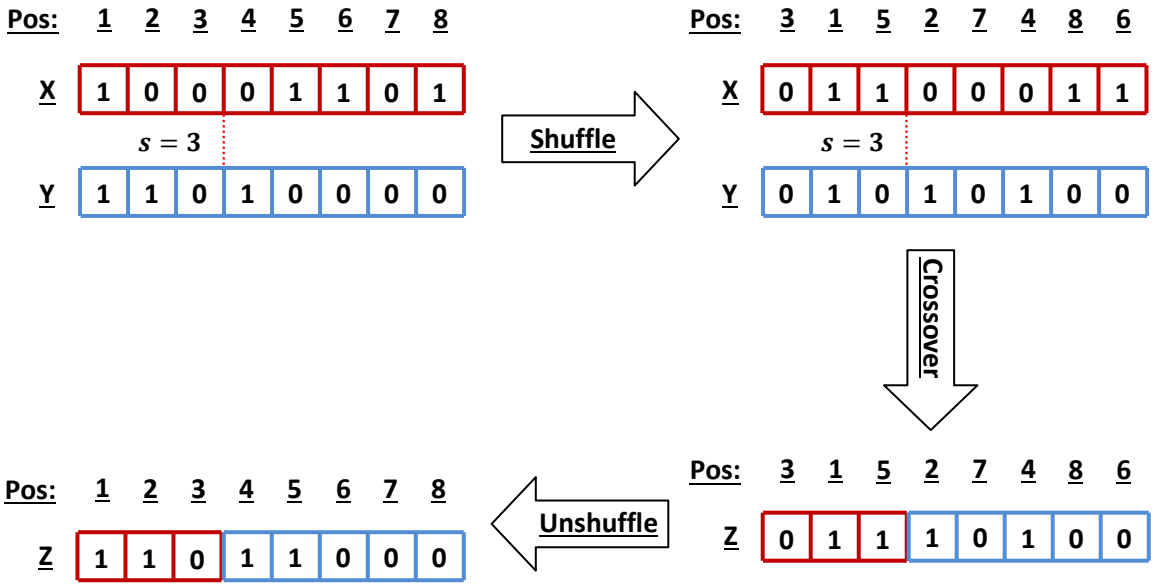

Abbildung 17 - Shuffle-crossover

einem One-Point-Crossover, welcher keinerlei Distribuational Bias besitzt, daher einen überaus vorteilhaften Rekombinationsalgorithmus. [28]

4.2.4.5. Positional- und Distributional Bias

Ein sinnvoller Vergleich der unterschiedlichen Crossover-Operatoren findet auf Basis zweier für die Entscheidungsfindung wesentlichen Problempunkte statt:

- **Positional Bias:** Wenn eine Abhängigkeit zwischen der Höhe der Wahrscheinlichkeit gewisse Gene bei einer Rekombination weiterzugeben, von der Position dieser Gene innerhalb des Chromosoms abhängt, spricht man von einem "positional bias". Ist dieser Bias also stark ausgeprägt, werden Gene mit bestimmter Positionsnummer mit wesentlich höherer Wahrscheinlichkeit ausgetauscht als andere. Ein gutes Beispiel dafür ist das One-Point-Crossover, bei dem die Wahrscheinlichkeit des Austausches mit steigender Positionszahl zunimmt.

- **Distributional Bias:** Unterliegt der Erwartungswert der Anzahl der ausgetauschten Gene bei der Rekombination keiner Gleichverteilung zwischen den Positionen, handelt es sich um einen distributional bias. Das One-Point-Crossover weist in dieser Kategorie keinerlei Verschiebung auf - jede Anzahl an Elementen wird daher mit gleich großer Wahrscheinlichkeit getauscht.

Crossover-Typ	Positional Bias	Distributional Bias
One-Point-Crossover	hoch	nicht vorhanden
Multi-Point-Crossover	Geringer als One-Point-Crossover	mit steigender Anzahl zunehmend
Uniform-Crossover	nicht vorhanden	Sehr hoch und zunehmend mit steigender Länge
Shuffle-Crossover	nicht vorhanden	Wie der distributional bias des gewählten Basisoperators (One- / Multi-Point)

Tabelle 3 - Vergleiche zwischen den Crossover-Typen

4.2.5. Bewertungs- und Fitnessfunktion

Die Unterscheidung in eine sogenannte Bewertungsfunktion und eine Fitnessfunktion ist dahingehend notwendig, als dass Erstere die Güte eines Individuums bezogen auf das Optimierungsziel misst, die Fitnessfunktion jedoch die Wahrscheinlichkeit der Fortpflanzung. Da im Regelfall ein direkter Zusammenhang zwischen Güte des Individuums und seiner Reproduktionswahrscheinlichkeit sinnvoll ist, werden die zwei Funktionen in manchen Implementierungsvarianten nicht voneinander unterschieden. Häufig kommt aber auch jener Fall vor, dass Bewertungs- und Fitnessfunktion in direkter Proportion zueinander stehen, sodass die Fitnessfunktion lediglich eine Funktion der Bewertung darstellt.

Die Bewertungs- und Fitnessfunktion eines genetischen Algorithmus kann nicht verallgemeinert werden und ist hochgradig problemspezifisch. Die Funktion repräsentiert den Kern des Prozesses und dient zur Bewertung eines Individuums. Formal ausgedrückt

weist, wie bereits beschrieben, eine Zielfunktion jedem $x \in \Omega$ einen Wert $\mathcal{F}(x)$ aus dem reellen Zahlenraum $\mathbb{R}$ zu

$$\boldsymbol{\mathcal{F}}: \boldsymbol{\Omega} \rightarrow \mathbb{R}.$$

Die "Umwelt" und sämtliche problemspezifischen Einflussfaktoren in welcher sich der genetische Algorithmus bewegt werden über diese Funktion gesteuert. Lässt sich die Bewertung eines Individuum auf Grund der komplexen Umweltbedingungen nicht auf lediglich einen Zahlenwert einschränken, so kann das Ergebnis der Bewertung auch in Ergebnis-Tupel münden, in welchen über die Methode der Vektoroptimierung das beste Ergebnis ermittelt wird.

4.2.6. Mutationsoperator

Ein Mutationsoperator entspricht einer Abbildung

$$\boldsymbol{M}: \boldsymbol{\mathcal{P}}(\mathbb{P}) \rightarrow \boldsymbol{\mathcal{P}}(\mathbb{P}), \boldsymbol{P} \longmapsto \boldsymbol{P'}$$

mit

$$|\boldsymbol{P}| = |\boldsymbol{P'}|.$$

Die Höhe des Einflusses des Mutationsoperators bestimmt die Breite der Streuung und Geschwindigkeit der Konvergenz. Umso Einflussreicher der Mutationsoperator umso größer die Erreichbarkeit sämtlicher Punkte im Suchraum - im Extremfall verhindert der Mutationsoperator allerdings jegliche Konvergenz, da durch die starke Veränderung kompetitiver Individuen nach jeder Rekombination, der Algorithmus seine Suchrichtung verliert. Bei geringem Einfluss des Operators, wird lediglich die direkte Umgebung des Individuums überprüft - große zufällige Sprünge in der Fitness der Individuen sind dadurch ausgeschlossen, jedoch auch eine zu breite Streuung und dadurch Unschärfe des Algorithmus. Da der Operator in der Regel ergänzend verwendet wird und keinen dominierenden Einfluss haben sollte, wird die Mutationswahrscheinlichkeit in der Regel sehr gering angegeben - häufig gewählte Wahrscheinlichkeiten sind $p_m = 0.01$ oder $p_m = 0.001$. Der Operator soll also lediglich den durch die Rekombination entstandenen Selektionsdruck etwas verringern.

In weiterer Folge werden die wichtigsten Umsetzungsmöglichkeiten nach [21] und [23] in Kürze erläutert.

4.2.6.1. Uniforme Verteilung

Die Mutationswahrscheinlichkeit bleibt über die Iterationen des genetischen Algorithmus hinweg stabil. Der Operator ersetzt das betreffende Gen mit einem Wert aus einem zuvor festgelegten Wertebereich, wobei jeder dieser Werte die gleiche Auswahlwahrscheinlichkeit besitzt. Wird das Individuum als Vektor implementiert, wird das betreffende Gen durch eine Alternative aus dem Gen-Pool ersetzt - die Auswahl aus dem Pool erfolgt mit einer Wahrscheinlichkeit von $\frac{1}{n}$.

4.2.6.2. Non-uniforme Verteilung

Im Gegensatz zur uniformen Verteilung wird die Mutationswahrscheinlichkeit eines Gens von Generation zu Generation geringer. Eine Anfangs höhere Wahrscheinlichkeit verhindert eine zu frühe Konvergenz in ein lokales Optimum, in den späteren Phasen des Algorithmus wird das Ergebnis daher nur noch geringfügig verändert. Mit jeder Iteration nähert sich die Mutationswahrscheinlichkeit den 0% an.

4.2.6.3. Invertieren eines oder mehrerer Bits

Ähnlich der uniformen Verteilung - besteht ein Individuum aus einem Bit-String und daher aus den möglichen Genvarianten "0" oder "1", wird lediglich aus dieser Bandbreite gewählt. Dies führt jedoch zu einer noch geringer als angegebenen Mutationswahrscheinlichkeit, da die Wahrscheinlichkeit einer Genveränderung bei lediglich 50% liegt. Bei einer Implementierungsvariante mit Bit-Strings wird daher im Normalfall lediglich ein "Bit-Flip" bzw. eine Invertierung des Bits durchgeführt.

4.2.6.4. Verteilung mit kleinerer Wahrscheinlichkeit, je höher die Bitposition

Eine weitere Adaption liegt in der Festlegung unterschiedlicher Mutationswahrscheinlichkeiten für Gene an verschiedenen Positionen im Bit-String was vor allem in Kombination mit der Standard-Binärkodierung sinnvoll erscheint. Dadurch kann die positionsabhängige Wertigkeit der Gene teilweise kompensiert werden.

In weiterer Folge sollen die oft als Randthemen behandelten Themen der Ersetzungsschemata zur Erneuerung der Population, sowie des gewählten

Abbruchkriteriums für den genetischen Algorithmus insgesamt, kurz diskutiert werden. Die Erläuterungen sind dabei angelehnt an [29].

4.2.7. Ersetzungs- und Ergänzungsschemata

Unterschiedliche Vorgehensweisen gibt es ebenfalls bei der Neubildung der Population auf Basis der bei der letzten Rekombination erzeugten Individuen. Jede der vorgestellten Möglichkeiten besitzt spezifische Eigenschaften und daher auch Nachteile.

4.2.7.1. Komplette Ersetzung

Das sogenannte "general replacement" ist die am weitesten verbreitete Technik. Hierbei werden sämtliche alten Individuen durch komplett neue Erzeugnisse ersetzt. Das hat jedoch den nachteiligen Effekt, dass das potentiell beste Individuum der letzten Iteration verloren geht wodurch es zu einer in Summe negativen Beeinflussung der Gesamtpopulation kommen kann. Als Vorteil ergibt sich daraus jedoch auch, dass einzelne Individuen welche in der Frühphase des Algorithmus relative Stärke besitzen, die Gesamtpopulation rasch dominieren.

4.2.7.2. Prinzip der Eliten

Behält man nun einen gewissen Prozentsatz der stärksten Individuen, dann spricht man auch von Elitismus. Der Hauptnachteil dieser Methode ist der oben behandelte Vorteil der kompletten Ersetzung. Um das Manko zu verringern, kann man die stärksten Individuen der letzten Iteration nochmals mit dem Mutationsoperator behandeln bevor sie in die neue Population übernommen werden um die Merkmale geringfügig zu verändern. Bei dieser Variante spricht man vom "schwachen Elitismus".

4.2.7.3. Delete-N-Last Schema

Die Herangehensweise dieser Methode ist logisch verkehrt. Man löscht von der bestehenden Population einen gewissen Teil der schwächsten N Individuen heraus und ersetzt sie durch neu generierte. Je nachdem wie hoch der Anteil der alten Generation an der Neuen ist, spricht man auch von einem "steadystate"-Ersetzungsschema.

4.2.8. Abbruchkriterium

Die Natur eines genetischen Algorithmus erfordert ein Abbruchkriterium. Da dieses nicht in der Bestimmung der absoluten Korrektheit der Lösung liegen kann, muss man sich beim Abbruch der Verarbeitung auf Indikatoren verlassen die das Erreichen einer nahezu

optimalen Lösung zumindest andeuten. Das Terminieren eines Prozesses basiert im realen Einsatz jedoch durchaus nicht ausschließlich auf objektiv messbaren Gründen.

4.2.8.1. Fixe Anzahl an Iterationen

Eine der simpelsten Lösungen ist es nach einer bestimmten Anzahl an Iterationen abzubrechen und die bis dahin beste Lösung auszugeben. Entstehen nach erfahrungsgemäß X Iterationen keine besseren Lösungen mehr und wird das Bewertungsverfahren nicht verändert, ist das eine legitime und simple Möglichkeit das Problem zu lösen.

4.2.8.2. Relative Veränderung der Fitnesswerte

Verändert sich die Datenbasis oder die Fitnessfunktion dynamisch, sollte die Verbesserungsquote der jeweils aktuell besten Lösung pro Iteration im Vergleich zur letzten Iteration betrachtet werden. Verbessert sich diese nach einer bestimmten Anzahl an Durchläufen nur noch minimal, ist die Wahrscheinlichkeit für das Auffinden einer noch optimaleren Lösung durch reine Mutation sehr gering.

4.2.8.3. Diversität der Gesamtpopulation

Eine ähnliche Variante ist die Analyse der Gesamtpopulation auf Diversität. Sind sämtliche Individuen genetisch ähnlich, der genetische Pool daher bereits stark eingeschränkt, ist das Auffinden einer besseren Lösung ebenfalls minimal. Oftmals wird auch eine Kombination der genannten Möglichkeiten gewählt.

5. Anwendung auf das Portfoliooptimierungsproblem

Das Ziel dieser Arbeit ist der Aufbau eines GA-Setups zur Entwicklung eines erweiterbaren Frameworks für Portfolio-Management-Zwecke. Aus einer vordefinierten Auswahl unterschiedlicher Assets, soll eine für den gewählten Testzeitraum optimale Zusammensetzung generiert werden. Diese Zusammensetzung von optimal gewichteten Komponenten soll im weiteren Verlauf stets als Portfolio bezeichnet werden - an diesem Punkt erfolgt eine lediglich logische Überführung eines technischen Problems in eine wirtschaftswissenschaftliche Nomenklatur. Das Optimierungsergebnis, generiert aus der Dynamik und dem Zusammenspiel sämtlicher Assets der Basis, entspricht daher einer Anlageempfehlung oder dient als einer von vielen zu beachtenden Indikatoren für den aktiven Portfoliomanager zur optimalen Portfolioallokation in der planbaren Zukunft.

Die entwickelte Optimierungsumgebung soll einem Portfolio-Manager die Möglichkeit bieten, die Datenbasis ad-hoc auszubauen oder komplett zu verändern und die Fitnessfunktion verhältnismäßig simpel ersetzen bzw. um beliebige Parameter erweitern zu können. Der Vorteil der Herangehensweise an die Portfoliooptimierung mittels genetischer Algorithmen, soll vor allem durch die beliebige Erweiterung der Optimierungsfunktion zur Geltung kommen. Durch die eingesetzte Heuristik ist es nicht mehr relevant ob die gewählte Optimierungsmethode zu komplex zu berechnen (NP-Vollständig) ist oder nicht. Durch die Implementierung optischer Feedbacks kann der Optimierungsfortschritt beobachtet und gegebenenfalls manuell beendet werden. Alle wesentlichen, in folgenden Unterkapiteln vorgestellten Optionen des genetischen Algorithmus können über eine grafische Oberfläche angepasst werden.

Neben der Festlegung der wichtigsten Rahmenbedingungen und des eigentlichen Optimierungszieles (inklusive Datenbasis), ist ein wesentlicher und zugleich besonders schwieriger Schritt, die Übersetzung eines für das Portfoliomanagement typischen Optimierungsproblems in eine für einen genetischen Algorithmus lösbare Form. Die dahingehenden Überlegungen zur Überführung der Problematik in einen verarbeitbaren Zustand betreffen folglich die

- Art der Kodierung eines Portfolios zur weiteren Verarbeitung,
- Ermittlung geeigneter und portfoliospezifischer Bewertungsfunktionen,
- Systematik der Rekombination zur Erzeugung neuer Individuen,
- Findung geeigneter Mutationssystematik,
- Sicherstellung der Erzeugung ausschließlich valider Individuen in jedem Schritt,
- Erstellung eines neuen Portfolio-Pools,
- Überlegung eines geeigneten Backtesting-Mechanismus für das optimierte Portfolio.

Die Überführung des Portfoliooptimierungsproblems in ein kombinatorisches Problem wurde durchaus in einigen wenigen Publikationen behandelt, jede der vorgeschlagenen Lösungen scheint jedoch exorbitant komplex oder methodisch fragwürdig. Aus diesem Grund wird eine eigens entwickelte Methode vorgestellt, mit welcher die Problematik auf verhältnismäßig simple Art gelöst werden soll. Die aus dem Testaufbau resultierenden Ergebnisse sowie die Dynamik der Konvergenz gegen ein scheinbares Optimum über die Zeit scheinen die gewählte Vorgehensweise zu bestätigen.

Ein Versuch der Zusammenstellung eines kompetitiven Portfolios mittels genetischer Algorithmen ist in [30] dokumentiert. Die dort erläuterten Ergebnisse erscheinen allerdings schwer nachvollziehbar - zudem ist der gewählte Ansatz der Kodierung der Individuen unnötig komplex wodurch sich in weiterer Folge eine ebenso umständliche Rekombination und Prüfung der daraus entstandenen Individuen auf Validität ergibt.

In weiterer Folge sollen die einzelnen Schritte die im Entwicklungsprozess des entwickelten Algorithmus durchlaufen wurden, sowie jene die diesem vorausgegangen sind, im Detail nachvollziehbar beschrieben werden.

5.1. Rahmenbedingungen und Datenbasis

Eine grundsätzliche Entscheidung über die groben Rahmenbedingungen und die sogenannte Umwelt in welcher der genetische Algorithmus operiert, musste direkt zu Beginn der Entwicklungsarbeit getroffen werden. Um ein Setup mittels Echtdaten zu erzeugen, welches sowohl realistische Bedingungen, einfache Zugänglichkeit wie auch eine mit begrenzten

Rechnerressourcen ausgestattetes Versuchsumfeld berücksichtigt, mussten gewisse Abstriche im Datenvolumen und der Granularität der Zeitreihen gemacht werden. Weiters sollte der Testzeitraum eine möglichst hohe Bandbreite an marktüblichen Stärke- und Schwächephasen aufweisen um den Algorithmus auf Perioden mit fundamental unterschiedlicher Eigendynamik zu testen.

Als Optimal geeignet erscheint hierfür der "Dow Jones-Index" vom Jahr 2000 bis Anfang 2012, welcher eine kapitalgewichtete Komposition aus 30 verschiedenen, hochkapitalisierten US-Industriebetrieben und Banken darstellt. Der Index repräsentiert einen der wichtigsten Börsenindizes der Welt und besteht zugleich aus einer überschaubaren Anzahl an Komponenten (vgl. S&P 500), deren Zeitreihen zudem vollständig verfügbar sind. [31]

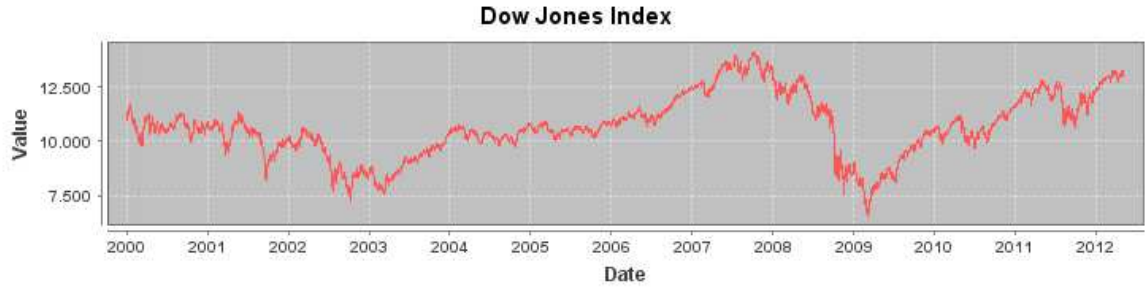

Abbildung 18 - Dow Jones Index 2000 bis 2012

Um ein besseres Verständnis für die fundamentalen Hintergründe der oben dargestellten Preisentwicklung, sowie den später vorgestellten Ergebnissen des genetischen Algorithmus zu erreichen, sollen auch in Kürze die markttreibenden Faktoren erläutert werden. De gewählte Zeitraum lässt sich also in vier durchwegs verschiedene, aufeinanderfolgende Phasen beschreiben:

1. **2000 bis Ende 2002:** Seitwärtsphase mit deutlichen Bärenmarkttendenzen. Die Aktien des Index entwickeln sich in Summe volatil Seitwärts mit einem Hang zur rezessiver bzw. deflationärer Tendenz. Diese Phase war vor allem geprägt durch das Platzen der sogenannten Dotcom-Blase - ein Vorgang der im Speziellen auf große Industriewerte einen verhältnismäßig geringen Einfluss hatte, sich jedoch auf Grund seiner problematischen deflationären und psychologischen Nebeneffekte auch hier wiederspiegelt.

2. **2003 bis Mitte 2007:** Bullenmarkt mit längeren Seitwärtsphasen. Hintergrund der positiven Marktentwicklung bestand in einer Lockerung der Geldpolitik der Zentralbanken und darauf folgender exzessiver Kreditvergabe von Banken an Schuldner mit geringer Bonität, besichert durch wenig werthaltige Immobilien. Die allgemeine Euphorie zeigt sich hier vor allem in Bankentiteln welche innerhalb weniger Jahre um mehrere 100% zulegen können - der Dow Jones Index in seiner gesamten Breite legt um ca. 70% (exklusive Dividendenzahlungen) zu. Eine Phase der Übertreibung beginnt.

3. **Mitte 2007 bis Anfang 2009**: Scharfer Bärenmarkt und systemische Krise. Die zuvor exzessiv auf Basis von zuletzt stark gestiegenen Immobilienpreisen vergebenen Kredite, werden nach und nach nicht mehr bedient. Durch den weltweiten Vertrieb der Schuldtitel mittels strukturierter Produkte kommt es nach dem Zusammenbruch des US-Häusermarktes zu einem Komplettverlust der Besicherungsbasis und vermuteten Totalabschreibungen in weltweiten Bank- und Versicherungsbilanzen. Durch gegenseitiges Misstrauen und dringenden Kapitalbedarf, trocknet der gesamte Interbankengeldmarkt aus - sämtliche Asset-Preise sinken durch erzwungene Verkäufe und Kapitalabzüge scharf. Die Phase der Untertreibung beginnt. Sämtliche Wertpapiere, unabhängig von ihrem tatsächlichen Wertgehalt, leiden unter dem Liquiditätsmangel an den Märkten. Die Cashquote in den Portfolios und das psychologische Negativsentiment erreichen ihren Höhepunkt.

4. **Anfang 2012 bis Ende Testzeitraum**: Sprunghafter Bullenmarkt mit scharfen Korrekturen. Durch unbeschränkte Anleihenkäufe der Notenbanken weltweit, der Senkung des US-Leitzinses auf quasi Null und der Notrettung eines Großteils des Banken- und Versicherungssystems, konnte eine rasche Erholung sämtlicher Asset-Preise erzielt und eine drohende Deflation vermieden werden. Fundamental erfolgte eine signifikante Marktbereinigung die sich auch in teils substantiellen Wertveränderungen einzelner Unternehmen innerhalb des Dow Jones Index nachvollziehen lässt. Vor allem Bankenwerte und bereits in Boomjahren kaum rentable Unternehmen gingen nun endgültig pleite (bzw. wurden verstaatlicht) oder wurden von der stärkeren Konkurrenz übernommen.

Unten stehende Tabelle verschafft einen Überblick über die im Index enthaltenen Einzelwerte. Die Auflistung soll verdeutlichen auf Basis welcher Werte die spätere Optimierung der Gewichte durch den genetischen Algorithmus erfolgt. Die Indexkomposition entspricht jener zum Stichtag 03.01.2000 laut Bloomberg (Informationsdienstleister), diese veränderte sich vor allem nach der Finanzkrise signifikant durch die seitdem deutlich abweichende Marktkapitalisierung einzelner Unternehmen.

Nr.	Unternehmen	Branche
1	3M	Konsumgüter
2	Alcoa	Energie / Rohstoffe
3	Altria Group	Konsumgüter
4	American Express Co	Finanzen
5	AT&T Inc	Telekommunikation
6	Boeing Co	Transport / Verkehr
7	Caterpillar Inc	Industrie
8	Citigroup Inc	Finanzen
9	Coca-Cola Co	Konsumgüter
10	Eastman Kodak	Technologie
11	DuPont Co	Chemie / Pharma
12	Exxon Mobil Corp	Energie / Rohstoffe
13	General Electric Co	Technologie
14	Hewlett-Packard Co	Technologie
15	Home Depot Inc/The	Handel
16	Honeywell International Inc	Technologie
17	Intel Corp	Technologie
18	IBM Co	Technologie
19	International Paper Co	Industrie
20	Johnson & Johnson	Chemie / Pharma
21	McDonald's Corp	Konsumgüter
22	Merck & Co Inc	Chemie / Pharma
23	Microsoft Corp	Technologie
24	Procter & Gamble Co	Konsumgüter
25	United Technologies Corp	Industrie
26	Wal-Mart Stores Inc	Handel
27	Walt Disney Co	Medien / Unterhaltung / Freizeit
28	AT&T Corp	Telekommunikation
29	JP Morgan	Finanzen
30	General Motors Company	Industrie

Tabelle 4 - Zusammensetzung Dow Jones Index zum Jahr 2000

Die Daten sämtlicher oben gelisteten Komponenten wurden auf Tagesbasis abgezogen und synchronisiert. Das tägliche Zeitintervall sichert die Möglichkeit der effektiven Optimierung auch auf einer zeitlich kürzeren Spanne (z.B. auf wenige Wochen) - eine Optimierung auf

mehrere Jahre wird durch die gezielte Auswahl von Wochen-Endkursen ermöglicht, wodurch auch die Berechnungsdauer in einer akzeptablen Zeitspanne erfolgen kann. Unter Umständen fehlende Einzelwerte innerhalb der Zeitreihen wurden mittels linearer Interpolation vervollständigt.

Wichtig ist, dass die Datenbasis zu experimentellen Zwecken quasi beliebig gewählt werden kann. Vorstellbar ist auch eine Anreicherung um Assets die einem Portfolio typischerweise beigemischt werden um die Volatilität zu verringern und die Diversifikation zu erhöhen. Das können Währungspaare (z.B. EUR/USD, EUR/JPY) aber auch Rohstoffe (z.B. Öl, Kupfer, Gold, Weizen) oder ein Volatilitätsindex (z.B. VIX) sein.

In weiterer Folge soll näher auf das eigentliche Setup und die wichtigsten Datenstrukturen eingegangen werden auf deren Basis die eigentliche Anwendung erstellt worden ist.

5.2. Experimentelle Definition und Implementierung

Das im letzten Kapitel beschriebene Basiskonstrukt muss nun in einen genetischen Algorithmus überführt werden. Die dafür notwendigen Überlegungen basieren vollständig auf den vorgestellten Theorien. Die Schritte der Selektion, Rekombination und Mutation werden Schritt für Schritt auf den Bereich der Portfoliooptimierung übersetzt.

5.2.1. Individuenkodierung und Initialisierung

Der erste Schritt liegt in der Überlegung der Kodierung einzelner Assets und eines Gesamtportfolios mit Asset-Gewichten, hierfür definieren wir ein **Gen** als ein **Asset** und ein **Chromosom** als **Kombination von Assets** (welche sich jedoch nicht zwingend voneinander unterscheiden müssen). Das Individuum repräsentiert daher eine mögliche Lösung des Optimierungsproblems.

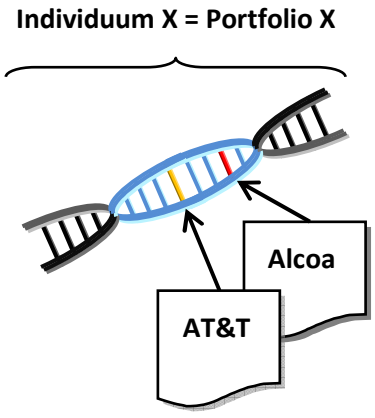

Abbildung 19 - Anwendungsbeispiel Individuum / Portfolio

im Kontext der Portfoliooptimierung valide verwendbar zu machen muss die Anzahl der Assets innerhalb eines Portfolios eingeschränkt werden. Um als Endergebnis eine gängige Asset-Allokation in Prozentpunkten zu erhalten, wurde diese auf 100 normiert. Jedes Asset spiegelt also ein einziges Gen wider welches wiederum 1% des Gesamtportfolios ausmacht. Die Gewichtung einzelner Assets wird also indirekt durch den genetischen Algorithmus über die Anzahl der Gene gesteuert in denen z.B. eine Aktie vorkommt. Folgende Grafik zeigt die Hülle eines Portfolios deren Plätze durch Assets besetzt werden sollen.

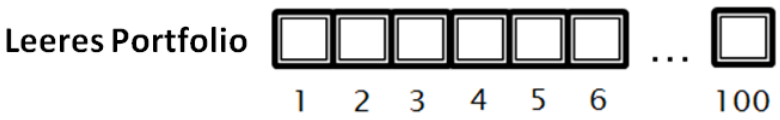

Abbildung 20 - Beispiel leeres Portfolio

Unten stehende Grafik zeigt einen beispielhaften Basisdatensatz von 4 Möglichen Assets und ein zufällig generiertes Portfolio.

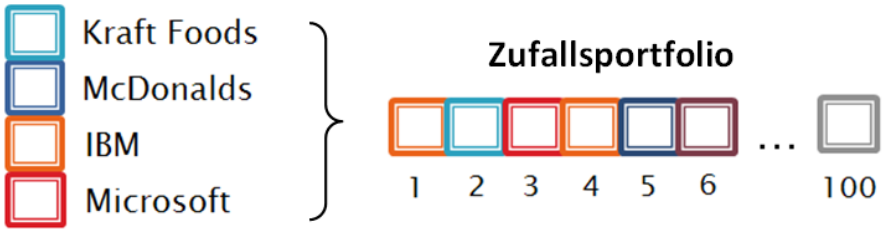

Abbildung 21 - Erstellung eines Zufallsportfolios

Im sichtbaren Teil des Portfolios befinden sich also 2% IBM, 1% Kraft Foods, 1% Microsoft und 1% McDonalds. Die Idee ist nun, dass sich die Allokation besonders optimaler Aktien, sowie deren Zusammensetzung im Laufe des Verfahrens zu einem Optimum über den gewählten Zeitraum entwickeln. Vorteilhaft an dieser Form der Strukturierung ist auch, dass bei jeder Form der Rekombination ein valides Individuum erzeugt wird – das prozentuelle Gewicht eines Wertes wird **nicht explizit** kodiert sondern ergibt sich **implizit**. Daraus ergibt sich quasi nebenbei ein dualer Optimierungsmechanismus, der neben der Aktienauswahl indirekt auch die Gewichtung vornimmt.

Die Erzeugung der Initialpopulation erfolgt durch die Generierung einer vordefinierten Anzahl an Chromosomen. Jedes Asset hat bei der Initialisierung eines Chromosoms eine Chance von 1/N pro Gen, ebendieses zu besetzen. N spezifiziert die Anzahl der möglichen

Assets - im hier gewählten Testsetting entspricht das Asset-Universum dem Inhalt des Dow-Jones-Index, d.h. N = 30. Zu betonen ist jedoch, dass die Ausgangsbasis vom Portfoliomanager frei gewählt werden kann, solange sie dem definierten Eingangsformat des Frameworks entspricht.

Um das Framework vielseitig verwendbar zu halten lässt sich die Zahl der Startchromosomen vor Beginn des genetischen Algorithmus über eine grafische Benutzeroberfläche festlegen. Auf das Interface über welches dieser Parameter gesteuert werden kann, wird im letzten Kapitel näher eingegangen.

5.2.2. Fitnessfunktion und Bewertungsansätze

Im Rahmen der hier vorgestellten prototypischen Implementierung werden zwei in der Finanzwirtschaft sehr häufig verwendete Kennzahlen herangezogen auf deren Basis die Bewertung eines Portfolios erfolgen kann.

Die vorgestellten Ansätze liefern beide eine reelle Zahl welche die Güte des Portfolios repräsentiert und daher auch gereiht werden kann - die vorgestellten Bedingungen werden von den implementierten Bewertungsfunktionen also durchgehend erfüllt. Die Bewertung der aus den einzelnen Assets zusammengesetzten Portfolios, kann in den hier vorgestellten Fällen mit höherem Initialaufwand durchaus exakt erfolgen und soll daher lediglich zur Illustrierung von Möglichkeiten dienen. Ein wichtiger Vorteil ist jedoch, dass durch das Setup des genetischen Algorithmus die Errechnung der Portfoliobewertungen an sich, vollkommen isoliert und daher modular gehalten wird. Die Fitnessfunktion kann daher beliebig kombiniert und erweitert werden, ohne dass ein Entwickler fundamental in den Gesamtprozess eingreifen müsste. Durch eine mögliche, individuelle Kombination von Bewertungsmaßen mit verschiedenen Gewichten die in Summe das Investitionsverhalten eines bestimmten Anwendungsbenutzers (bzw. Portfoliomanagers) optimal trifft, kann es durchaus zu NP-vollständigen Problemen kommen - dies muss bei der Erstellung der Fitnessfunktion dann aber in keinster Weise beachtet werden. Die Approximation des heuristischen Lösungsverfahrens wird in Summe ähnlich zuverlässige Ergebnisse liefern wie bei berechenbaren Kombinatorik-Problemen.

Bevor auf die eigentlichen Funktionen eingegangen werden kann, sollen zwei simple Methoden erläutert werden, mit welcher die Portfolio-Rendite zur Vereinfachung von weiterführenden Berechnungen üblicherweise standardisiert wird:

- **Logarithmierung:** Errechnet man die Rendite auf logarithmierter Basis, ergibt sich daraus die sogenannte stetige Rendite bzw. die prozentuale Veränderung des Wertes in einer gegebenen Periode. Diese logarithmierte Rendite lässt sich sehr einfach durch Addition kumulieren.

$$R_{ln} = \ln \frac{R_t}{R_{t-1}} = \ln R_t - \ln R_{t-1}$$

  Im Gegensatz zur diskreten Rendite kann die stetige Rendite als normalverteilt angesehen werden. Auf der Annahme der Normalverteilung von Asset-Renditen, welche durch die Logarithmierung erreicht wird, basieren in weiterer Folge sämtliche weiteren Berechnungen von Risikomaßen.

- **Annualisierung:** Durch die theoretische Streckung bzw. Hochrechnung einer entsprechenden Kennzahl auf ein Jahr, kann eine permanente Vergleichbarkeit von Rendite- und auch Risikokennzahlen gewährleistet werden, welche sich über unterschiedliche Zeiträume erstrecken.

$$\boldsymbol{R_{ann} = \sqrt{\frac{250}{N}} * R_N}$$

  Die Anzahl der Handelstage für ein Jahr wird üblicherweise mit 250 angegeben. **N** steht für die Anzahl der Tage über welche die angegebene Rendite erzielt wurde. Für die Standardabweichung kann dieselbe Logik angewandt werden.

Sämtliche Kalkulationen wurden also auf Basis einer annualisierten, stetigen Rendite berechnet.

5.2.2.1. Sharpe-Ratio

Die Sharpe-Ratio [32] ist ein weit verbreitetes Maß mit dessen Hilfe die Rendite eines Assets in Relation zur Schwankungsbreite der Rendite, also seiner Volatilität, gesetzt werden kann. Analysiert man diese zwei Komponenten daher im Verhältnis zueinander, kann ein weit besseres Bild von der Attraktivität einer Investmentmöglichkeit erzeugt werden (zumindest

retrospektiv), als die bloße Betrachtung nur eines dieser Werte. Am Ergebnis kann festgemacht werden, wie hoch die Überrendite pro Einheit des übernommenen Risikos ist. Risiko wird in seinen einfachsten Definitionen als direkt äquivalent zur Volatilität betrachtet, welche statistisch wiederum als eine Schwankung der Renditezeitreihe ausgedrückt werden kann. Da üblicherweise eine gewisse (theoretisch) risikofreie Rendite existiert, wird dieser zuvor festgelegte Renditesatz zu jedem Mess-Zeitpunkt der Asset-Rendite um diese bereinigt. Als risikofreie Rendite wird üblicherweise die Rendite eines Wertpapiers ohne Ausfallswahrscheinlichkeit herangezogen, wie beispielsweise eine 3-Monats-Schatzanleihe der US-Regierung (3-month US treasurybill) die zum 01.08.2012 bei historisch niedrigen 0.09% rentiert hat.

Für die Volatilität $\boldsymbol{\sigma_D}$ gilt nun

$$\boldsymbol{\sigma_D = \sqrt{\frac{\sum_{t=1}^{T}(D_t - \bar{D})^2}{T-1}}}$$

wobei $\boldsymbol{D_t}$ einen um die risikofreie Rendite bereinigten Renditewert („excessreturn“) zum Zeitpunkt $\mathbf{t}$ bezeichnet, also

$$\boldsymbol{D_t = R_t^a - R_t^f}$$

$\boldsymbol{\bar{D}}$ ist der Durchschnittswert der Zeitreihe über $\boldsymbol{D_t}$ in der gesamten Zeitspanne und $\mathbf{T}$ die Anzahl der gemessenen Zeitpunkte. Durch die Normierung (mittels Quadrierung) der Abweichungen heben sich negative und positive Schwankungen nicht auf, sondern werden als absolute Abweichungen berücksichtigt und zusammengeführt. Die eigentliche Sharpe-Ratio wird dann definiert als

$$\boldsymbol{SR \coloneqq \frac{\bar{D}}{\sigma_D}}$$

woraus sich ein einfach verwendbares und reihbares Maß zur Abschätzung eines Preis/Leistungs-Verhältnisses ergibt. Unten stehende Grafik zeigt an, für welche Preisentwicklungen sich besonders hohe und besonders niedrige Sharp-Ratios ergeben.

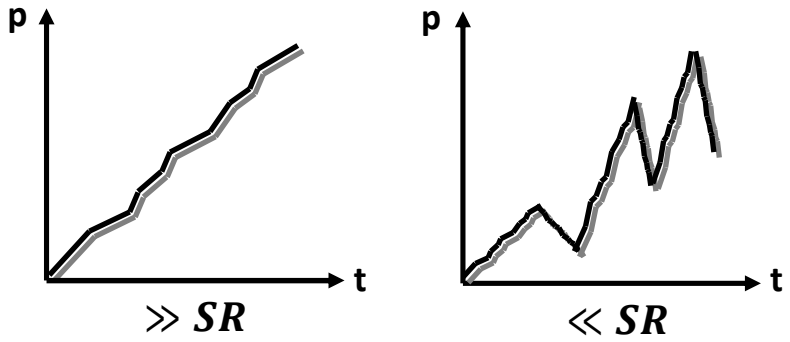

Abbildung 22 - Hohes Sharpe Ratio vs. niedriges Sharpe Ratio

5.2.2.2. Calmar-Ratio

Im Vergleich zum Sharpe-Ratio ist das Calmar-Ratio [33] ein sehr simples Maß, welches die Volatilität eines Assets nicht explizit in Betracht zieht. Es ist lediglich definiert durch das Ratio von annualisierter, um den risikofreien Return bereinigter Rendite $\bar{\boldsymbol{D}}$ und dem maximalen „drawdown“ $\boldsymbol{d}$ eines Assets über eine gegebene Periode.

$$\boldsymbol{CR} = \frac{\bar{\boldsymbol{D}}}{\boldsymbol{d}}$$

Diese Maßzahl ist ein wichtiger Anhaltspunkt für Investoren, für welche starke Kurseinbrüche problematischer erscheinen als eine konstant höhere Volatilität. Es liegt auf der Hand, dass sich beide Maße bis zu einem gewissen Grad deckungsgleich verhalten, da sich ein starker Kurseinbruch in der Gesamtvolatilität und daher auch im Sharpe-Ratio wiederspiegelt - das Calmar-Ratio legt jedoch andere Schwerpunkte und wird oftmals ergänzend zum Sharpe-Ratio betrachtet. Zum besseren Verständnis sollen zwei Beispielverläufe mit besonders niedrigen und hohen Calmar-Ratio gezeigt werden.

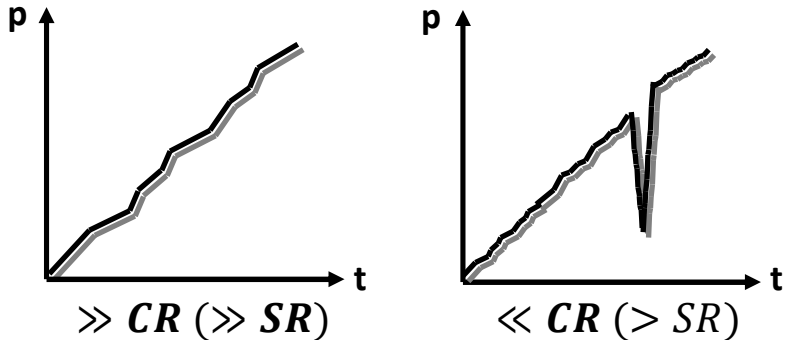

Abbildung 23 - Hohes Calmar Ratio vs. niedriges Calmar Ratio

Die Notierung des Sharpe-Ratio zum Vergleich (innerhalb der Klammer), soll den Unterschied verdeutlichen.

5.2.3. Selektions- und Rekombinationsmechanismus

Innerhalb der bewerteten Individuen muss nun eine Auswahl an Portfolios getroffen werden, die in den Rekombinationsmechanismus übergehen können. Die Bewertung der Portfolios geschieht in Abhängigkeit von der zuvor gewählten Fitnessfunktion. Sowohl die Selektion als auch die Rekombination sollten so simpel wie möglich und komplex wie nötig gewählt werden, um ein übliches Konvergenzverhalten zu erreichen.

Nach der Bewertung werden die Portfolios mittels ihrer Sharpe Ratio gereiht und anschließend die besten 25% des Portfolios ausgewählt – es wird also auf eine simple rangbasierte Selektion zurückgegriffen.

Der prozentuelle Wert der die Anzahl der übernommenen Individuen bestimmt, kann über die zur Verfügung gestellte Applikation dynamisch den Bedürfnissen angepasst werden. Zu beachten ist jedoch, dass sich eine Adaption des Parameters massiv auf den Optimierungsvorgang niederschlägt. Umso niedriger der gewählte Prozentsatz, umso stärker ist der bereits diskutierte Selektionsdruck was die Gefahr zu früher Konvergenz deutlich erhöht.

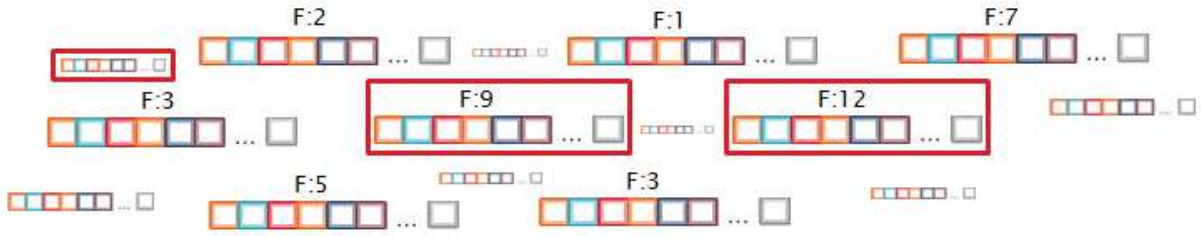

Abbildung 24 - Auswahl der besten Individuen

Nachdem eine kompetitive Subpopulation an Portfolios ausgewählt wurde, folgt die Rekombination der Einheiten. Durch den bereits zuvor erläuterten Kodierungsmechanismus, erfolgt die Gewichtung der Assets im Portfolio implizit über die Häufigkeit der vorhandenen Gene (=Assets). Es werden stets zwei Elternindividuen miteinander rekombiniert, die Auswahl des Elternpaares erfolgt auf Zufallsbasis – es werden also keinerlei Ähnlichkeiten bei der Zusammenstellung der Paare beachtet.

Um die Suche jedoch insgesamt zielgerichteter zu gestalten, werden äquivalente Gene der gewählten Einheiten gruppiert bzw. geordnet.

Abbildung 25 - Sortieren der Gene (Zusammenstellen gleicher Gene)

Nach der Ordnung der Gene, werden möglichst gleich große Genblöcke gegeneinander ausgetauscht um auch tatsächlich Portfolio-Eigenschaften gegeneinander zu tauschen und effektiv weiterzugeben.

Programmtechnisch wird bei der Erzeugung der Kinderindividuen erst jeweils ein Elternteil kopiert um in die Kopie dann eigene Genblöcke mit jenen des jeweils Anderen Elternteils zu ersetzen. In Summe wird also ein Multi-Point-Crossover mit Möglichkeit zur dynamischen Parameteranpassung implementiert.

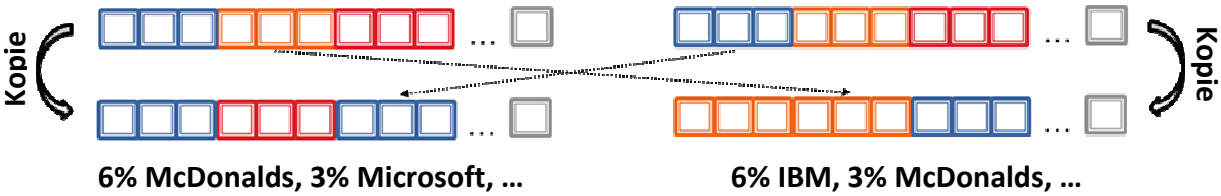

Abbildung 26 - Multi-Point-Crossover der Individuen

Die Population wird über die Generationen hinweg stets konstant gehalten, es stehen daher mit jeder Iteration dieselbe Anzahl an Individuen zur Verfügung wie bei der vorangegangenen.

Da die für die Rekombination wesentlichen Parameter in besonderer Abhängigkeit vom konkreten Anwendungsgebiet stehen, sind die Größe des Chromosomenpools, die Anzahl der Crossovers die bei einer Rekombination getätigt werden, sowie die Länge der auszutauschenden Genblöcke dynamisch einstellbar.

5.2.4. Mutation und Abbruchkriterium

Im Lauf des Algorithmus konvergiert der gesamte Portfolio-Pool gegen eine optimale Zusammensetzung an Aktienwerten das auch ein ausdünnen des Aktienpools impliziert. Um die Wahrscheinlichkeit einer zu frühen Konvergenz in ein lokales Optimum zu vermeiden, werden die Kinderindividuen einem Mutationsprozess unterworfen.

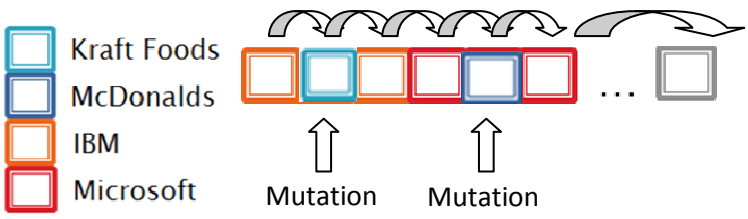

Abbildung 27 - Mutation einzelner Gene

Obenstehende Grafik zeigt den Ablauf. Jede Position innerhalb eines Portfolios wird durchlaufen und mit einer zuvor festgelegten Wahrscheinlichkeit mutiert. Tritt der Fall einer Mutation ein, wird mit uniformer Wahrscheinlichkeit eine der Positionen aus dem Asset-Pool gewählt und in das Portfolio integriert.

Die Mutationsrate $\boldsymbol{M}$ ist Teil der dynamischen Parametrisierung und wird vom Benutzer manuell über die implementierte Benutzeroberfläche festgelegt. Die Standardwahrscheinlichkeit beträgt 1%. Wie auch bei den Einstellungsmöglichkeiten der Rekombination, gilt bei der Mutation das Augenmaß zu behalten. Ist diese zu hoch, findet der Suchalgorithmus keine konkrete Richtung was auch am Konvergenzgraphen abzulesen ist. Eine zu niedrige Mutationsrate zeigt sich in einer sehr raschen Konvergenz und daher höchstwahrscheinlich suboptimalen Lösung.

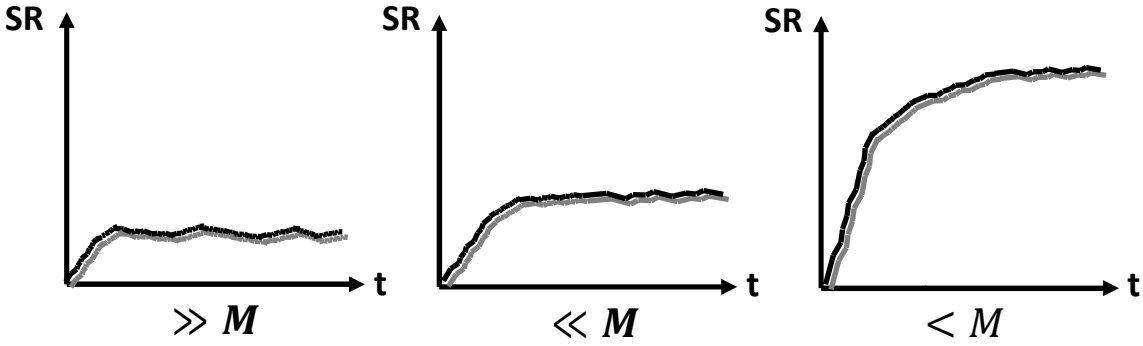

Abbildung 28 - Optimierungsgraphen mit unterschiedlichen Mutationsraten

Ebenso wie die Mutationsrate wird das Abbruchkriterium konstant initialisiert um die Möglichkeit einer einfachen Anpassung über die Anwendung zuzulassen. Der Algorithmus wird abgebrochen sobald die festgelegte Anzahl an Durchläufen erreicht wird, diese wurde für den Testaufbau mit 120 bestimmt und durch simples Probieren eruiert. Da der Anstieg in der Qualität eines Suchergebnisses bei jedem heuristischen Such- bzw. Optimierungsverfahren mit der Zeit abflacht, sollte die Zahl der Iterationen bei statischer

Einstellung so gewählt werden, dass bei mehrmaliger Wiederholung des Algorithmus zum Abbruchszeitpunkt keine signifikante Verbesserung mehr erzielt werden kann.

5.3. Backtesting- und Benchmarkingprozedur

Unter dem Begriff **"Backtesting"** versteht man im weitesten Sinn eine Überprüfung der über einen Trainingsbereich ermittelten Strategie auf seine Gültigkeit mit Hilfe eines historischen Datensatzes. Der beschriebene Testdatensatz diente bei der Entwicklung des Algorithmus sowohl zu Trainings- als auch zu Backtesting-Zwecken bezogen auf jeweils verschiedene Zeiträume. Der Testzeitraum kann auch hier über eine Benutzeroberfläche vollkommen autark gewählt werden. Da die gewählte Zeitreihe eine Vielzahl verschiedener Marktphasen enthält, kann die Validität einer eigenen Bewertungsfunktion über vollkommen konträre Marktphasen hinweg getestet werden. Ist die aus der Portfoliobewertung resultierende Aktienauswahl und Gewichtung valide, wird sich das auch in der Performance des Portfolios auf dem Backtesting-Zeitraum erkennbar machen. Anzumerken ist jedoch, dass sich hieraus selbstverständlich keinerlei allgemein gültige Handlungsanweisungen ergeben können. Da es sich um empirische Versuche handelt, können die hier gewonnen Erkenntnisse durchaus von großem Nutzen sein, müssen jedoch vom jeweiligen Portfoliomanager richtig interpretiert und in einen größeren Kontext gesetzt werden. Da die marktbestimmenden Parameter im Zeitverlauf in ihrer Zusammensetzung und Gewichtung einem stetigen Wandel unterliegen, benötigen die Ergebnisse eine korrekte Einordnung.

Da die Performance des optimierten Portfolios jedoch nichts über dessen Güte aussagt, muss diese über den Vergleich mit einem angemessenen Referenzportfolio ermittelt werden. Je nachdem wie stark das eigens optimierte Portfolio im selben makroökonomischen Umfeld neben ähnlichen Referenzportfolios bspw. in Bezug auf Performance und Volatilität abgeschnitten hat, ergibt sich daraus die eigentliche Güte.

Bei der Selektion von Aktien gibt es üblicherweise zwei Ansätze das eigens optimierte Portfolio zu vergleichen.

5.3.1. Standard-Index

Die Benchmark stellt ein standardisierter Index mit eigener, vorgegebener Gewichtung dar. Beliebte Vergleichsindizes für aktiv gemanagte Aktienfonds sind je nach Länderallokation die

Aktienindizes MSCI-World, S&P 500 bzw. der zuvor vorgestellte Dow Jones Index oder einer der EUROSTOXX-Indizes.

5.3.2. 1/N generischer Index

Eine der simpelsten und zugleich am schwierigsten zu schlagenden Benchmarks repräsentiert ein generischer Index der zu gleichen Teilen aus all jenen Komponenten besteht, welche im Auswahlpool des Optimierungsalgorithmus verfügbar sind. Jeder Wert wird dabei zu gleichen Teilen in den Index genommen und je nach Umsetzungsvariante entweder im Zeitablauf neu gewichtet oder "laufen" gelassen. Durch die unterschiedliche Preisentwicklung der einzelnen Werte erfolgt eine natürliche Verschiebung der Gewichte über die Zeit - soll diese Verschiebung korrigiert werden (stark gestiegene Aktien verkaufen, billiger gewordene Aktien nachkaufen), muss eine Adaptierung der Stückzahlen der Einzelwerte erfolgen.

Beide Alternativen sind im Rahmen der Implementierung umgesetzt worden - es steht daher sowohl die originäre DowJones-Indexreihe wie auch ein generischer Index zur Verfügung, der zu gleichen Teilen aus den dargestellten Elementen besteht. Dieser wird während der Backtesting-Dauer jedoch nicht umgewichtet um die Benchmark bewusst kompetitiver zu gestalten und an die Handhabung des optimierten Portfolios anzupassen. Da bei einer Umgewichtung eines Portfolios in der Regel hohe Spesen (im Wert von ca. 1% des Gesamtportfolios) anfallen, sowie dadurch auch implizit eine nachgewiesen weniger optimale Handelsstrategie implementiert werden würde (Stärke verkaufen, Schwäche kaufen) - verläuft der Vergleichsindex ebenso wie das durch den genetischen Algorithmus optimierte Portfolio statisch.

5.4. Ergebnisse, Erkenntnisse und Kritik

Um das zuvor beschriebene Framework und die dahinterliegende Methode auf seine Tauglichkeit zu überprüfen, wurde auf Basis der vorgestellten Daten eine Reihe von Testläufen generiert. Wichtig ist allerdings, dass die aus den Testläufen generierten Portfolioergebnisse auf Basis der vorgestellten Optimierungsfunktionen keine direkte Aussagekraft über die Validität der Funktionen in Bezug auf ihre allgemeingültige Prognosefähigkeit bzgl. Performance und Volatilität eines dadurch erstellten Portfolios erlauben. Die dargestellten Szenarios stellen also vor allem Funktionstests des im Rahmen

dieser Diplomarbeit entwickelten Prototyps eines erweiterbaren GA-Frameworks dar und sollen die bestehenden Analysemöglichkeiten und den durch die Software ermöglichten Analyseprozess näher erläutern.

Sowohl die Tests als auch die dabei erzeugten Ergebnisse und aufgekommenen Probleme der heuristischen Optimierung sollen nun in weiterer Folge näher beschrieben werden. Im ersten Schritt wird ein erster Versuchsaufbau mit Trainings- und Backtestingperiode erzeugt –unten stehende Grafik zeigt einen der ersten Schritte nach der Spezifikation der Basisdaten die dafür notwendig sind.

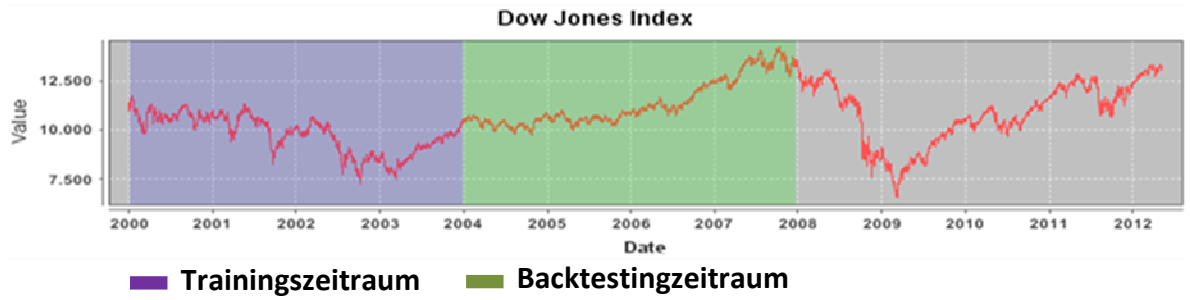

Abbildung 29 - Beispielhafter Trainings- und Backtestingzeitraum

Der Trainingszeitraum (blau markiert) reicht von Anfang 2000 bis Ende 2003 und beträgt damit vier volle Jahre. Der gewählte Bereich der zur Generierung des Portfolios genutzt wird, fällt in eine Bärenmarktphase. Der Performancevergleich mit den entsprechenden Benchmarks erfolgt direkt anschließend über eine Bullenmarktphase von Anfang 2004 bis zum Start der Finanzmarktkrise Ende 2007. Der Gen-Pool setzt sich aus der Gesamtheit der Einzelaktien des DowJones-Index zum Stand 01.01.2000 zusammen. Die Performance- und Volatilitätsberechnungen erfolgen wegen des ansonsten zu hohen Rechenaufwandes auf wöchentlicher Basis.

Die Parametrisierung des Algorithmus zum Startzeitpunkt kann der unten stehenden Tabelle entnommen werden.

Wert	Parameter
100	Mutationswahrscheinlichkeit pro Gen in Promille
120	Anzahl der Iterationen
50	Anzahl der Chromosomen
3	Anzahl der Crossover-Punkte
10	Länge (Genanzahl) der Crossover-Teile
25%	%-Anteil der Top-Chromosomen die für nächste Iteration gewählt werden

Tabelle 5 - Beispielparameter des genetischen Algorithmus

Zu betonen ist jedoch, dass die Wahl der Parameterwerte lediglich das scheinbar optimale Ergebnis mehrerer Versuche mit dem vorliegenden Test-Set repräsentiert. Hierfür wurde im Speziellen der Vorgang der Portfoliooptimierung und das Verhalten sowie die Form des Optimierungsgraphen beobachtet. Eine derart dynamische Verwendung eines genetischen Algorithmus erfordert also zwingend eine vom Benutzer geführte Testphase in welcher die besten Parameterwerte eigenständig ermittelt werden müssen. Unten erscheint das Ergebnis auf Basis des vorgestellten Algorithmus und der angegebenen Parameter - die gewählte Darstellung soll den zuvor angesprochenen Optimierungsfortschritt verdeutlichen.

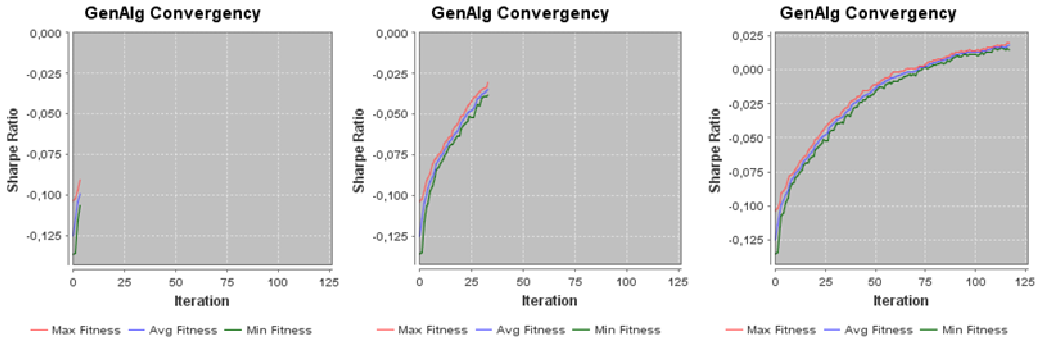

Abbildung 30 - Konvergenzgraph des genetischen Algorithmus

Achse die aktuelle Sharpe-Ratio. Die eingezeichneten Graphen beschreiben die Entwicklung der Fitness des stärksten (in rot) und des schwächsten (in grün) Portfolios der Population, sowie der mittleren Fitness der Gesamtpopulation (in blau).

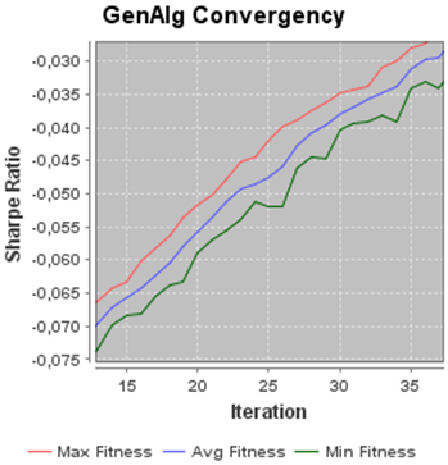

Abbildung 31 - Beispielhaft Max / Avg / Min Fitness innerhalb der Population

Diese Darstellung erlaubt einen guten Einblick in die Diversität des Individuenbestandes. Ist die genetische Vielfalt rasch sehr gering, verengt sich der Abstand der einzelnen Portfolios ebenfalls schnell. Diese Entwicklung erfolgt oftmals in Zusammenhang mit früher Konvergenz und daher vorzeitigem Abflachen des Optimierungsgraphen. Ist das der Fall, sollten die Parameter in Entsprechung ihrer fundamentalen Wirkung adaptiert werden (z.B. Erhöhung der Mutations- oder Selektionsrate).

Das oben gezeigte Ergebnis entspricht jedoch einem üblichen Optimierungsverlauf und zeigt, dass das gewählte Setup für die Individuenkodierung sowie der implementierte Rekombinationsalgorithmus funktionsfähig und valide zu sein scheinen. Das eigentliche Optimierungsergebnis, welches bisher eher abstrakt definiert und behandelt wurde, wird in unten stehender Grafik gezeigt. Die Skala der Y-Achse zeigt dabei an, in welcher Iteration sich das aktuell beste Portfolio der Population befindet, das Tortendiagramm zeigt die Portfolioallokation an. Zu Beginn sind auch die besten verfügbaren Portfolios noch extrem stark diversifiziert und daher auch die genetische Vielfalt sehr breit. Im Laufe der Zeit beginnen sich erste größere Blöcke zu formieren und den Vorgang zu dominieren – die in diesem Stadium resultierenden Portfolios verändern sich nur noch verhältnismäßig geringfügig bis weitestgehend Stagnation eintritt. Die oben gezeigten Optimierungsgraphen stellen die jeweiligen Entsprechungen zu den Tortendiagrammen dar. Die nur noch geringfügigen Veränderungen in der Portfolioallokation in der späteren Phase, stimmen also mit der immer flacher werdenden Steigung der Fitness überein.

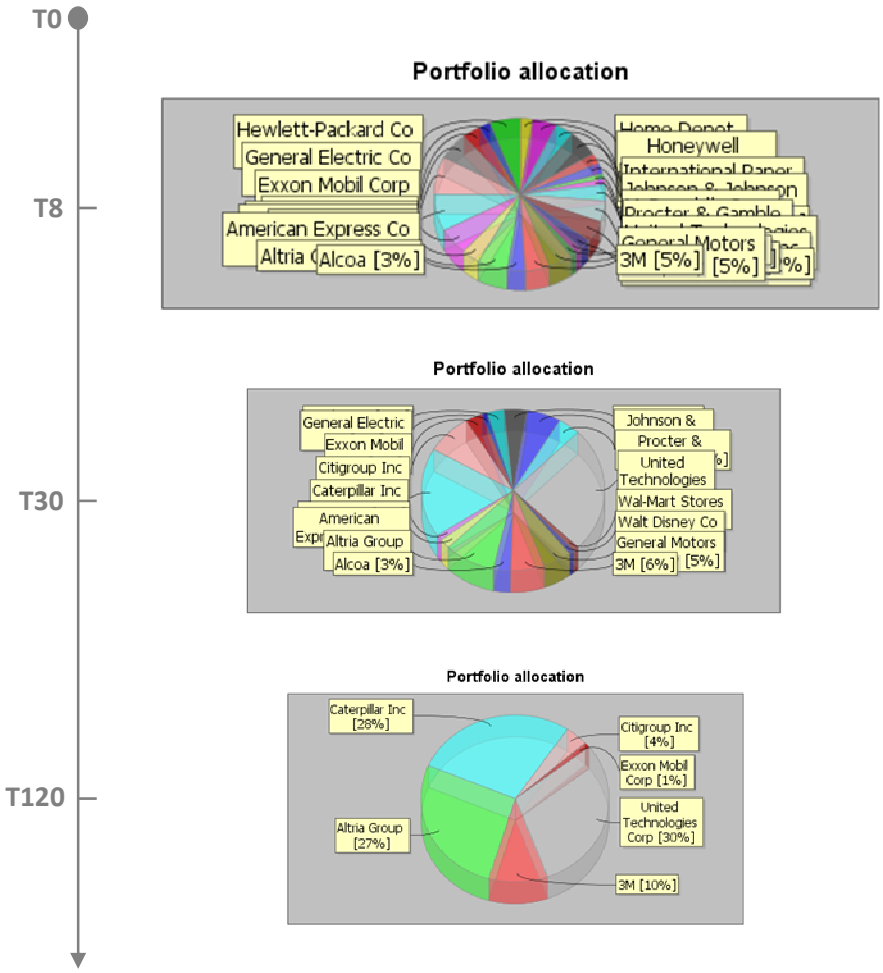

Abbildung 32 - Portfolioallokation über die Zeit

Die fertige Portfolioallokation für den dargestellten Trainingszeitraum von Anfang 2000 bis Ende 2003 ist letztendlich wie folgt:

Gewichtung	Asset
30%	United Technologies Corp
28%	Caterpillar Inc
27%	Altria Group
10%	3M
4%	Citigroup Inc
1%	Exxon Mobil Corp

Tabelle 6 - Resultierende Gewichtung der Assets

Das für den Zeitraum per Sharpe-Ratio optimierte Portfolio soll nun über unterschiedliche Testzeiträume bezüglich Performance und Volatilität gegen ein synthetisches Vergleichsportfolio mit einer Portfoliogewichtung von $\frac{1}{N}$ pro Asset getestet werden. Weiters

wird die Entwicklung jenes Index als Vergleich herangezogen aus dessen Einzelwerten sich das optimierte Portfolio ergeben hat. Unten stehende Grafik zeigt das Backtesting-Ergebnis für den oben angegebenen Zeitraum von Anfang 2004 bis Ende 2007. Die in weiterer Folge angeführten Liniendiagramme zeigen die reine Performanceentwicklung sämtlicher Portfolios. Die X-Achse beschreibt dabei die zeitliche Komponente, die Y-Achse die Renditeentwicklung in %, wobei der Bezugspunkt zur Berechnung der Rendite jeweils als der erste angezeigte Datenpunkt definiert ist.

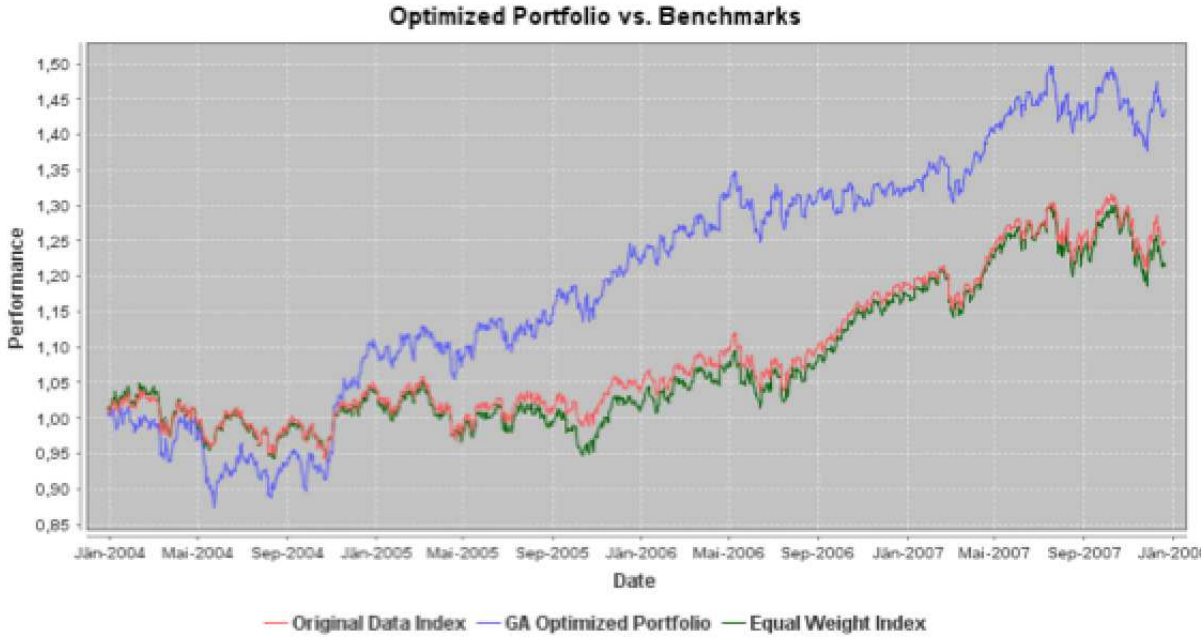

Abbildung 33 - Entwicklung der Vergleichsindizes 2004-2007

In Summe zeigt sich ein deutlich optimaleres Ergebnis im Vergleich zu Benchmarks desselben Bezugspunktes.

Performance	Sharpe-Ratio	Portfolio
43%	-0.13	GA-Optimized Portfolio
25%	-0.86	Original Data Index
22%	-0.9	Equal Weight Index

Tabelle 7 - Messzahlen der Indizes von 2004-2007

In positivem Börsenumfeld scheint eine Outperformance des Portfolios im Vergleich tendenziell gegeben. Da das Hauptinteresse der meisten Portfoliomanager jedoch in der Analyse des Portfolioverhaltens in grundlegend unterschiedlichen Marktphasen liegt, soll dieselbe Asset-Auswahl nun also über die zuvor ausführlich beschriebenen, anschließenden Zeiträume bis Mitte 2012 getestet werden.

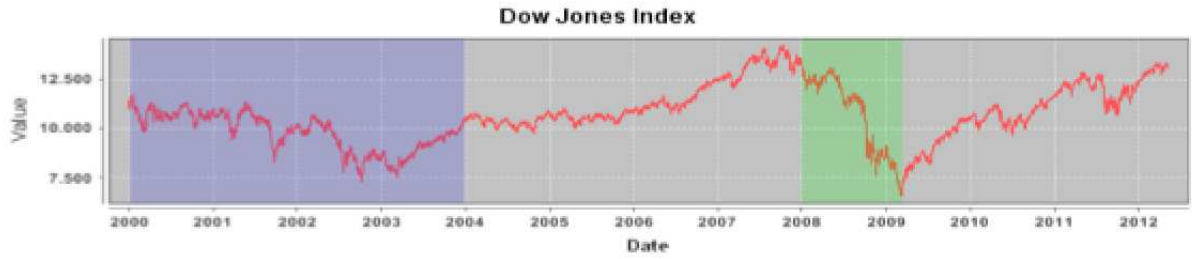

Abbildung 34 - Backtesting im Zeitraum der Finanzkrise 2008 / 2009

Zu beobachten ist, dass sich vor allem in außergewöhnlichen Marktphasen wie von Anfang 2008 bis März 2009, sehr interessante Effekte ergeben.

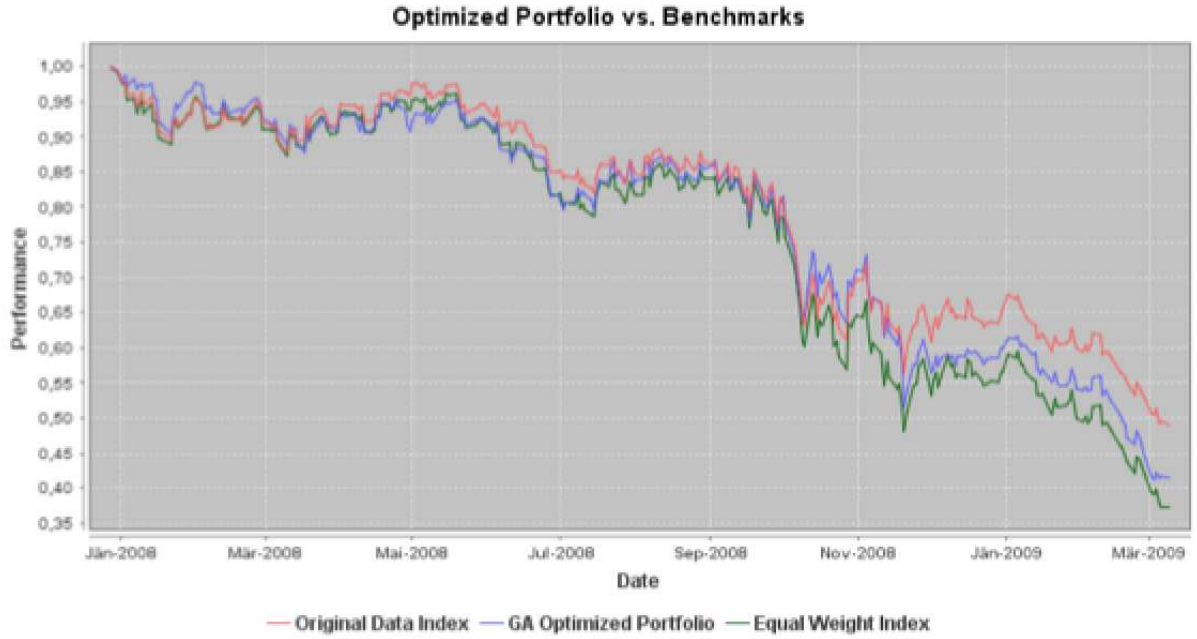

Abbildung 35 - Entwicklung der Vergleichsindizes 2008 / 2009

Die Optimierung ergibt keinerlei Vorteile im Indexvergleich – ein struktureller Bruch scheint sämtliche Assets unabhängig von ihrer Güte in ähnlicher Höhe abzustrafen. Die konkret errechneten Vergleichswerte ergeben ein ähnliches Bild wie eine intuitive optische Beurteilung des oben angezeigten Charts.

Performance	Sharpe-Ratio	Portfolio
-51%	-6.21	Original Data Index
-58%	-6.66	GA-Optimized Portfolio
-63%	-7.07	Equal Weight Index

Tabelle 8 - Messzahlen der Indizes von 2008 / 2009

in ähnlichem Muster wie auch während der ersten Backtesting-Phase, verläuft nun auch der Vergleich über den anschließenden Zeitraum der Erholung der Asset-Preise insgesamt. Die Phase verläuft von April 2009 bis Mai 2012.

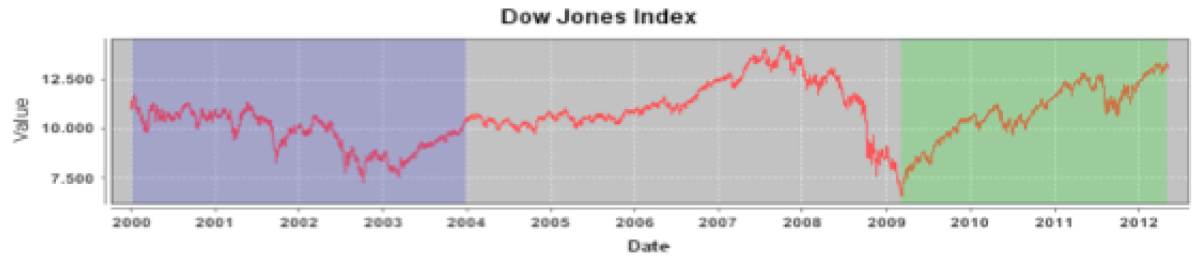

Abbildung 36 - Backtesting im Zeitraum der Erholung 2009-2012

Auf Grund der Stärke der Aufwärtsbewegung ergibt sich in Summe auch ein deutlich höheres „Premium“ des optimierten Portfolios zu den Vergleichsportfolios.

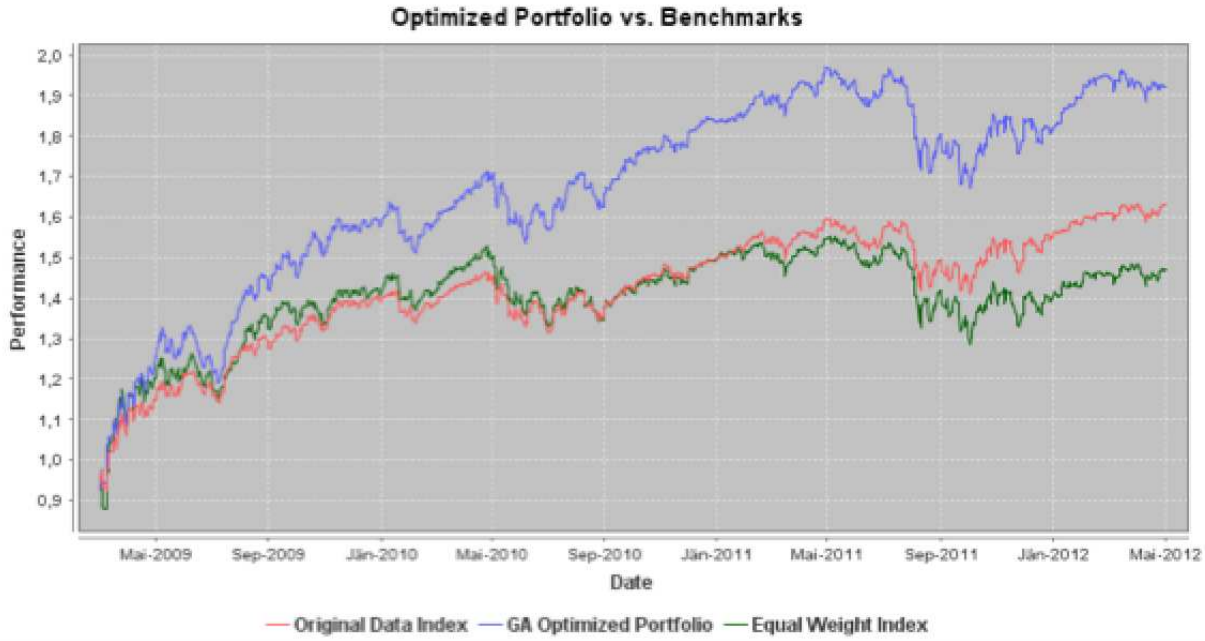

Abbildung 37 - Entwicklung der Vergleichsindizes 2009-2012

Die Ergebnisse sind in diesem Fall in Summe bis zu 50% besser als jene des Dow-Jones Index. Deutlich erkennbar ist eine kontinuierliche Outperformance über den selektierten Zeitraum.

Performance	Sharpe-Ratio	Portfolio
92%	1.49	GA-Optimized Portfolio
63%	1.02	Original Data Index
47%	1.47	Equal Weight Index

Tabelle 9 - Messzahlen der Indizes von 2009-2012

Eine abschließende Betrachtung der Portfolioentwicklung über den gesamten Zeitraum gibt einen Überblick über sämtliche Marktphasen hinweg.

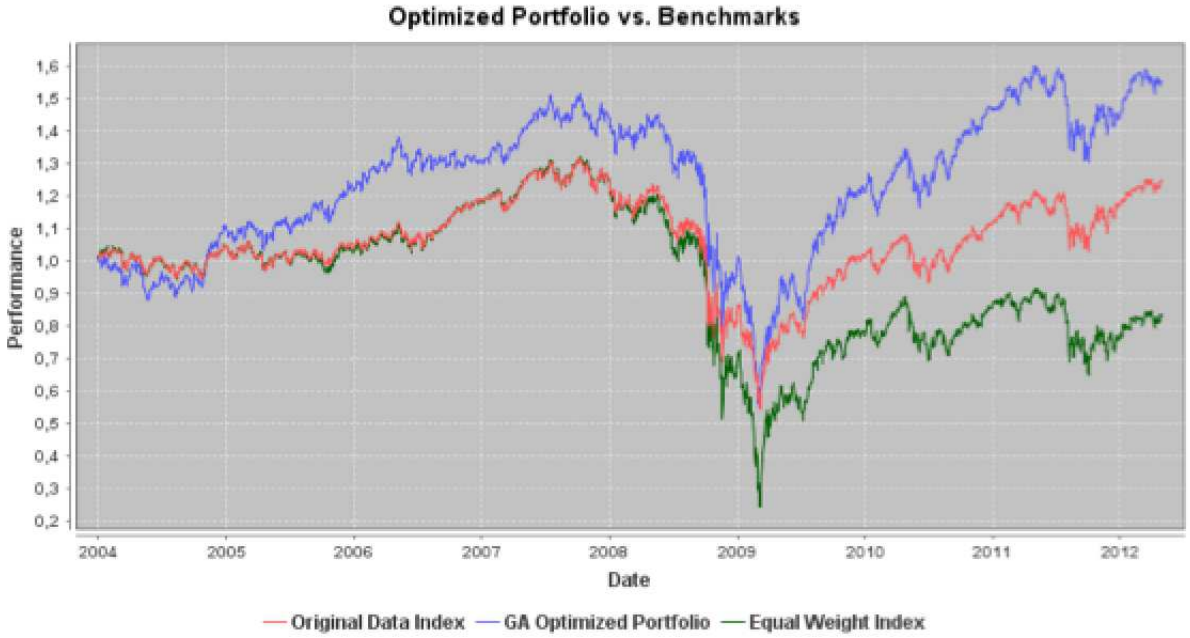

Tabelle 10 - Optimiertes Portfolio vs. Benchmarks über gesamten Zeitraum

Auffällig schwach verhält sich das $\frac{1}{N}$ Portfolio, welches sich aus den Bestandteilen des Dow-Jones Index zum 01.01.2000 zusammensetzt. Das liegt vor allem an der außergewöhnlichen Marktsituation ab 2008 in welcher eine Reihe großer Industriebetriebe und Banken (z.B. General Motors, Citibank) verstaatlicht wurden und daher kurz vor dem Konkurs standen. Jene Werte auf denen der Dow-Jones Index basiert, haben sich in dieser Zeit besonders stark in Hinblick auf die Zusammensetzung und Gewichtung verändert. Der seit Anfang 2000 gleichgewichtete Index beinhaltet daher eine Reihe von Werten welche dauerhaft über 90% ihres Börsenwertes verloren haben – folglich wird dieser entsprechend stark negativ verzerrt.

Performance	Sharpe-Ratio	Portfolio
55%	-0,94	GA-Optimized Portfolio
25%	-1,25	Original Data Index
-17%	-1,44	Equal Weight Index

Tabelle 11 - Messzahlen der Indizes von 2004-2012

Üblicherweise sind konstant gleichgewichtete Indizes auf Dauer eine kaum zu schlagende Benchmark. Ähnliche Resultate ergeben sich beim selben Test-Setup (d.h. mit gleicher Trainings- und Backtesting-Zeitraum) bei der Optimierung mit Hilfe des Calmar-Ratios.

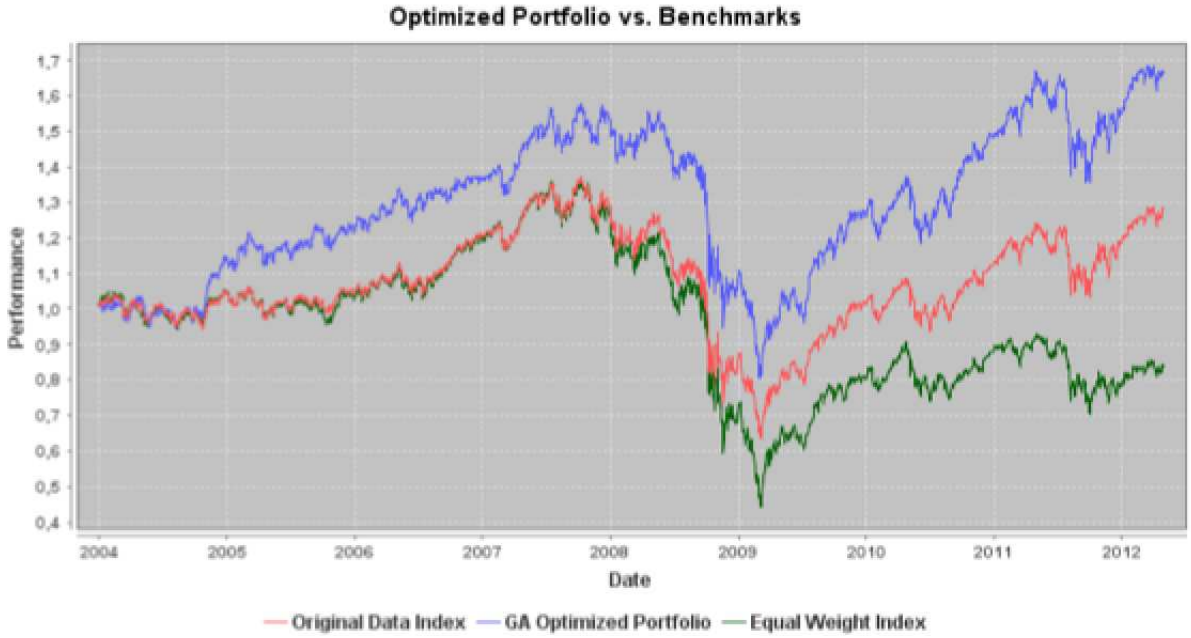

Abbildung 38 - Optimierung mittels Calmar Ratio

Das dazugehörige Portfolio ist bei gleicher Anzahl an Iterationen jedoch deutlich stärker fragmentiert und würde auf Grund beständig hoher Transaktionskosten eine real wesentlich geringere Rendite bringen als dies den Anschein im gezeigten Testszenario macht.

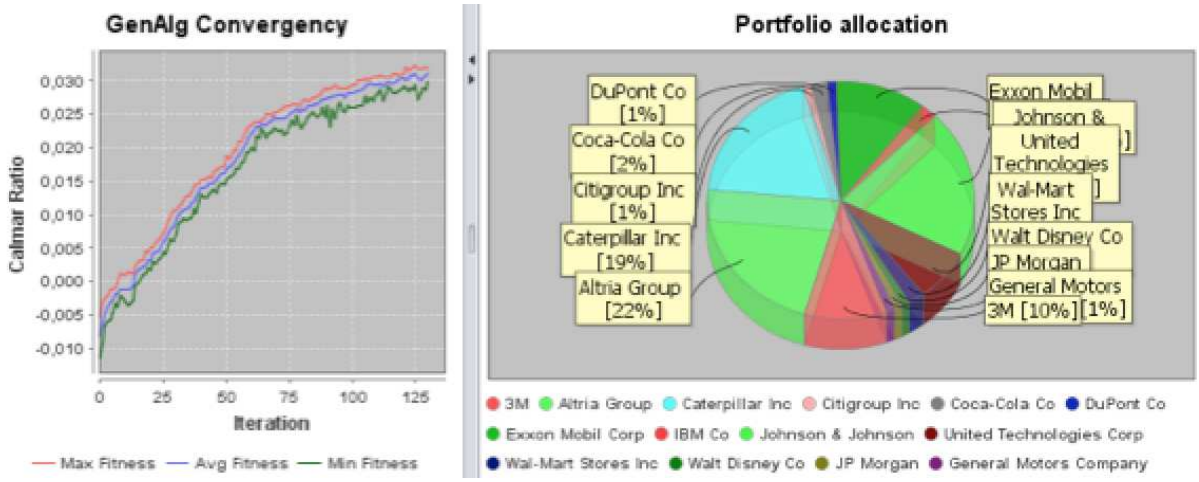

Abbildung 39 - Konvergenzgraph und Portfolio-Allokation mit Calmar Ratio

Problematisch erscheinen auch die in Summe deutlich unruhiger verlaufenden, sowie stärker auseinanderliegenden Konvergenzgraphen. Dies deutet darauf hin, dass das Calmar-Ratio für sich genommen, keine hinreichende Bewertungsfunktion darstellt mit deren Hilfe auf ein dominantes globales Optimum im Testzeitraum optimiert werden kann. Der Gedanke liegt also Nahe, dass eine ausschließliche Optimierung eines Portfolios über die Rendite und eine Ratio-Bildung mit dem „maximumdrawdown“ möglicherweise kein hinreichend starkes Kriterium darstellt.

Ein weiteres Optimierungsbeispiel welches mit Hilfe der zuvor präsentierten Testdaten erzeugt wurde, ist unten angefügt.

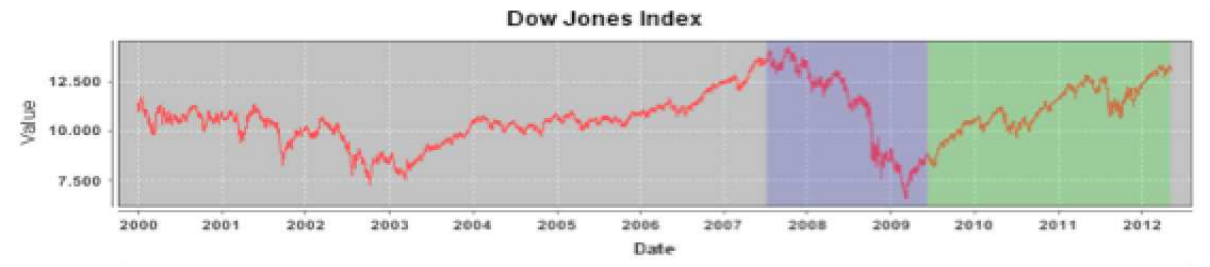

Abbildung 40 - Optimierung über systemkritische Phase 2008/2009

Getestet wurde, ob die implementierte Optimierung mittels Sharpe-Ratio über eine systemkritische Phase hinweg in der Lage ist, ein Portfolio zu erzeugen welches selbst in der anschließenden Erholung, Vorzüge gegenüber der bereits erwähnten Vergleichsindizes besitzt. Die ermittelte Zusammensetzung und Gewichtung ist auch jene welche Intuitiv zu erwarten wäre – vor allem etablierte und defensiv geführte Großunternehmen mit nur gering zyklischem Geschäftsmodell finden sich hier wieder.

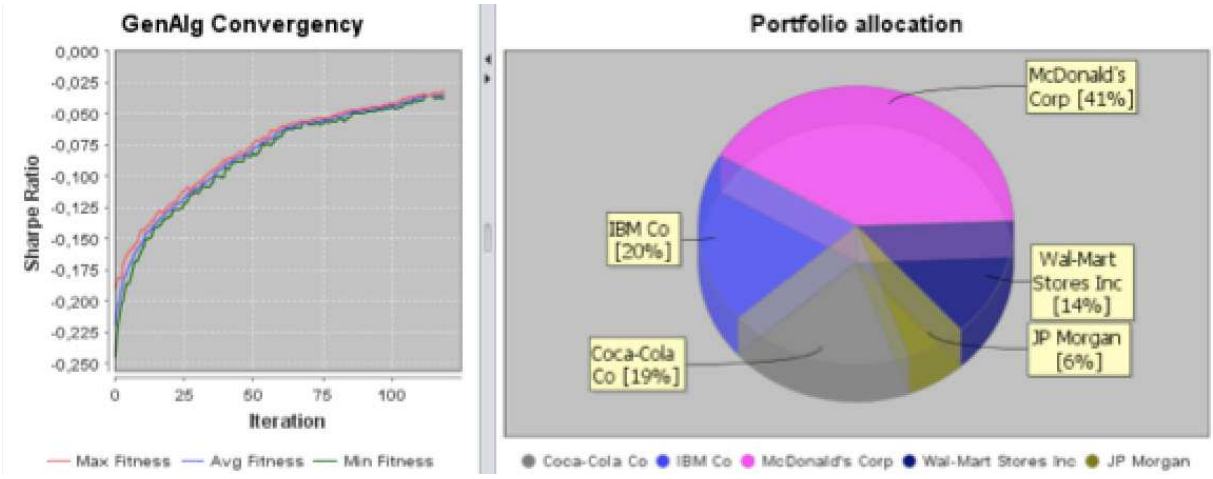

Abbildung 41 - Konvergenzgraph und Portfolio-Allokation

Das resultierende Ergebnis im Benchmark-Vergleich, zeigt ein Portfolio mit zwar nur geringfügig besserer Performance, jedoch deutlich geringerer Schwankungsbreite.

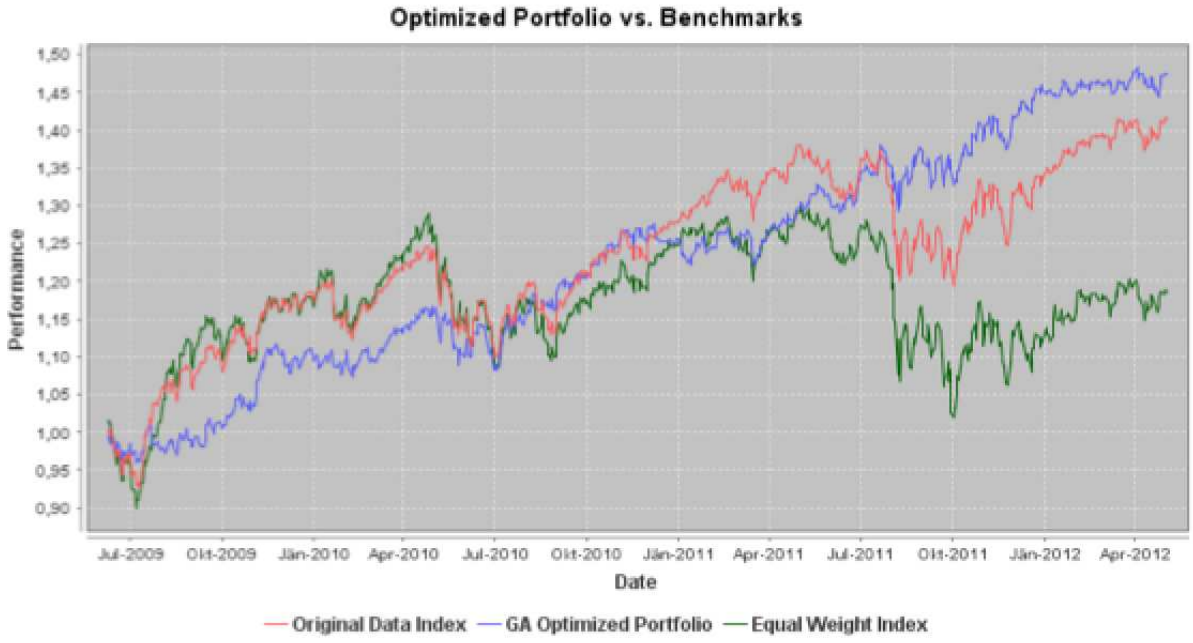

Abbildung 42 - Optimiertes Portfolio vs. Benchmarks

Dieser Umstand wird in der angeführten Tabelle deutlich. Die Kombination aus besonders guter Performance bei sehr geringer Volatilität ergibt ein deutlich besseres Sharpe Ratio des optimierten Portfolios bei ähnlicher Performance im Vergleich zum Dow-Jones Index bzw. „Original Data Index“.

Performance	Sharpe-Ratio	Portfolio
48%	1,00	GA-Optimized Portfolio
41%	0,45	Original Data Index
19%	-0,33	Equal Weight Index

Tabelle 12 - Kennzahlen der Portfolios im Vergleich

Die spezielle Problematik von möglichen Totalverlusten in einzelnen Anlagewerten, führt direkt zu einem Hauptproblem bei Portfoliooptimierungen mittels genetischer Algorithmen bzw. heuristischen Optimierungsverfahren im Allgemeinen. Da ein erzieltes Ergebnis der optimalen Lösung im Normalfall nur Nahe kommt, existieren tendenziell eine Vielzahl von Lösungsvarianten die der Algorithmus auf Basis der historischen Testdaten als annähernd gleichwertig bewertet.

Aus den vorangegangenen Erläuterungen sollte in Summe klar hervorgehen, dass die vorliegende Arbeit lediglich den Grundstein für ein breiter aufgestelltes Framework und

einen diffizileren Algorithmus liefern soll. Es existieren unzählige noch unbedachte Möglichkeiten, sowohl die dargestellte Vorgehensweise zu verbessern, als auch die Methodik zu erweitern.

6. Implementierung einer dynamischen Testumgebung

Auf Basis des zuletzt vorgestellten Algorithmus sollte in weiterer Folge eine dynamische Applikation entwickelt werden, die diesen zugänglich und adaptierbar macht. Dabei sollte es möglich sein die beschriebenen Parameter über eine simple Benutzerschnittstelle anzupassen. Eine einfache Datenschnittstelle soll einem potentiellen Portfoliomanager die Option geben, die Anwendung zur Portfoliooptimierung und Analyse zu verwenden. Das Programm wurde in Java 6.0 als eigenständig lauffähige „Fat-Client“ Anwendung entwickelt. Zur Darstellung der Ergebnisse wurde auf der unter LGPL stehenden Bibliothek JFreeChart zurückgegriffen und deren Komponenten erweitert oder für eigene Zwecke adaptiert.

6.1. Eingabedaten und Schnittstelle

Die Eingabedaten mit denen die Anwendung nach Programmstart initialisiert werden kann, müssen gewisse Rahmenbedingungen erfüllen um einen reibungsfreien Ablauf zu gewährleisten. Werden die unten stehenden Bedingungen nicht erfüllt, kann das zu unerwarteten Ergebnissen oder Fehlermeldungen führen.

6.1.1. Struktur

Die Eingabedaten müssen als „commaseparatedfile“ zur Verfügung gestellt werden. Im konkreten Fall wird ein Datensatz in folgender Form erwartet:

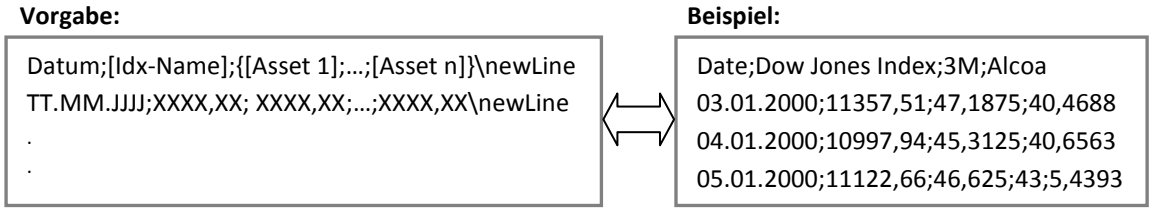

Abbildung 43 - Schema und Beispiel Importdatei

Spalte wird stets das als Index zu klassifizierende Asset erwartet und in der Anwendung selbst auch automatisch so verwendet. Auf Basis des Index erfolgt schließlich das Backtesting im entsprechenden Programmbereich (siehe Punkt X). Die Asset-Werte können beliebig viele Nachkommastellen aufweisen, sollten als Kommata jedoch einen Beistrich verwenden.

6.1.2. Synchronisation

Zu beachten ist vor allem die Synchronität der Zeitpunkte. Da nur eine einzige Datumsspalte zulässig ist, müssen die Werte der Assets dem jeweiligen Datum der Zeile entsprechen.

6.1.3. Vollständigkeit

Es erfolgt keine automatische Interpolation der Daten oder Standardisierung der Zeitpunkte um keine ungewollte Verfälschung der Ergebnisse zu provozieren. Fehlen also Datenpunkte für die angegeben Zeitpunkte, sollten diese nach beliebiger Methode ergänzt werden bevor die Datenbasis zur Optimierung herangezogen wird.

6.2. Benutzeroberfläche

In weiterer Folge werden die einzelnen Komponenten der Applikation vorgestellt. Die unten stehende Abbildung zeigt die Benutzerschnittstelle nachdem die Ursprungsdaten importiert und das Optimierungsverfahren durchgeführt wurde.

Abbildung 44 - Benutzeroberfläche des Prototyps

Der Prototyp ist also im Wesentlichen in die in weiterer Folge erläuterten Punkte untergliedert.

6.2.1. Auswahl der Basisdaten

Der Pfad zum Basisdatensatz kann entweder direkt über das zur Verfügung gestellte Textfeld angegeben werden oder über eine zusätzliche Benutzeroberfläche.

Diese kann über einen Klick auf „…“ aktiviert werden kann.

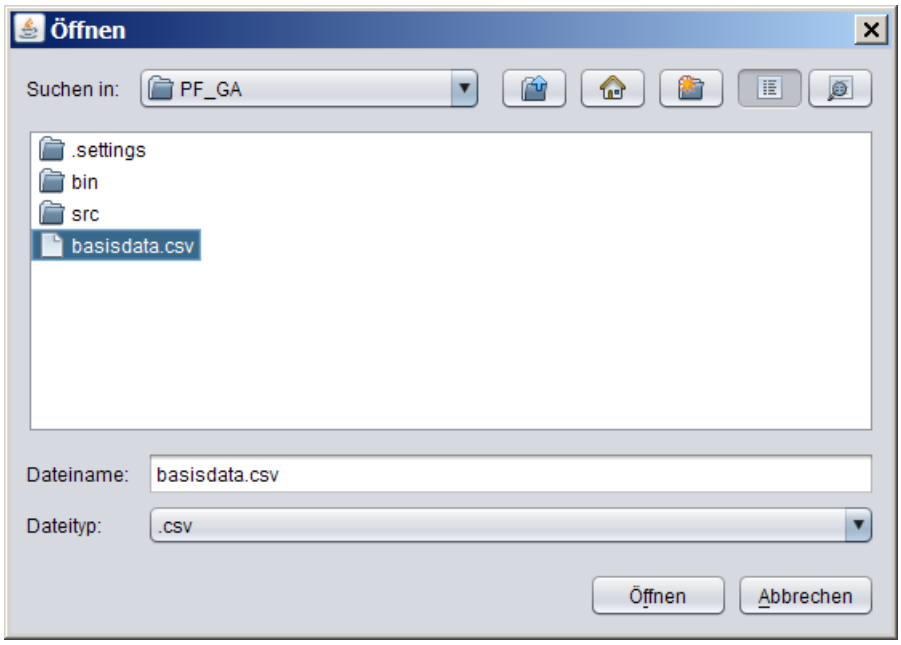

Abbildung 45 - Importieren eines Datensatzes

Durch das Öffnen der Datei wird der Pfad für den Import bereitgestellt.

6.2.2. Auswahl der Training und Backtesting-Bereiche

Nachdem der Datensatz importiert wurde, wird der Index der Basisdaten als Liniendiagramm geplottet. Initial wird dem Benutzer ein vorselektiertes Intervall grau hinterlegt und als „Training-Range“ angeboten. Über den darunterliegenden Auswahlbalken kann das Intervall dynamisch eingestellt werden. Das Datum der jeweiligen Bezugsfelder ändert sich mit der Veränderung des Ortes der Knöpfe am Auswahlbalken.

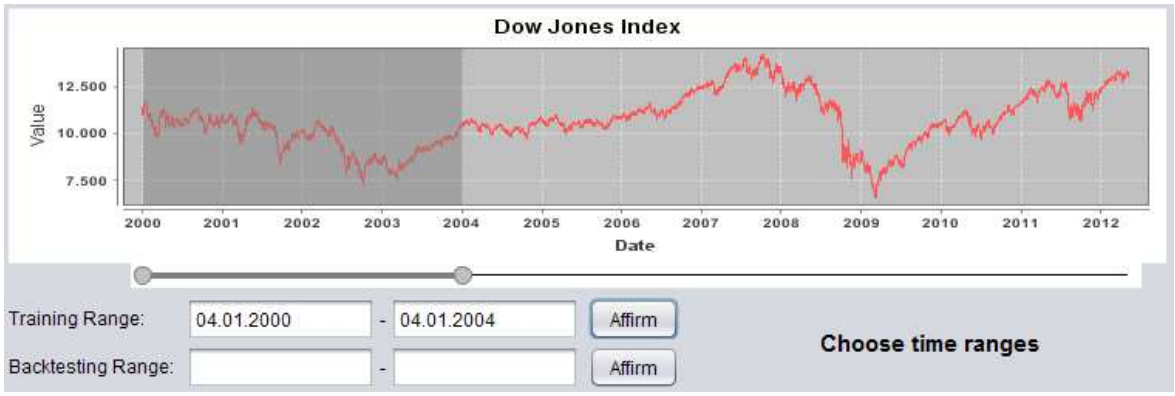

Abbildung 46 - Fixierung der Zeitbereiche

Soll ein Bereich fixiert werden, kann das über den „Affirm“-Button erfolgen.

Durch die Fixierung des Trainingsintervalls wird der markierte Bereich farblich gekennzeichnet. Zugleich verändert man auch den Bezug des Auswahlbalkens und bezieht sich in weiterer Folge auf den zu wählenden Backtesting-Bereich.

Abbildung 47 - Fixierter Trainings- und Backtestingzeitraum

Wie das zuvor erläuterte Intervall, lässt sich auch der Backtesting-Bereich vollkommen beliebig festlegen (aufeinanderfolgend, überschneidend, etc.). Sind beide Intervalle festgelegt, folgt die Parametrisierung des genetischen Algorithmus.

6.2.3. Konfiguration des genetischen Algorithmus

In den Algorithmus kann über die zur Verfügung gestellten Optionen weitreichend eingegriffen werden. Als zentrale Fitness-Funktion stehen sowohl das Sharpe-Ratio als auch das Calmar-Ratio zur Verfügung. Die Mutationswahrscheinlichkeit ist in Promille angegeben und bezieht sich auf jeweils ein einzelnes Gen. Die Voreinstellungen erzielen bei den im Rahmen der Arbeit analysierten Daten bei beliebigen Zeiträumen sinnvolle Ergebnisse.

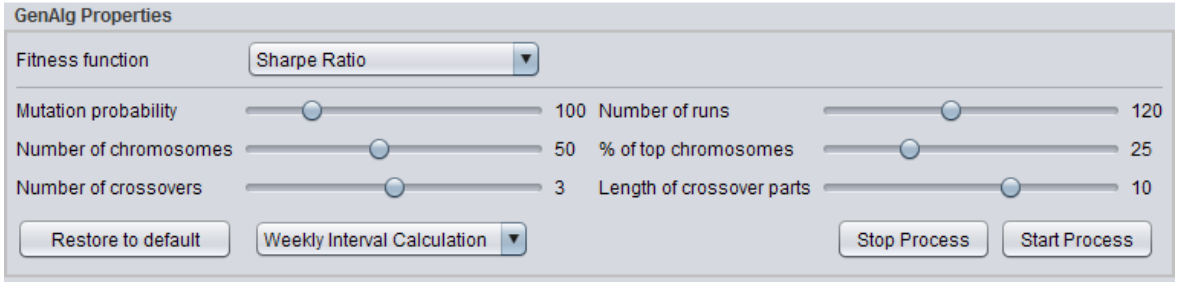

Abbildung 48 - Konfiguration der Parameter

Wird die Anzahl der zu analysierenden Assets massiv erweitert (bspw. von 20 auf 40), wird jedenfalls eine Anpassung der Anzahl der Chromosomen (daher der möglichen Startportfolios) notwendig um sicherzustellen, dass bei der Zufallsinitialisierung auch tatsächlich alle möglichen Asset-Kombinationen erzeugt werden. Als mögliches Berechnungsintervall steht eine Berechnung auf Tagesbasis wie auch auf Wochenbasis zur Verfügung.

Der Prozess kann jederzeit abgebrochen werden, das resultierende Ergebnis entspricht dann dem besten zum Zeitpunkt des Abbruchs ermittelten Portfolio. Wird der Prozess neu

gestartet, beginnt das gesamte Optimierungsverfahren von vorne. Da sämtliche Berechnungsprozesse im Rahmen eigener Threads implementiert sind, ist die Benutzeroberfläche zu keiner Zeit blockiert - eine Anpassung der Algorithmus-Parameter während der Laufzeit ist jedoch nicht vorgesehen.

6.2.4. Visualisierung der Optimierungsanalyse

Zur Optimierungsanalyse werden zwei verschiedene Diagramme verwendet. Das Liniendiagramm im linken Bereich des Panels zeigt den Konvergenzfortschritt des Prozesses. Die X-Achse beschreibt die Anzahl der Iterationen, die Y-Achse den aktuellen Fitness-Wert in Abhängigkeit von der zuvor gewählten Fitness-Funktion. Die drei Konvergenzlinien zeigen die Fitness des besten (rot), des schwächsten (dunkelgrün) und eines generischen Durchschnittsportfolios (blau) innerhalb der Population zum jeweiligen Zeitpunkt an. Umso weiter die Linien voneinander entfernt sind, umso inhomogener die Gesamtpopulation. Insbesondere direkt nach dem Start des Algorithmus zeigt sich stets eine große Divergenz welche bei fortlaufender Annäherung an das Portfolio-Optimum mit jeder Iteration tendenziell abnimmt.

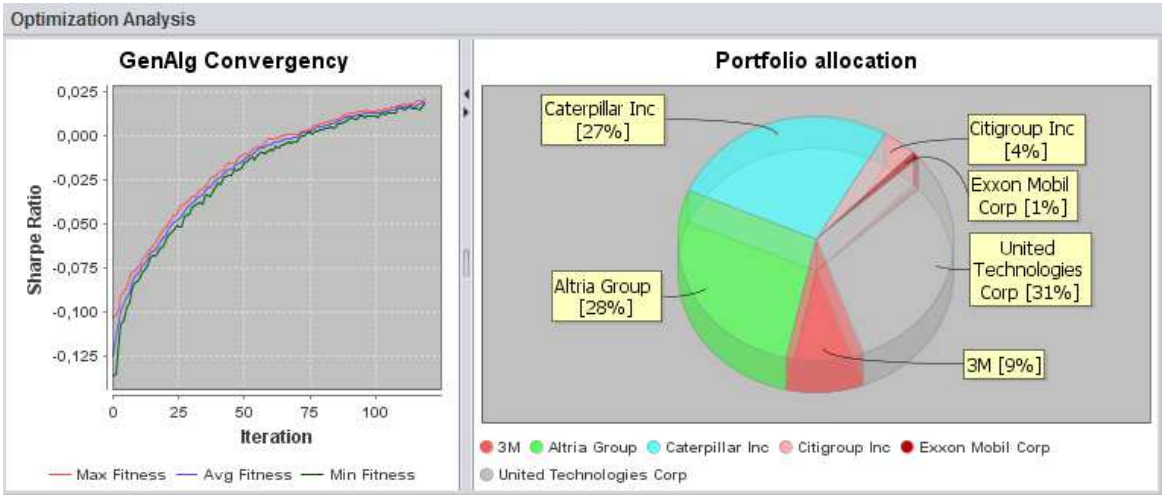

Abbildung 49 - Optimierungsanalyse

Die Portfolio-Allokation zeigt die Zusammensetzung des aktuell besten Portfolios (also jenes auf das sich die rote Linie im Konvergenzdiagramm bezieht).

Beide Diagramme werden nach jeder durchgeführten Iteration aktualisiert, der Optimierungsvorgang kann also durchaus beobachtet werden falls der gewählte Zeitraum und die Anzahl der Assets nicht vielfach höher sind als mit den gewählten Testdaten. Die

Optimierung mit 20 Werten über einen Zeitraum von vier Jahren auf wöchentlicher Berechnungsbasis nimmt nur wenige Minuten in Anspruch.

6.2.5. Visualisierung der Backtesting-Ergebnisse

Unter dem Tab „Backtesting“ im unteren Bereich der Anwendung können direkt nach Abschluss oder Abbruch des Optimierungsvorgangs jene Ergebnisse betrachtet werden, die sich aus der Zusammensetzung des optimierten Portfolios ergeben. Die sogenannte „Benchmark“ wird stets auf Basis des in den importierten Daten angegebenen Index berechnet. Hier steht sowohl der in den Daten angegebene Direktwert zur Verfügung, wie auch ein generischer Index der sämtliche vorhandenen Assets in einem generischen Portfolio annähernd gleichgewichtet.

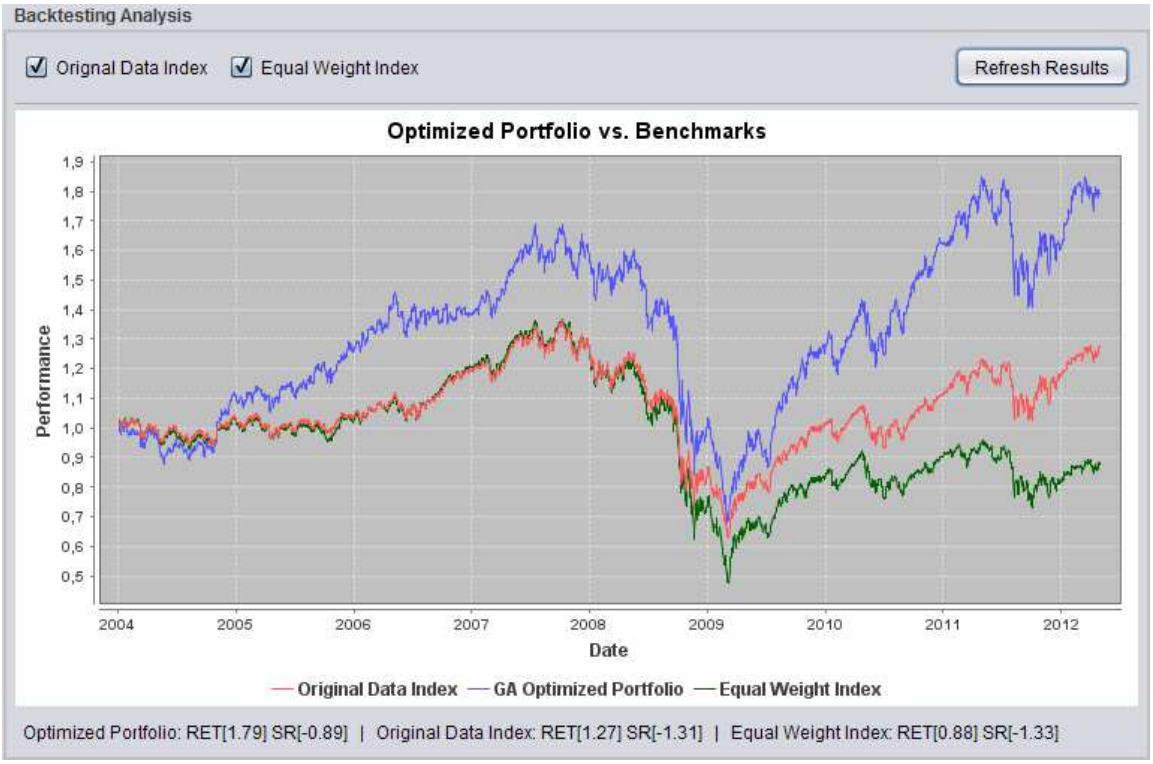

Abbildung 50 - Visualisierung des Backtesting

Erfolgt eine Anpassung des Backtesting-Zeitraums (siehe Punkt X) kann mit einem Klick auf „Refresh Results“ die Portfolioentwicklung über beliebige Zeiträume ad-hoc betrachtet werden. Unterhalb des angezeigten Charts werden die jeweiligen Eckdaten, also das zuvor gewählte Optimierungsmaß sowie der Gesamt-Return angezeigt.

6.2.6. Sonstige Einstellungsmöglichkeiten

Innerhalb sämtlicher Liniendiagramme kann beliebig „gezoomt“ werden. Durch einen Klick und anschließendem Zug innerhalb des Diagramms kann die zu vergrößernde Oberfläche ausgewählt werden. Nach dem Loslassen der Maustaste wird der gewählte Bereich neu gezeichnet und angezeigt.

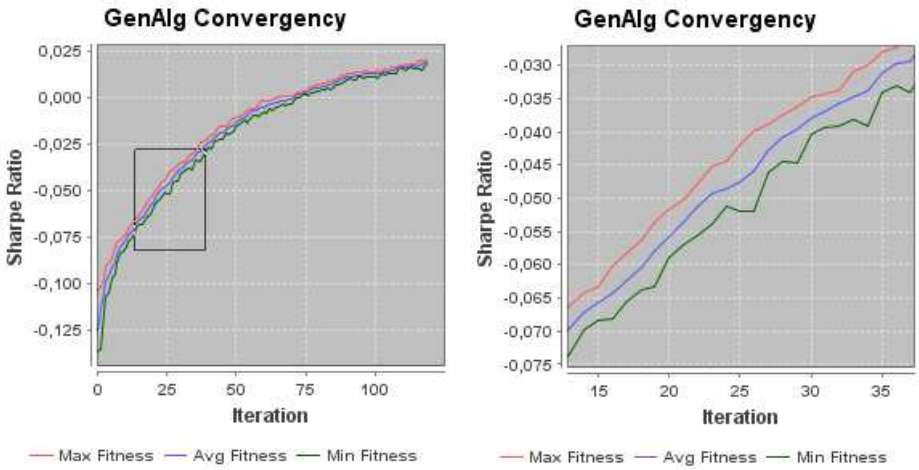

Abbildung 51 - Zooming in der Anzeige des Optimierungsgraphen

Diagramm, dann „Speichern unter“ erfolgen.

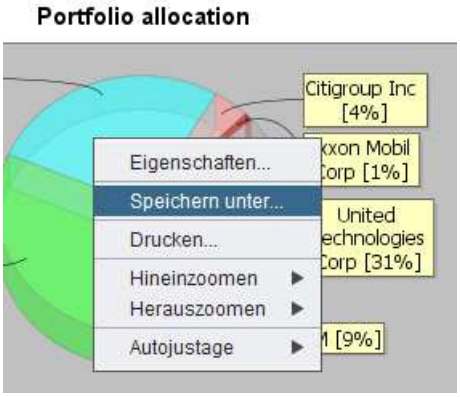

Abbildung 52 - Speichern den der generierten Grafiken

7. Conclusio und Ausblick

Genetische Algorithmen stellen als effizientes heuristisches Optimierungsverfahren eine gute Basis für die Finanzmarktanalyse dar. Durch die Dynamik des Verfahrens und seine verhältnismäßig simple Erweiterbarkeit über die Bewertungsfunktion, lassen sich eine Vielzahl von Analysen durchführen ohne auf die analytische Errechenbarkeit der Funktion achtgeben zu müssen. Insbesondere in der Praxis der aktiven Vermögensverwaltung entspricht dies einer verhältnismäßig hohen Arbeitszeitersparnis.

Eine wesentliche Herausforderung bei der Arbeit mit genetischen Algorithmen ist der grundsätzliche Aufbau des Settings. Dies umfasst vor allem jene Überlegungen die angestellt werden müssen um ein vorliegendes Problem, in diesem Fall das der Portfoliooptimierung in eine für GA lösbare Form zu übertragen. Die Definition und Kodierung der Individuen, sowie die Mechanismen zur Vererbung- und Verkreuzung müssen in einem relativ aufwändigen Entwicklungsprozess durchdacht und ausprobiert werden – die reine Bewertung der kodierten Individuen ist dann frei wählbar. Im Zuge der vorliegenden Diplomarbeit wurde eine im Vergleich zu bisherigen Ansätzen extrem simple aber nicht weniger funktionelle Methode erdacht.

Gut erweiterbare Lösungen auf Basis heuristischer Optimierungsverfahren stellen jedenfalls ein im Zusammenhang mit Portfoliooptimierung und Finanzmarktanalyse mächtiges Werkzeug dar. Bereits das vorliegende Setting bietet unbegrenzte Möglichkeiten zur Erweiterung des Algorithmus. Ein weiterer Schritt wäre die Implementierung eines gleitenden Zeitfensters bei welchem in Abhängigkeit der neu verfügbaren Information (= Kursdaten) Umschichtungen im Portfolio ausgelöst werden. Fließen dabei entsprechende Handelskosten mit in die Berechnung ein und werden dadurch Umschichtungsbarrieren erzeugt, kann das zu einer realitätsbezogenen Handelsunterstützung führen. Denkbar ist auch die Erweiterung um das Konzept der „Cash“-Position. Wertpapiere könnten antizyklisch in einer Umgebung mit sukzessiv sinkender Volatilität mit Strafpunkten belegt werden um dadurch eine Umschichtung in Bargeld auszulösen. Ob das langfristig auch tatsächlich im Vergleich zum Gesamtmarkt zu einer signifikant höheren Rendite führt wäre mittels Backtests zu evaluieren.

Literaturverzeichnis

[1] **Steiner, Manfred und Bruns, Christoph.** *Wertpapiermanagement.* Stuttgart : Schäffer-Poeschl Verlag, 2007.

[2] **Bodie, Zvi, Kane, Alex und Marcus, Alan J.** *Investments.* New York : McGraw Hill, 2009.

[3] **Weiss, David M.** *Financial Instruments: Equities, Debt, Derivatives, and Alternative Investments.* s.l. : Portfolio, 2009.

[4] **Bösch, Martin.** *Derivate: Verstehen, anwenden und bewerten.* München : Franz Vahlen Verlag, 2011.

[5] **Levinson, Marc.** *Guide to Financial Markets.* [Hrsg.] The Economist. s.l. : John Wiley & Sons, 2010.

[6] **SIFMA.** Securities Industry and Financial Markets Association. *US Bond Market Issuance and Outstanding.* [Online] [Cited: 07 01, 2013.] http://www.sifma.org/uploadedFiles/Research/Statistics/StatisticsFiles/CM-US-Bond-Market-SIFMA.xls?n=36045.

[7] **Monetary and Economic Department.** *Triennial Central Bank Survey of foreign exchange and derivatives market activity in April 2013.* s.l. : Bank for International Settlements, 2013.

[8] *Efficient Capital Markets: A Review of Theory and Empirical Work.* **Fama, Eugene.** s.l. : Journal of Finance, 1970. Bd. 25 [2], S. 383-417.

[9] *Efficient Capital Markets II.* **Fama, Eugene.** s.l. : Journal of Finance, 1991. Bd. 46 [5], S. 1575-1617.

[10] *The Analysis of Economic Time-Series-Part I: Prices.* **Kendall, Maurice und Hill, Bradford.** 1953. Journal of the Royal Statistical Society. Bd. 116 [1].

[11] *Market efficiency in real time.* **Jeffrey A. Busse, T. Clifton Green.** 2002. Journal of Financial Economics. S. 415-437.

[12] *Foundations of portfolio theory.* **Markowitz, Harry M.** 1952. Journal of Finance. Bd. 7 [1], S. 77-91.

[13] *The Volatility Effect: Lower Risk Without Lower Return.* **David Blitz, Pim Van Vliet.** 2007. Journal of Portfolio Management. S. 102-113.

[14] *What Hath MPT Wrought: Which Risks Reap Rewards?* **Arnott, Robert D.** 1983. Journal of Portfolio Management. S. 5-11.

[15] *The Cross-Section of Expected Stock Returns.* **Fama, Eugene F. und French, Kenneth R.** 1992. Journal of Finance. Bd. 47 [2], S. 427–465.

[16] *Evidence to support the four-factor pricing model from the Canadian stock market.* **L'Her, Jean-François, Masmoudi, Tarek und Suret, Jean-Marc.** 2004. Journal of International Financial Markets, Institutions and Money. Bd. 14 [4], S. 313-328.

[17] *Profitability of Momentum Strategies: An Evaluation of Alternative Explanations.* **Jegadeesh, Narasimhan und Titman, Sheridan.** 2001. Journal of Finance. Bd. 56 [2], S. 699-720.

[18] *Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency.* **Jegadeesh, Narasimhan und Titman, Sheridan.** 1993. Journal of Finance. Bd. 48 [1], S. 65-91.

[19] *Dangers of data mining: The case of calendar effects in stock returns.* **Sullivan, Ryan, Timmermann, Allan und White, Halbert.** 2001. Journal of Econometrics. Bd. 105 [1], S. 249-286.

[20] **Berkshire Hathaway Inc.** *2012 Annual Report.*

[21] **Holland, John H.** *Adaptation in Natural and Artificial Systems.* s.l. : MIT Press, 1992.

[22] **Weicker, Karsten.** *Evolutionäre Algorithmen.* Wiesbaden : Teubner Verlag, 2007.

[23] **Pohlheim, Hartmut.** *Evolutionäre Algorithmen: Verfahren, Operatoren und Hinweise für die Praxis.* Kleinmachnow : Springer Verlag, 1999.

[24] *Different Approaches to Solve the 0/1 Knapsack Problem.* **Hristakeva, Maya und Shrestha, Dipti.** 2005. Midwest Instruction and Computing Symposium.

[25] **Gerdes, Ingrid, Klawonn, Frank und Kruse, Rudolf.** *Evolutionäre Algorithmen.* s.l. : Vieweg Verlag, 2004.

[26] *Error-detecting and error-correcting codes.* **Hamming, Richard W.** 1950. Bell Systems Technical Journal. Bd. XXVI (2), S. 147-160.

[27] *Uniform Crossover in Genetic Algorithms.* **Syswerda, Gilbert.** s.l. : Morgan Kaufmann Publishers, 1989. International Conference on Genetic Algorithms. S. 2-9.

[28] *Representation and Hidden Bias II: Eliminating Defining Length Bias in Genetic Search via Shuffle Crossover.* **Caruana, Richard A., Eshelman, Larry J. und Schaffer, David.** San Mateo : Morgan Kaufmann Publishers, 1989. Eleventh International Joint Conference on Artificial Intelligence. Bd. 1, S. 750-755.

[29] **Mitsuo Gen, Runwei Cheng.** *Genetic Algorithms & Engineering Design.* s.l. : John Wiley & Sons, Inc, 1997.

[30] *A Combination Genetic Algorithm with Applications on Portfolio Optimization.* **Jiah-Shing Chen, Jia-Leh Hou.** s.l. : IEA/AIE Proceedings of the 19th international conference on Advances in Applied Artificial Intelligence, 2006. S. 197-206.

[31] Wikipedia, the free encyclopedia. *Dow Jones Index.* [Online] [Zitat vom: 10. 10 2012.] http://en.wikipedia.org/wiki/Dow_Jones_Industrial_Average.

[32] *Mutual Fund Performance.* **Sharpe, William F.** s.l. : The Journal of Business, 1966. Bd. 39, S. 119-138.

[33] **Young, Terry W.** Calmar Ratio: A smoother Tool. *Managed Money.* 1991, Bd. 20 [1], S. 40.

[34] **Rawlins, Gregorry J.E.** *Foundations of Genetic Algorithms.* San Mateo : Morgan Kaufmann Publishers, 1991.

[35] **Markowitz, Harry M.** *Portfolio Selection: Efficient Diversification of Investments.* New York : John Wiley and Sons, 1959.

[36] *How Inefficient is the 1/N Asset-Allocation Strategy?* **DeMiguel, Victor, Garlappi, Lorenzo und Uppal, Raman.** 2005.

[37] **Mitchell, Melanie.** *An Introduction to Genetic Algorithms.* Cambridge : MIT Press, 1999.