



SEMINARARBEIT

# TRADING SYSTEM TOOLS

vorgelegt von

**Johannes Messner**

betreut von

**Stefan Gerhold**

Wintersemester 2019/20

# Inhaltsverzeichnis

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Einleitung</b>  | <b>3</b>  |
| <b>2</b> | <b>Regression und Prognose für Handelsstrategien</b>               | <b>4</b>  |
| 2.1      | Analyse mit Hilfe von klassischen statistischen Methoden . . . . . | 4         |
| 2.2      | Parametrische- und Non-Parametrische Modelle . . . . .             | 4         |
| 2.3      | Glättungsmethoden . . . . .  | 5         |
| 2.3.1    | Glättung durch Spline . . . . .                                    | 5         |
| 2.3.2    | Kernel-Regression . . . . .  | 6         |
| 2.3.3    | LOESS Regression . . . . .   | 7         |
| 2.3.4    | k-Nächste Nachbarn Regression . . . . .                            | 7         |
| 2.4      | Prognose von finanziellen Zeitreihen . . . . .                     | 8         |
| 2.4.1    | Simple und Exponential Moving Average Prediction . . . . .         | 8         |
| 2.4.2    | Nächste Nachbar Prognose . . . . .                                 | 9         |
| <b>3</b> | <b>Testen und Verbessern von Handelsstrategien</b>                 | <b>10</b> |
| 3.1      | Testen von Handelssystemen . . . . .                               | 10        |
| 3.2      | Verbessern von Handelssystemen . . . . .                           | 10        |
| <b>4</b> | <b>Portfoliotheorie</b>  | <b>11</b> |
| 4.1      | Notation . . . . .   | 11        |
| 4.2      | Verteilungsfunktionen . . . . .                                    | 12        |
| 4.3      | Erweiterung zu Mehrperiodenmodell . . . . .                        | 13        |
| 4.4      | Markowitz's Ansatz . . . . .                                       | 13        |
| 4.5      | Effiziente Portfolios . . . . .                                    | 14        |
| 4.6      | Risiokolose Anlagen . . . . .                                      | 15        |
| <b>5</b> | <b>Das Capital Asset Pricing Model</b>                             | <b>16</b> |
| 5.1      | Annahmen . . . . .   | 16        |
| 5.2      | Effizienter Markt und die Kapitalmarktlinie . . . . .              | 16        |
| 5.3      | Das Preismodell . . . . .  | 17        |
| 5.4      | Schwachpunkte des CAPM . . . . .                                   | 18        |
| 5.5      | Vergleich zur Arbitrage Pricing Theory . . . . .                   | 18        |

# 1 Einleitung

In einer immer digitalisierteren Welt ist Trading als Form der Kapitalinvestition nicht mehr wegzudenken. Vor allem das Erstellen oder Kaufen eines sogenannten Trading Systems (dt: Handelsstrategie) wird immer beliebter, die Gründe sind natürlich das einfache Verbessern und Testen der Handelstrategie, der Wegfall des emotionalen Risikos und die Zeitersparnis.

Diese Arbeit widmet sich einiger Methoden zur Vorhersage der Entwicklung einzelner Handelssysteme. Außerdem werde ich auch auf die verschiedenen mathematischen Grundzüge der Kapitalmarkttheorie, die in im Laufe der Geschichte in der Finanzwelt vorherrschten, eingehen.

Grundlage dieser Seminararbeit ist das Kapitel *Trading System Tools* aus dem Buch *The Strategic Analysis of Financial Markets 2 - Trading System Analytics* von STEVEN D. MOFFITT [1].

## 2 Regression und Prognose für Handelsstrategien

Viele allgemeine Analyseverfahren von statistischen Daten gab es bis zur zweiten Hälfte des 20. Jahrhunderts noch nicht (z.B. Mustererkennung, data mining, artificial intelligence etc.). Heutzutage wird zwischen zwei Techniken zum Analysieren von Daten unterschieden: Überwachtes Lernen und Unüberwachtes Lernen. Überwachtes Lernen beinhaltet das Erstellen eines bestimmten Algorithmusses, der später einen genauen Zielwert vorhersagen soll. Beim Unüberwachten Lernen wird Muster zu erkennen und künftige Werte zu berechnen ohne vorher bestimmten Zielwert. In diesem Kapitel werden eine Methoden zum überwachtem Lernen vorgestellt.

### 2.1 Analyse mit Hilfe von klassischen statistischen Methoden

Statistik ist ein sehr weit entwickeltes Fachgebiet und eignet sich oftmals hervorragend als Grundlage zur Analyse von Datensätzen und Zeitreihen. Allerdings werden sehr ungünstige Modelle häufig verwendet, da der Zufall meist einen allzu großen Faktor darstellt. Betrachten wir als Beispiel folgendes Modell

$$x_t = x_{t-1} + \varepsilon_t,$$

wobei  $\varepsilon_t$  ein Weißes Rauschen ist. Offensichtlich ist dieses Modell zu ungenau um präzise Vorhersagen zu treffen und sollte nie als solches verwendet werden.

### 2.2 Parametrische- und Non-Parametrische Modelle

Ein Parametrisches Modell berechnet Ergebnisse unter Verwendung einer Verteilung von beliebigen Familien von Verteilungen  $\{T_\theta | \theta \in \Theta\}$ , wobei  $\Theta$  ein endlichdimensionaler Parameterraum ist. Das Modell sieht folgendermaßen aus:

$$Y = \mu + \beta'X + e,$$

wobei  $e \sim N(0, \delta^2)$  eine stochastische Variable ist,  $X$  und  $Y$  sind unabhängige Zufallsvariablen und  $\mu, \beta, \delta(>0)$  sind Parameter.

Non-Parametrische Modelle bedienen sich keiner speziellen Familie von Verteilungen (z.B. Histogramme). Semi-Parametrische Modelle enthalten, wie der Name vermuten lässt, Teile von beiden obigen Modellen:

$$Y = \mu + \beta'X + g(Z) + e,$$

wobei  $X, Y$  und  $Z$  eine gemeinsame Verteilung besitzen,  $g$  eine frei wählbare Funktion ist und für  $e$  muss gelten:  $\mathbf{E}[e|X, Y] = 0$ . Außerdem sind  $\mu$  und  $\beta$  endlichdimensional und  $g$  und  $e$  unendlichdimensional.

Größter Kritikpunkt am Parametrischen Modell ist, dass immer entweder der Parameter  $\theta$  oder dessen Verteilung geschätzt werden muss, weshalb sich ein subjektiver Einfluss nicht verhindern lässt. Dies wird bei Semi- bzw. Non-Parametrischen Modellen verhindert. Semi- und Non-parametrische Modelle lassen vor allem auch Glättungsmethoden (sog. *smoothing*) zu.

## 2.3 Glättungsmethoden

Mit Hilfe von Glättungsmethoden lässt sich Verlauf von Daten besser einschätzen, indem man versucht Beobachtungen, die abhängig von jeweiligen (Zeit-)Werten sind, einer gewissen Kurve zuzuordnen. Durch verschiedene Glättungsmethoden kommt man für dieselben Beobachtungen zu (meist geringfügig) unterschiedliche Kurven und damit auch unterschiedlichen Prognosen. Die bekanntesten Glättungsmethoden sind:

- Glättung durch Spline
- Kernel-Regression
- LOESS Regression
- k-Nächste Nachbar Regression

*Anmerkung:* Als Quelle der Daten der folgenden Grafiken wurde die 30-jährige Eurodollar-Entwicklung genommen (quartalsweise).

### 2.3.1 Glättung durch Spline

Ein häufiges Problem bei Splines (meistens werden kubische Splines verwendet) ist, dass bei Verwendung aller Daten als Stützstellen eine nicht allzu sehr glatte Funktion entsteht:

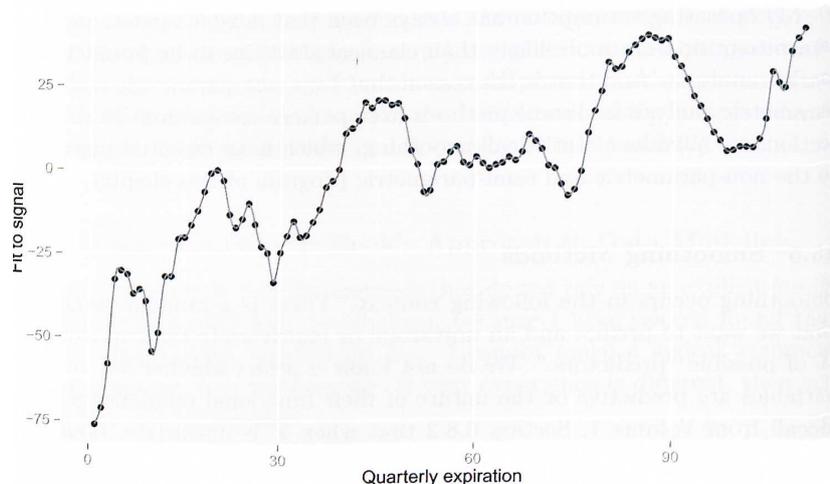


Abbildung 1: Spline mit allen Werten als Stützstellen

Daher werden kubische Spline oft mit der Methode der kleinsten Quadrate verwendet, so dass eine geringere Anzahl an geeigneten Stützpunkten gewählt wird. Dadurch erhält man außerdem eine glattere Funktion als zuvor:

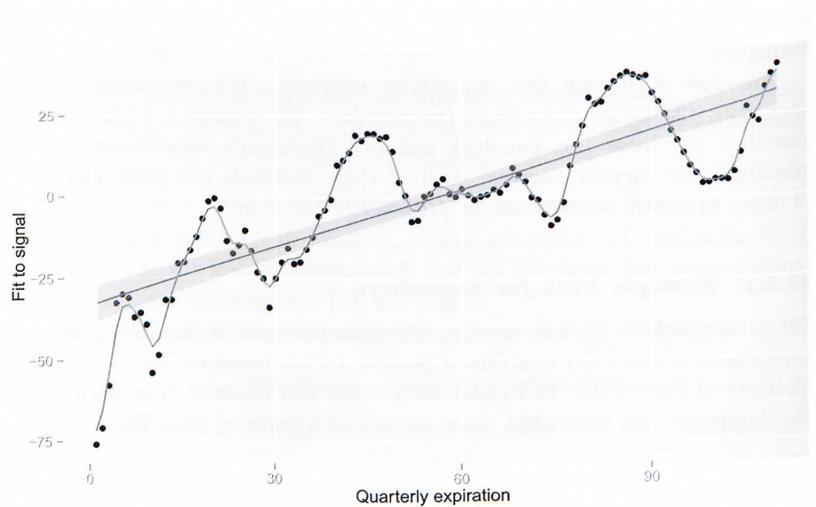


Abbildung 2: kubischer Spline mit weniger Stützstellen

### 2.3.2 Kernel-Regression

Bei der Kernel-Regression berechnet man die Funktion  $\hat{f}(x)$  mit der Formel:

$$\hat{f}(x) = \frac{1}{n} \sum_{j=1}^n K_h(x - x_j) = \frac{1}{nh} \sum_{j=1}^n K\left(\frac{x - x_j}{h}\right)$$

Hier ist  $h > 0$  eine beliebig gewählte Bandweite,  $x_1, \dots, x_n$  eine Stichprobe und  $K$  ein Kern (für den folgenden Graphen wurde der Gauss-Kern verwendet:  $K(u) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}u^2)$ ). Die folgende Kurve ist die glatteste von allen Bisherigen und Kommenden, daher wirken sich aber Ausreißer nicht allzu sehr auf den Verlauf des Graphen aus:

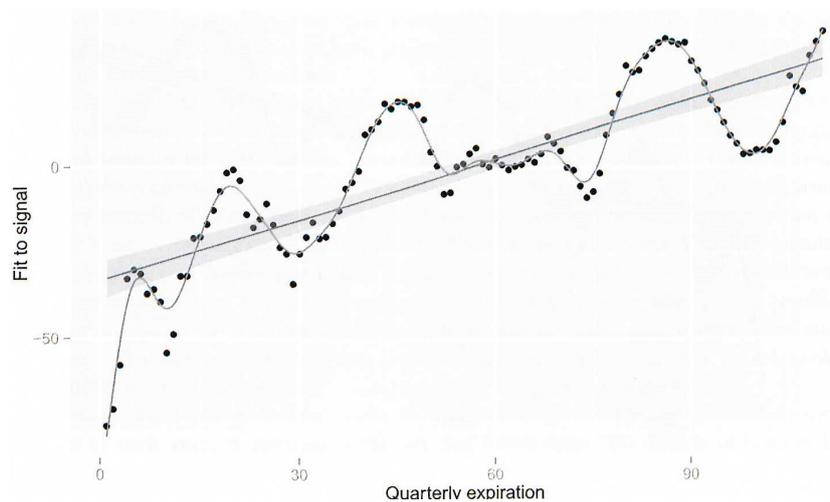


Abbildung 3: Kernel-Glättung

### 2.3.3 LOESS Regression

Bei dieser Art der Regression wird folgendes Modell verwendet:

$$y_i = f(x_i) + \epsilon_i$$

Hier ist  $f$  die unbekannte Funktion,  $(x_i, y_i)_{i=1, \dots, n}$  eine Stichprobe bzw. die bekannten Daten und  $\epsilon_i$  der Fehler. Zur Berechnung von  $\hat{f}(x)$  wählen wir zuerst eine Spannweite  $q$ . Nun nimmt man die ersten  $q$   $x_i$ 's mit dem kleinsten Abstand zu  $x$ . Dann ordnet man besagte  $x_i$ 's nach der Distanz zu  $x$ , beginnend mit dem Nächsten:  $(z_1, z_2, \dots, z_q)$ . Nun kann man mithilfe der *Trikubischen Gewichtsfunktion*:

$$T(u; t) = \mathbf{1}_{0 \leq u < t}(u; t) \left(1 - \left(\frac{u}{t}\right)^3\right)^3$$

die Gewichte  $w_i$  berechnen:

$$w_i = T(d(x, z_i), d(x, z_q))$$

Mit diesen Gewichten kann man nun eine gewichtete polynomiale Regression unter Verwendung des Modells  $y \sim p(x)$  durchführen um den den Wert des Schätzers  $\hat{f}(x)$  und dessen Varianz zu erhalten:

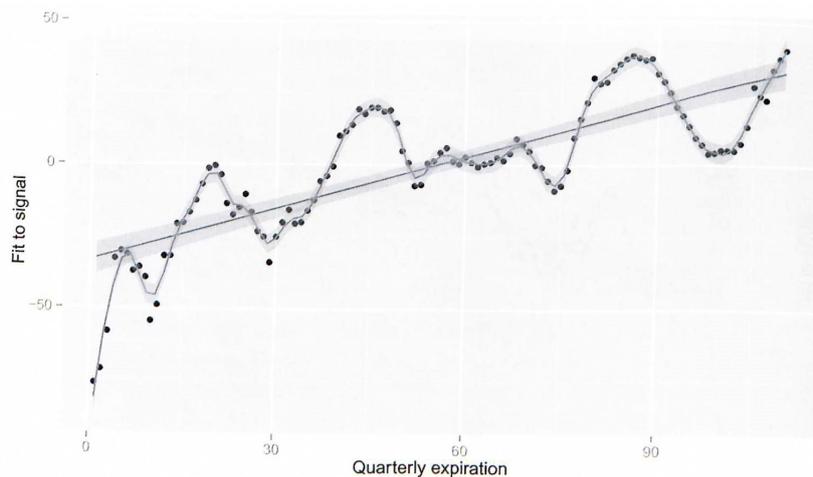


Abbildung 4: LOESS-Glättung

### 2.3.4 k-Nächste Nachbarn Regression

Diese Regressionsmethode ist der Kernel-Regression sehr ähnlich. Man berechnet hier den Schätzer  $\hat{f}(x)$  mit der Formel:

$$\hat{f}(x) = \sum_{i \in \mathcal{N}(x)} w_i y_i$$

Hier sind  $\mathcal{N}(x)$  die Indizes von den  $k$  nächsten  $x_i$ 's zu  $x$  und die Werte der  $w_i = K(x_i; x)$  erhält man mithilfe einer beliebigen *Kernel – Funktion* die an der Stelle  $x$  zentriert und in dessen Nähe symmetrisch ist. Eine häufige Wahl für  $K$  ist eine normale Dichtefunktion mit geeignetem  $\sigma^2$ . Da der Graph fast identisch zum vorherigen (LOESS-Regression) ist,

kann man obige Kurve betrachten.

Das größte Problem bei Glättungsmethoden ist vor allem die Ungenauigkeit an den Grenzen (aufgrund der verminderten Anzahl an Daten) weshalb sie für Prognosen oft ungeeignet sind. Allerdings werden Glättungsmethoden oft zum Erkennen von vorliegenden Trends benutzt, vor allem die LOESS-Regression wird hierfür sehr häufig als Hilfsmittel verwendet.

## 2.4 Prognose von finanziellen Zeitreihen

Wie man zuvor erkannt hat, eignen sich Glättungsmethoden selten um Prognosen von finanziellen Zeitreihen aufzustellen. Allerdings müssen bzw. sollten Prognosen nicht glatt sein, da finanzielle Zeitreihen dies auch nicht sind. Daher gibt es spezielle Methoden zur Prognose von Zeitreihen.

### 2.4.1 Simple und Exponential Moving Average Prediction

Die *Simple* und *Exponential Moving Average Prediction* haben weniger Parameter und unterliegen keinem speziellen Modell. Modernere Methoden benötigen gewisse Bedingungen (z.B. unabhängige identische Verteilung der Zufallsvariablen), sind aber trotzdem nicht immer für eine finanzielle Zeitreihe geeignet sind, weshalb deren Vorteile eher fragwürdig erscheinen.

$X_1, X_2, \dots, X_T$  ist eine Zeitreihe. Der Moving Average Prozess  $MA_t$  wird mithilfe der Formel

$$MA_t = \sum_{s=0}^{s=\infty} \alpha_s X_{t-s}$$

berechnet, wobei die  $\alpha_s$ 's fixierte Gewichte mit  $\sum \alpha_s = 1$  sind.

Der Simple Moving Average Prozess der Länge  $k$  ist die Zeitreihe  $SMA_t$ , die sich mit den Mittelwerten der vorherigen  $k$  Daten berechnen lässt:

$$SMA_t = \frac{1}{k} \sum_{s=t-k+1}^{s=t} X_s$$

wobei  $t = k, k+1, \dots, T$  gilt. Da für die ersten  $k-1$  Daten kein Simple Moving Average existiert, definiert man für diese den Prozess als

$$SMA_t = \frac{1}{t} \sum_{s=1}^{s=t} X_s$$

wobei  $t = 1, 2, \dots, k-1$ . Größter Kritikpunkt am Simple Moving Average Prozess ist, dass eine größere Schwankungen die vor mehr als  $k$  Perioden auftraten keinen Einfluss auf den Prozess hat.

Ein Exponential Moving Average Prozess  $EMA_t$  mit Parameter  $\beta$  wird rekursiv definiert als

$$EMA_t = \mathbf{1}_{t=1} X_1 + \mathbf{1}_{t>1} \beta X_t + (1 - \beta) EMA_{t-1}$$

Ein Vorteil vom *Exponential Moving Average Prozess* gegenüber dem *Simple Moving Average Prozess* ist natürlich der größere Einfluss von Schwankungen auf die Prognose.

### 2.4.2 Nächste Nachbar Prognose

Folgende Methode ist der bereits erwähnten *k – Nächsten Nachbar Regression* ähnlich. Sei also der Schätzer  $\hat{y}(x)$ :

$$\hat{y}(x) = \frac{\sum_{i=1}^k K(x, x_i) y_i}{\sum_{i=1}^k K(x, x_i)}$$

wobei  $K(\cdot, \cdot)$  eine geeignete Kernel-Funktion ist und die  $x_i$ 's die  $k$ -nächsten Werte von  $x$  sind. Ersetzen wir nun die  $x_i$ 's mit dem Vektor  $z_t \in \mathbb{R}^n$ :

$$z_t = (x_t, x_{t-1}, \dots, x_{t-n+1})^T$$

dieser besteht also aus  $x_t$  und den vorherigen  $n - 1$  Werten von  $x$ . Nun setzt man

$$\hat{x}_{t-1} = \frac{\sum_{i=1}^k K(z_t, z_{s_i}) x_{s_i+1}}{\sum_{i=1}^k K(z_t, z_{s_i})}$$

wobei die  $z_{s_i}$ 's die  $k$   $n$ -Werte in  $\{z_s | s < t\}$  mit größtem  $K(z_t, z_s)$ . Man betrachte, dass das Gewicht für den  $i$ -ten Wert  $x_{s_i+1}$  ist

$$\frac{K(z_t, z_{s_i})}{\sum_{i=1}^k K(z_t, z_{s_i})}$$

Die Summe dieser Gewichte ist klarerweise 1.

## 3 Testen und Verbessern von Handelsstrategien

In diesem Kapitel werde ich kurz auf die wichtigen Kriterien vom Testen von Handelsstrategien und Methoden zum Verbessern dieser eingehen.

### 3.1 Testen von Handelssystemen

Jedes Handelssystem sollte vor Inbetriebnahme immer gründlich getestet werden. Die vier wichtigsten Merkmale eines Handelssystems sind:

1. *Homogenität*: Ein Handelssystem sollte immer einem gewissen Trend folgen. Auch wenn ein Handelssystem hauptsächlich Gewinne einbringt, schränkt ein nicht homogenes Handelssystem den Trader in seinen Verbesserungsmöglichkeiten ein
2. *Vorhersagbarkeit*: Ein vorhersagbares Handelssystem gibt dem Trader die Möglichkeit nicht allzu oft in das System eingreifen zu müssen.
3. *Stabilität*: Ein stabiles Tradingsystem wird benötigt, da es auch während des Bärenmarktes Profite erzielt und damit nicht durch die aktuelle Wirtschaftslage beeinflusst wird

### 3.2 Verbessern von Handelssystemen

Zwei sehr beliebte und effektive Methoden sind die *Stop-Loss-Order* und der *Hebeleffekt*:

Mithilfe einer *Stop-Loss-Order* kann der Trader oder auch Anleger eine automatische Aktion beim Erreichen eines gewissen Kurses einer Aktie festlegen: Mit einer Stop-Verkauforder wird die Aktie vom Besitzer beim Erreichen eines bestimmten Kursniveaus verkauft. Mit einer Stop-Kauforder wird eine bestimmte Anzahl an Aktien beim Erreichen eines bestimmten Niveaus erworben.

Mithilfe des *Hebeleffekts* wird eine Kursänderung maximiert, d.h. hat ein Optionschein einen Hebel von 2 wirken sich Kursschwankungen doppelt so stark auf den Preis des Optionsscheins als auf den Preis der eigentlichen Aktie aus. Auch wenn sich dies nach einer Methode anhört mit der man sein Kapital schnell vervielfachen kann, ist hier höchste Vorsicht geboten, oft kommen durch den Hebeleffekt große Verluste, wenn nicht sogar ein Totalverlust zustande.

## 4 Portfoliotheorie

Bisher erfuhren wir nur über Analysemethoden für einzelne Systeme. Da man es meistens mit mehreren Systemen, die Teil eines Portfolios sind, zu tun hat, handelt dieses und nächstes Kapitel über die verschiedenen Theorien und Modelle zum Betrachten eines Portfolios als Ganzes.

Die *Portfoliotheorie*, engl. *Mean – Variance Portfolio Theory* (kurz MVPT) stammt vom us-amerikanischen Ökonom Harry Max Markowitz aus dem Jahre 1952 und war Grundlage für das von William Forsyth Sharpe, John Lintner und Jan Mossin entwickelte *Capital Asset Pricing Theorem* (kurz CAPT) aus den 1960ern. Sie legte den Grundstein für die klassische Kapitalmarkttheorie und galt lange Zeit mit dem CAPM als Grundlage für Berechnungen in der Finanzwelt. Heutzutage wird sie allerdings nicht mehr allzu häufig verwendet.

### 4.1 Notation

Um die später folgenden Formeln verstehen zu können, betrachte man als erstes die Notation: Es gibt  $n$  mögliche Investments ( $n \in \mathbb{N}$ ), ein zur Verfügung stehendes Kapital  $W$ , Beteiligungen  $h_i$ , wobei  $i = 1, \dots, n$  und den Wert des Portfolios zum Zeitpunkt  $T$   $V_t$ . Klarerweise muss gelten  $V_1 \leq W$ . Die Preise der  $i$ -ten Anlage zum Zeitpunkt  $t = 1$  bzw.  $t = 2$  sind  $X_{i1}$  bzw.  $X_{i2}$  und der entstehende Cashflow der  $i$ -ten Anlage zwischen den Perioden 1 und 2 ist  $C_{i1}$ .

Fasst man nun die Variablen zu Vektoren zusammen erhält man:  $X_t = (X_{1t}, X_{2t}, \dots, X_{nt})^T$  mit  $t = 1, 2$  als Preisvektor,  $h = (h_1, h_2, \dots, h_n)^T$  als Portfolio und  $C = (C_{11}, C_{21}, \dots, C_{n1})$  als Cashflowvektor.

Mit dieser Notation beträgt der Wert der Portfolios  $h$  für  $t = 1, 2$ :

$$V_t = \sum_{i=1}^n h_i X_{it} = h^T X_t$$

Die Preisdifferenz zwischen den Perioden 1 und 2 beträgt

$$\Delta X_{i1} = X_{i2} - X_{i1},$$

also beträgt der *Portfoliozuwachs* für diese Periode

$$V_2 - V_1 + C = \sum_{i=1}^n h_i (\Delta X_{i1} + C_{i1})$$

Für  $C$  gilt daher:

$$C = \sum_{i=1}^n h_i C_{i1} = h^T C_1$$

Diese Cashflows können z.B. Dividenden, ausgezahlte Zinsen, bezahlte Zinsen etc. sein (sie können u.a. auch negativ sein).

Schreiben wir nun den *Portfoliozuwachs* etwas um, erhalten wir

$$\sum_{i=1}^n h_i X_{i1} R_{i1}$$

wobei  $R_{i1}$  die Rendite der  $i$ -ten Anlage in der Periode zwischen 1 und 2 ist, also

$$R_{i1} = \frac{\Delta X_{i1} + C_{i1}}{X_{i1}}$$

Definiert man nun  $R_1$  als den Vektor der Renditen mit

$$R_1 = (R_{11}, R_{21}, \dots, R_{n1})^T$$

so kann die Rendite vom Portfolio  $h$  mit

$$R_h = \frac{V_2 - V_1 + C}{V_1} = \frac{\sum_{i=1}^n h_i X_{i1} R_{i1}}{\sum_{j=1}^n h_j X_{j1}} = \sum_{i=1}^n w_i R_{i1} = w^T R_1$$

berechnet werden, wobei  $w = (w_1, w_2, \dots, w_n)^T$  mit

$$w_i = \frac{h_i x_{i1}}{\sum_{j=1}^n h_j X_{j1}} = \frac{h_i x_{i1}}{V_1}$$

wobei  $V_1 \neq 0$  angenommen wird. Man sieht auch sofort, dass die Summe der Gewichte  $w_i$  gleich 1 ist.

## 4.2 Verteilungsfunktionen

Seien  $X_t = (X_{1t}, X_{2t}, \dots, X_{nt})^T$  der Zufallsvektor der Preise für  $t = 1, 2$  und  $C = (C_{11}, C_{21}, \dots, C_{n1})$  der Zufallsvektor der Cashflows, wie bereits bekannt. Unter der Annahme, dass  $(X_1, X_2, C_1)$  die Dichtefunktion  $g(x_1, x_2, c_1)$  besitzt, versucht man die Verteilung der Renditen  $R_h$  für das Portfolio  $h$  zu finden:

$$R_h = \frac{h^T (X_2 - X_1 + C_1)}{h^T X_1}$$

bedingt auf  $X_1 = x_1$ . Zuerst kann man die Verteilungsfunktion von  $U = X_2 - X_1 + C_1$  bedingt auf  $X_1$  mit der Funktion  $g$  berechnen und dann erhält man mit  $w = \frac{h}{h^T X_1}$ :

$$R_h = w^T U | X_1 = \sum_{i=1}^n \frac{h_i}{\sum_{i=1}^n h_i X_{i1}} U_i X_1$$

Ein häufiges Problem war früher die Annahme, dass die gemeinsame Verteilungsfunktion  $g$  normalverteilt wäre, also

$$U | X_1 \sim N_n(\mu, \delta)$$

weshalb

$$R_h | X_1 \sim N(w^T \mu, w^T \delta w)$$

Allerdings ist diese Annahme falsch, da die Renditen im allgemeinen nicht normalverteilt sind.

### 4.3 Erweiterung zu Mehrperiodenmodell

Die Erweiterung auf mehrere Perioden lässt sich unter Betrachtung von 2 Einflüssen problemlos ausführen: die sich verändernde Verteilung der Renditen (zufälliger Einfluss) und die Veränderung der Gewichte im Portfolio (nicht zufällig, lässt sich vom Portfoliomanager verändern).

### 4.4 Markowitz's Ansatz

Sei  $R_1$  ein Zufallsvektor der Länge  $n$  mit den Renditen der  $n$  Anlagen, wobei die erwartete Rendite  $\mu$  und dessen Standardabweichung bzw. Risiko  $\delta$  bereits bekannt sind. Dann wählt man einen Wert  $r$  als erwünschte Rendite des Portfolios, welche man in der nächsten Periode erreichen will. Markowitz Idee war es das Risiko zu minimieren bei zu erwartender Rendite  $r$ .

Dies heißt dann also:

$$\text{Minimiere } w^T \delta w, \quad \text{wobei } w^T \mu = r$$

wobei die Summe der Einträge des Gewichtsvektors  $w$  (Aufteilung des Kapitals) hier wieder 1 sein muss.

Löst man nun dieses Optimierungsproblem erhält man den Vektor  $w$ :

Als Nebenbedingung wählt man  $Aw = b$  mit:

$$A = \begin{bmatrix} \mu^T \\ \mathbf{1}^T \end{bmatrix} \quad \text{und} \quad b = \begin{bmatrix} r \\ 1 \end{bmatrix}$$

Damit kommt man auf die Lagrange-Funktion (die Verwendung von  $-2\lambda^T$  erleichtert die späteren Umformungen):

$$L(w, \lambda) = w^T \delta w - 2\lambda^T (Aw - b)$$

Daraus erhält man die partiellen Ableitungen

$$L'_w(w, \lambda) = 2\delta w - 2A^T \lambda = 0 \quad \text{und} \quad L'_\lambda(w, \lambda) = Aw - b = 0$$

Damit erhält man nun

$$w = \delta^{-1} A^T \lambda, \quad b = Aw = A\delta^{-1} A^T \lambda$$

und

$$\lambda = (A\delta^{-1} A^T)^{-1} b$$

woraus man nun für  $w$  folgendes erhält:

$$w = \delta^{-1} A^T (A\delta^{-1} A^T)^{-1} b \tag{1}$$

Damit  $A\delta^{-1} A^T$  invertierbar ist, muss die Matrix  $A$  vollen Rang haben, wodurch der Vektor  $\mu$  unproportional zum Einervektor sein muss, also müssen sich mindestens 2 Einträge in  $\mu$  unterscheiden. Wenn  $\delta$  invertierbar sein muss, heißt dies dass jegliche Anlage risikobehaftet sein muss.

Das Problem bei diesem Ansatz ist natürlich die Bekanntheit der Werte von  $\mu$  und  $\delta$ , also die Erwartungswerte der Renditen und das Risiko bzw. die Streuung, diese kann man nur die bisherigen Daten abschätzen.

## 4.5 Effiziente Portfolios

Um das optimale Portfolio zu finden, setzt man (1) in  $w^T \delta w$  ein:

$$\begin{aligned} w^T \delta w &= b^T (A \delta^{-1} A^T)^{-1} A \delta^{-1} \delta \delta^{-1} A^T (A \delta^{-1} A^T)^{-1} b = \\ &= b^T (A \delta^{-1} A^T)^{-1} b \end{aligned}$$

Setzt man nun

$$(A \delta^{-1} A^T) = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

erhält man

$$w^T \delta w = g_{11} r^2 + 2g_{12} r + g_{22} > 0$$

wobei  $r \neq 0$ , also die festgelegte Rendite nicht 0 sein darf. Diese Ungleichung gilt, da  $w^T \delta w$  positiv definit ist und  $g_{12} = g_{21}$  gilt aufgrund der Symmetrie.

Erstellt man nun anhand dieser Erkenntnis eine Kurve der Form (Standardabweichung, Erwartungswert), also

$$(\sqrt{g_{11} r^2 + 2g_{12} r + g_{22}}, r)$$

erhält man folgende Graphik:

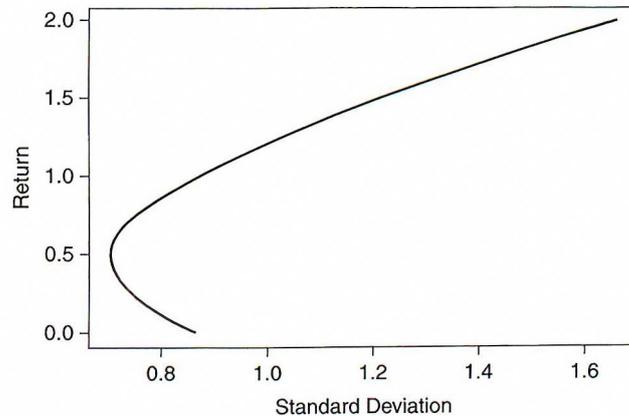


Abbildung 5: Effizienzlinie

Da der obere Teil der entstandenen Kurve den unteren dominiert, ist nur die Kurve ab einer Rendite von ca. 0.5 für den Anleger interessant. Im Allgemeinen gilt dies für eine Rendite  $r$  für die gilt  $r \geq \frac{g_{12}}{g_{11}}$ . Dieser Teil der Kurve nennt man auch *Effizienzlinie*.

*Anmerkung:* Angenommen man findet zwei Lösungen  $w_1, r_1$  und  $w_2, r_2$  für das Optimierungsproblem.

Folgendes Portfolio

$$w = \alpha w_1 + (1 - \alpha) w_2$$

mit Erwartungswert

$$r = \alpha r_1 + (1 - \alpha)r_2$$

ist ebenfalls eine Lösung, also auch ein optimales Portfolio. Dass Linearkombinationen von optimalen Porfolios wieder optimale Portfolios sind, wird auch als *Two-Fund Theorem* bezeichnet.

## 4.6 Risiokolose Anlagen

Angenommen man hat eine risikolose Anlage mit Rendite  $r_f$ , also

$$E[r_f] = r_f$$

$$Var[r_f] = 0$$

Betrachtet man nun noch ein risikobehaftetes Portfolio  $R$  mit

$$E[R] = r_r$$

$$Var[R] = \sigma_r^2 > 0$$

Kann man nun mithilfe eines  $\alpha \in (0, 1)$  das zu investierende Kapital auf die risikolose Anlage  $r_f$  und das risikobehaftete Portfolio  $R$  wie folgt zusammenfügen:

$$X = \alpha r_f + (1 - \alpha)R \tag{2}$$

Dadurch ergeben sich Erwartungswert und Standardabweichung:

$$\mu_X = E[X] = \alpha r_f + (1 - \alpha)r_r$$

$$\sigma_X = \sqrt{Var[X]} = (1 - \alpha)\sigma_r$$

Mit diesem Ansatz kan man eine Tangente der Form  $r_\alpha = \alpha r_f + (1 - \alpha)r_r$  auf die Effizienzlinie legen, welche auf der y-Achse den Punkt  $r_f$  trifft. Eine ähnlich Gegebenheit wird uns später noch einmal begegnen.

## 5 Das Capital Asset Pricing Model

### 5.1 Annahmen

Das Capital Asset Pricing Model, CAPM, hatte von allen Theorien den bisher größten Einfluss auf die Finanzwelt, auch wenn es heute nicht mehr allzu verbreitet ist. Es beruht auf der Portfoliotheorie von Markowitz. Um es anwenden zu können werden allerdings folgende Annahmen über die Investoren getroffen:

1. Eine risikolose Anlage existiert
2. Alle ersten und zweiten Momente der Rendite sind bekannt
3. Jedes Portfolio hat eine Standardabweichung die nicht *Null* ist
4. Die erwartete Rendite eines Portfolios mit minimaler Standardabweichung ist positiv
5. Alle Investoren sind risikoabgeneigt und haben alle dasselbe Verständnis über den Markt
6. Alle Investoren optimieren das Verhältnis zwischen Erwartungswert und Standardabweichung

Es lässt sich allerdings schnell erkennen, dass in der Praxis die meisten Annahmen nie vollkommen zutreffen:

1. Im Laufe der letzten 120 Jahre kam es immer wieder zu Fällen in denen Staatsanleihen nicht das einhielten was sie versprochen hatten, da dies allerdings selten der Fall, kann man diese Annahme trotzdem treffen
2. Diese Momente lassen sich zwar anhand historischer Daten berechnen, können aber meistens damit nur angenähert werden
3. Wahrscheinlich wahr, allerdings kann dies nicht bewiesen werden
4. Kann auch als wahr angenommen werden, es kann aber auch nicht nachgewiesen werden
5. Der erste Teil kann als wahr angenommen werden aber Letzterer trifft in der Praxis natürlich nicht zu
6. Trifft nicht zu, da sich in der Praxis immer wieder das Gegenteil beobachten lässt

### 5.2 Effizienter Markt und die Kapitalmarktlinie

Da Investoren das Eigenwert-Abweichungsverhältnis optimieren, wird jeder ein Portfolio mit einer erwarteten Rendite zu minimalen Risiko wählen. Nun sollen alle Preise der Anlagen auf ein faires Niveau gelangen: Zu teure Anlagen werden verkauft und deren Preis sinkt, zu billige Anlagen werden gekauft bis sie den angemessenen Preis erreichen. Dadurch wird die Standardabweichung minimiert.

Wie man bereits bei (2) gesehen hat, hat dieses Portfolio die Form

$$r_e = \alpha r_f + (1 - \alpha)r_m,$$

wobei hier wieder  $\alpha \leq 0$  gilt und die realisierte Rendite des risikobehafteten Portfolios  $r_m$  ist. Sei  $R_m$  nun die Zufallsvariable des risikobehafteten Portfolios, dann gilt

$$R_e = \alpha r_f + (1 - \alpha)R_m = \alpha r_f + (1 - \alpha)(r_m + \epsilon) =$$

$$\alpha r_f + (1 - \alpha)r_m + (1 - \alpha)\epsilon$$

wobei  $\epsilon = R_m - r_m$  die zufällige Abweichung der Rendite ist, also  $E[\epsilon] = 0$ . Bildet man wie zuvor eine Tangente an der Effizienzlinie die durch  $r_f$  läuft, erhält man folgende Geradengleichung:

$$r(\sigma) = \frac{r_m - r_f}{\sigma_m} \sigma + r_f$$

Diese Gerade wird auch *Kapitalmarktlinie* genannt und stellt das Risiko-Renditenverhältnis des effizienten Portfolios, bestehend aus risikolosen und risikofreien Anlagen, dar. Nach dieser sollte sich der Anleger richten und sich dann entscheiden wie viel Risiko er bereit ist einzugehen:

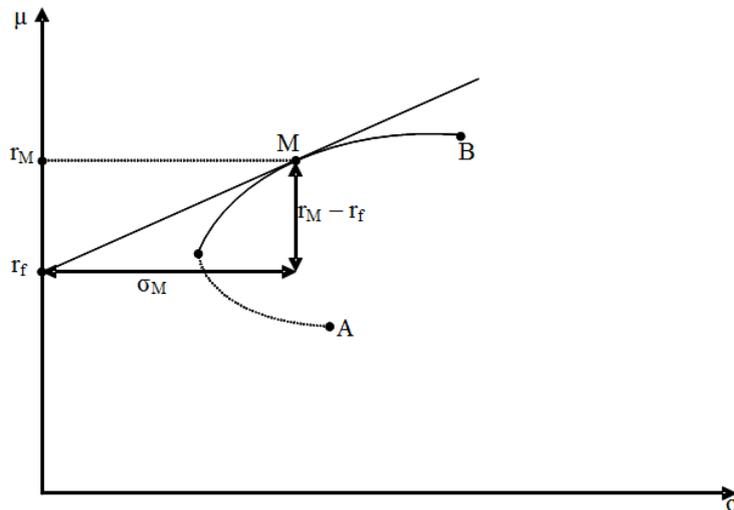


Abbildung 6: Kapitalmarktlinie

Sehr vielsagend ist außerdem die Steigung dieser Geraden

$$K = \frac{r_m - r_f}{\sigma_m}$$

die die Veränderung der Rendite bei größer werdendem Risiko modelliert.

### 5.3 Das Preismodell

Betrachtet man nun die erwartete Rendite  $r_i$  einer Anlage  $i$ , kommt man zu folgender, wahrscheinlich wichtigster Gleichung des CAPMs:

$$r_i - r_f = \beta_i(r_m - r_f), \tag{3}$$

wobei gilt, dass

$$\beta_i = \frac{Cov(R_i, R_m)}{\sigma_m^2}$$

Dies heißt, dass die erwartete Rendite nur von deren Verhältnis zum gesamten risikobehaftetem Portfolios abhängt.

*Anmerkung:* Das  $\beta_i$  ist die Veränderung der erwarteten Rendite von  $i$  im Verhältnis zur Veränderung der erwarteten Rendite des risikobehaftetem Portfolios.

Obige Formel (3) lässt sich auch auf ein Teilportfolio anwenden:

$$r_w - r_f = \beta_w(r_m - r_f),$$

wobei

$$r_w = w^T \vec{r} \text{ und } \beta_w = w^T \vec{\beta}$$

$\vec{r}$  ist der Vektor der erwarteten Renditen und  $\vec{\beta}$  ist der Vektor der Betas.

## 5.4 Schwachpunkte des CAPM

Obwohl das CAPM in den Jahren nach seiner Veröffentlichung (1960er) revolutionär war und noch heute in einigen Fällen Verwendung findet, vor allem bei Performancebmessungen von Investmentfonds, gilt die Anwendung im Allgemeinen Fall aufgrund der unrealistischen Annahmen an die Investoren und die Voraussetzung, dass Erwartungswerte und Standardabweichung a priori bekannt sein müssen, als problematisch.

## 5.5 Vergleich zur Arbitrage Pricing Theory

Die Arbitrage Pricing Theory, oder APT, wurde in den letzten Jahrzehnten zur gängigsten Methode für die Berechnung wichtiger Kennzahlen im Kapitalmarkt. Die APT wurde im Jahre 1976 vom Ökonomen Stephen Ross entwickelt, mit dem Ziel eine Theorie aufzustellen, welche keine derartig strikten Bedingungen an den Markt und an die Investoren stellte.

Größter Unterschied zum CAPM ist, dass hier kein Marktgleichgewicht gefordert wird, sondern in abgeschwächter Form von einem arbitragefreiem Markt ausgegangen wird. Dies gibt der APT einen realistischeren und glaubwürdigeren Ansatz. Ein anderer signifikanter Unterschied ist die Anzahl der Faktoren bzw. Einflüsse: Die APT ist ein Mehrfaktorenmodell, d.h. es gibt mehrere verschiedene Marktfaktoren, die alle andere Auswirkungen auf den Kurs eines Wertpapiers haben können. Hat man nun  $n$  Marktfaktoren:

$$F_1, F_2, \dots, F_n$$

Nun gilt für die Rendite  $r_i$  des Wertpapiers  $i$ :

$$r_i = R_i + \beta_{i1}F_1 + \beta_{i2}F_2 + \dots + \beta_{in}F_n + \epsilon_i,$$

wobei  $R_i$  die erwartete Rendite von  $r_i$  ist, die  $\beta_{ij}$ 's der Einfluss des  $j$ -ten Faktors ist und  $\epsilon_i$  der Fehler ( $E[\epsilon_i] = 0$ ). Damit kommt man auf die von Ross entwickelte ATP-Formel:

$$E[r_i] = r_f + \beta_{i1}RP_1 + \beta_{i2}RP_2 + \dots + \beta_{in}RP_n,$$

wobei  $r_f$  die risikolose Anlage ist und die  $RP_j$ 's die  $j$ -te Risikoprämie darstellen. Das CAPM hingegen geht nur von einem Faktor, dem Marktrisiko aus:

$$r_w = r_f + \beta_w(r_m - r_f)$$

Dadurch hat man mit der APT ein viel präziseres, aber auch ein zeitaufwändigeres Modell als das CAPM.

## Abbildungsverzeichnis

|   |   |    |
|---|---|----|
| 1 | Spline mit allen Werten als Stützstellen . . . . .  | 5  |
| 2 | kubischer Spline mit weniger Stützstellen . . . . . | 6  |
| 3 | Kernel-Glättung . . . . .                           | 6  |
| 4 | LOESS-Glättung . . . . .                            | 7  |
| 5 | Effizienzlinie . . . . .                            | 14 |
| 6 | Kapitalmarktlinie . . . . .                         | 17 |

## Literatur

- [1] STEVEN MOFFITT: *The Strategic Analysis of Financial Markets 2 - Trading System Analytics*, World Scientific Series in Finance - Vol. 11 (2017)
- [2] CAPM - Modell der Kapitalmarktlinie <https://www.wiwiweb.de>, eingesehen am 24.02.2020
- [3] SEBASTIAN STEYER: *Crashkurs Trading: Das Einsteigerwerk für Trader - und alle, die es werden möchten!*, Börsenbuchverlag - 3. Auflage (2014)