vorne ziehen; wir erhalten

$$\int f(x)\log_2 \frac{\langle b_f, x\rangle}{\langle b_g, x\rangle}\frac{g(x)}{f(x)}\,dx \le \log_2 \int f(x)\frac{\langle b_f, x\rangle}{\langle b_g, x\rangle}\,dx$$
$$= \log_2 \int g(x)\frac{\langle b_f, x\rangle}{\langle b_g, x\rangle}\,dx \le \log 1 = 0\,,$$

da b_g das log-optimale Portfolio zur Wahrscheinlichkeitsdichte g ist. Somit ist $\Delta W \le D(f\|g)$, wie behauptet. ■

Diesen Satz wollen wir nun anwenden auf den Fall, daß wir zusätzliche Informationen über die Entwicklung der Börse haben. Diese Information beschreiben wir wieder durch eine Zufallsvariable Y. Die gemeinsame Wahrscheinlichkeitsverteilung von X und Y sei durch die Wahrscheinlichkeitsdichte $f(x,y)$ gegeben, f_X und f_Y seien die Wahrscheinlichkeitsdichten zu X und Y allein. Ohne Zusatzinformation definieren wir das log-optimale Portfolio anhand von f_X, mit der Zusatzinformation $Y = y$ verwenden wir stattdessen f mit zweitem Argument y.

Satz: Der Anstieg ΔW der Verdoppelungsrate durch die in Y kodierte Zusatzinformation ist höchstens gleich $I(X;Y)$.

Beweis: Wenn wir das log-optimale Portfolio auf einen konkreten Wert $Y = y$ abstimmen, steigt die Verdoppelungsrate nach dem vorigen Satz höchstens um

$$D\big(f(x|X = y)\|f(x)\big) = \int_x f(x|Y = y)\log_2 \frac{f(x|Y = y)}{f_X(x)}\,dx\,.$$

ΔW ist das mit f_Y gewichtete Mittel über diese Anstiege, also kleiner oder gleich

$$\int_y f_Y(y)\int_x f(x|Y = y)\log_2 \frac{f(x|Y = y)}{f_X(x)}\,dx\,dy$$
$$= \int_y\int_x f_Y(y)f(x|Y = y)\log_2 \frac{f(x|Y = y)}{f_X(x)}\frac{f_Y(y)}{f_Y(y)}\,dx\,dy$$
$$= \int_y\int_x f(x,y)\log_2 \frac{f(x,y)}{f_X(x)f_Y(y)}\,dx\,dy = I(X;Y)\,.$$ ■

Man beachte, daß der Zuwachs hier im Gegensatz zum Fall der Pferdewetten nicht gleich der wechselseitigen Information sein muß, sondern auch kleiner ausfallen kann.

f) Verallgemeinerung auf stationäre Märkte

Bislang sind wir davon ausgegangen, daß die verschiedenen Börsentage (oder auch die hypothetischen Wiederholungen eines Börsentags) durch voneinander unabhängige Zufallsvariablen beschrieben werden. In letzter Konsequenz bedeutet diese Annahme, daß sich die Börse am Tag nach einem großen Crash verhält, als sei nichts geschehen, daß es keine Gewinnmitnahmen nach Höhenflügen einer Aktie gibt, und so weiter.

Für den Rest dieses Paragraphen wollen wir daher unser Modell erweitern und die Kursentwicklung beschreiben durch einen stochastischen Prozess $\mathcal{X} = X^{(1)}, X^{(2)}, \ldots$ aus Zufallsvariablen $X^{(i)}$ mit Werten in $\mathbb{R}^m$, wobei die k-te Komponente von $X^{(i)}$ angibt, wie sich die Aktie k am i-ten Börsentag entwickelt, d.h. mit welchem Faktor ihr Wert im Laufe des Tages multipliziert wird. Dieser stochastische Prozess soll stationär sein, insbesondere haben also weiterhin alle $X^{(i)}$ dieselbe Wahrscheinlichkeitsverteilung F; wir gehen aber nicht mehr davon aus, daß die Wertentwicklungen der einzelnen Tage unabhängig voneinander sind.

Für die Anlage betrachten wir weiterhin kausale Strategien, d.h. das Portfolio $b^{(i)}$ am i-ten Tag hängt ab von den Werten der $X^{(j)}$ mit $j < i$. Um herauszustellen, daß $b^{(i)}$ eine Funktion dieser Zufallsvariablen ist, schreiben wir gelegentlich auch $b^{(i)}(X^{(1)}, \ldots, X^{(i-1)})$; den Funktionswert für konkrete Werte $X^{(j)} = x^{(j)}$ bezeichnen wir mit $b(x^{(1)}, \ldots, x^{(i-1)})$.

Auch in der neuen Situation wollen wir den Erwartungswert für die langfristige Wertentwicklung der Anlage maximieren, also den Erwartungswert des Logarithmus von

$$S_n = \prod_{i=1}^{n} \big\langle b^{(i)}(X^{(1)}, \ldots, X^{(i-1)}), X^{(i)}\big\rangle\ .$$

Das Maximum dieses Erwartungswert über alle kausalen Anlagestrategien ist die Summe über die maximalen Erwartungswerte für die einzelnen Tage; wir wählen also für jeden Tag dasjenige Portfolio, von dem wir das maximale logarithmische Wachstum erwarten, d.h. das log-optimale Portfolio. Da die Zufallsvariablen $X^{(i)}$ nicht mehr als unabhängig vorausgesetzt sind, dürfen wir allerdings nicht mehr einfach jeden Tag das log-optimale Portfolio zur Wahrscheinlichkeitsverteilung F nehmen, denn wie sich $X^{(i)}$ entwickelt hängt ja ab von den Entwicklungen der Vortage. Das log-optimale Portfolio $b^{*(i)}(x^{(1)}, \dots, x^{(i-1)})$ muß daher bestimmt werden über die bedingte Wahrscheinlichkeitsverteilung unter der Voraussetzung, daß $X^{(1)} = x^{(1)}, \dots, X^{(i-1)} = x^{(i-1)}$ ist. Den entsprechenden Erwartungswert für den i-ten Börsentag bezeichnen wir als die bedingte Verdoppelungsrate dieses Tages und schreiben sie als

$$W^*(X^{(i)}|x^{(1)}, \dots, x^{(i-1)}) =$$
$$\mathbb{E}\Big(\log_2 \big\langle b^{*(i)}(x^{(1)}, \dots, x^{(i-1)}), X^{(i)}\big\rangle \;\Big|\; X^{(1)} = x^{(1)}, \dots, X^{(i-1)} = x^{(i-1)}\Big)\,.$$

Der Erwartungswert des Logarithmus von $S(X^{(i)})$ ist der gewichtete Mittelwert $W^*(X^{(i)}|X^{(1)}, \dots, X^{(i-1)})$ dieser bedingten Verdoppelungsraten über alle möglichen Entwicklungen der ersten $i-1$ Tage, und für den Erwartungswert des Logarithmus von S_n erhalten wir die Kettenregel

$$W^*(X^{(1)}, \dots, X^{(n)}) = \mathbb{E}(\log_2 S_n) = \sum_{i=1}^{n} W^*(X^{(i)}|X^{(1)}, \dots, X^{(i-1)})\,.$$

In Analogie zur Entropierate eines stochastischen Prozesses sagen wir

Definition: Die optimale Wachstumsrate ist

$$W^*_\infty \underset{\text{def}}{=} \lim_{n\to\infty} \frac{W^*(X^{(1)}, \dots, X^{(n)})}{n}\,,$$

sofern dieser Grenzwert existiert.

Genau wie im Falle der Entropierate können wir auch hier zeigen

Satz: In einen stationären Markt existiert die Wachstumsrate und ist gleich $\lim\limits_{n\to\infty} W^*(X^{(n)}|X^{(1)}, \dots, X^{(n-1)})$.

Beweis: Da die Wachstumsrate bei mehr Information nicht kleiner werden kann und wir einen stationären Prozess vorausgesetzt haben, ist

$$\begin{aligned} W^*(X^{(n+1)}|X^{(1)}, \dots, X^{(n)}) &\geq W^*(X^{(n+1)}|X^{(2)}, \dots, X^{(n)}) \\ &= W^*(X^{(n)}|X^{(1)}, \dots, X^{(n-1)})\,, \end{aligned}$$

die Folge der $W^*(X^{(n)}|X^{(1)}, \dots, X^{(n-1)})$ ist also monoton wachsend. Somit ist sie entweder konvergent oder divergiert bestimmt gegen $+\infty$. Da nach der Kettenregel

$$\frac{W^*(X^{(1)}, \dots, X^{(n)})}{n} = \frac{1}{n} \sum_{i=1}^{n} W^*(X^{(i)}|X^{(1)}, \dots, X^{(i-1)})$$

ist, hat die linke Seite nach der Mittelwertsatz von CESÁRO *(Kap. 1, §5c), Schritt 2)* denselben Grenzwert. ■

Um auch bei stationären Märkten die log-optimale Strategie mit anderen Anlagestrategien vergleichen zu können, müssen wir uns an einen Begriff aus der Stochastik erinnern:

Definition: $\mathcal{Y} = Y_1, Y_2, \dots$ und $\mathcal{Z} = Z_1, Z_2, \dots$ seien zwei stochastische Prozesse. $\mathcal{Z}$ heißt *Martingal* bezüglich $\mathcal{Y}$, wenn für alle $k < \ell$ gilt $\mathbb{E}(Z_\ell|Y_1, \dots, Y_k) = Z_k$. Er heißt *Supermartingal*, wenn $\mathbb{E}(Z_\ell|Y_1, \dots, Y_k) \leq Z_k$ ist; entsprechend heißt er *Submartingal*, falls $\mathbb{E}(Z_\ell|Y_1, \dots, Y_k) \geq Z_k$.

(Tatsächlich betrachtet man in der Stochastik Martingale meist allgemeiner bezüglich einer beliebigen Filtration auf dem Wahrscheinlichkeitsraum; für uns genügt aber dieser Spezialfall.)

Satz: $\mathcal{X}$ sei ein stationärer stochastischer Prozess und S_n^* beschreibe die Wertentwicklung der ersten n Tage einer Anlage bezüglich der bedingt log-optimalen Strategie, S_n die bezüglich irgendeiner beliebigen kausalen Strategie. Dann ist die Folge $\mathcal{Q}$ der Quotienten S_n/S_n^* ein positives Supermartingal bezüglich $\mathcal{X}$.

Zum *Beweis* müssen wir den Erwartungswert

$$\mathbb{E}\left(\frac{S_{n+1}}{S_{n+1}^*}\,\middle|\, X^{(1)}, \dots, X^{(n)}\right)$$

abschätzen. Da wir bei Kenntnis von $X^{(1)}, \dots, X^{(n)}$ die Wertentwicklungen S_n und S_n^* kennen, können wir diese vor den Erwartungswert ziehen und erhalten

$$\mathbb{E}\left(\frac{S_{n+1}}{S^*_{n+1}}\,\middle|\, X^{(1)},\dots,X^{(n)}\right) = \mathbb{E}\left(\frac{\langle b^{(n+1)},X^{(n+1)}\rangle\, S_n}{\langle b^{*(n+1)},X^{(n+1)}\rangle\, S^*_n}\,\middle|\, X^{(1)},\dots,X^{(n)}\right)$$

$$= \frac{S_n}{S^*_n}\mathbb{E}\left(\frac{\langle b^{(n+1)},X^{(n+1)}\rangle}{\langle b^{*(n+1)},X^{(n+1)}\rangle}\,\middle|\, X^{(1)},\dots,X^{(n)}\right) \le \frac{S_n}{S^*_n}\,,$$

denn wie wir in Abschnitt *c)* gesehen haben, ist der letzte Erwartungswert für das log-optimale Portfolio kleiner oder gleich eins. Induktiv folgt die entsprechende Aussage für Index $n+k$ statt $n+1$ und damit die Supermartingaleigenschaft. ■

Nach dem Martingalkonvergenzsatz von JOSEPH DOOB (1910–2004) folgt aus der Tatsache, daß die Folge der S_n/S^*_n ein Supermartingal ist, daß diese Folge gegen eine Zufallsvariable konvergiert deren Erwartungswert höchstens gleich dem von S_1/S^*_1 ist, also kleiner oder gleich eins. Aus KOLMOGOROVs Verallgemeinerung der MARKOV-Ungleichung folgt daraus wiederum, daß für alle $t>1$ gilt

$$p\left(\sup_n \frac{S_n}{S^*_n} \ge t\right) \le \frac{1}{t}\,.$$

Die Wahrscheinlichkeit, daß irgendeine andere Strategie langfristig dramatisch bessere Ergebnisse liefert als die bedingte log-optimale ist somit relativ gering.

Kurzfristig freilich gibt es Fälle, in denen andere Strategien mit hoher Wahrscheinlichkeit zumindest etwas bessere Ergebnisse liefern als die log-optimale: An einer „Börse“ mit nur zwei Aktien und

$$X = (X_1,X_2) = \begin{cases} \left(1,\frac{1}{1-\varepsilon}\right) & \text{mit Wahrscheinlichkeit } 1-\varepsilon \\ (1,0) & \text{mit Wahrscheinlichkeit } \varepsilon \end{cases}$$

ist für das Portfolio $b=(1,0)$

$$\mathbb{E}\left(\frac{X_1}{\langle b,X\rangle}\right) = \mathbb{E}\left(\frac{X_1}{X_1}\right) = 1 \quad \text{und}$$

$$\mathbb{E}\left(\frac{X_2}{\langle b,X\rangle}\right) = \mathbb{E}\left(\frac{X_2}{X_1}\right) = \frac{1-\varepsilon}{1-\varepsilon} = 1\,,$$

nach dem Kriterium aus Abschnitt *c)* ist b also log-optimal. Trotzdem erzielt man mit dem Portfolio $(0,1)$ mit Wahrscheinlichkeit $1-\varepsilon$ ein besseres Ergebnis. Nach n Börsentagen ist allerdings die Wahrscheinlichkeit dafür, daß man mit dieser Strategie noch nicht pleite gegangen ist, gleich $(1-\varepsilon)^n$, was für große n eine ziemlich kleine Zahl ist.

Zusammenfassend läßt sich sagen, daß log-optimale Portfolios relativ sicher sind und auf lange Sicht die besten Erwartungswerte liefern; mit relativ hoher Wahrscheinlichkeit – aber keinesfalls sicher! – liefern sie zumindest langfristig auch ein besseres Ergebnis. Wie allerdings unter anderem der Wirtschaftsnobelpreisträger von 1979 PAUL A. SAMUELSON (1915–2009) mehrfach gegen sie argumentierte, hat jeder Anleger seine eigenen Vorstellungen vom Umgang mit Risiken, so daß sie nicht unbedingt *die* Anlagestrategie für jedermann sind. *Risikofrei* sind sie definitiv nicht; wie SAMUELSON in seiner Arbeit *Why we should not make mean log of wealth big though years to act are long* (Journal of Banking and Finance **3** (1979), 305–307) sagt (Sein N ist unser n):

> When you lose – and you *sure can* lose – with N large, you can lose real big. Q.E.D.

§ 3: Universelle Portfolios

Die log-optimalen Portfolios des letzten Paragraphen waren definiert in Bezug auf die Verteilungsfunktion für die Wertentwicklung an der Börse. Wie bereits dort erwähnt, ist über diese Funktionen nur wenig bekannt, und wie wir gesehen haben, führt eine falsche Schätzung schnell zu einer deutlich kleineren Verdoppelungsrate.

In diesem Paragraphen wollen wir Ansätze betrachten, die keinerlei Informationen über die Verteilungsfunktionen voraussetzen und als Information für die Wahl eines Portfolios am n-ten Tag höchstens die Wertentwicklungsvektoren $x^{(1)},x^{(2)},\dots,x^{(n-1)}$ der Vortage benutzen.

Wir suchen somit ein kausales Portfolio, d.h. eine Familie b von Portfolios $b^{(i)}(x^1,\dots,x^{(i-1)})$ derart, deren Wertentwicklungen

$$S_n = \prod_{i=1}^n \sum_{k=1}^m b_k^{(i)}(x^1,\dots,x^{(i-1)})x_k^{(i)}$$

in einem noch zu präzisierenden Sinne „möglichst gut“ sein sollen.

Entsprechende Anlagestrategien werden als *universelle Portfolios* bezeichnet; je nachdem, ob wir von einem festen n ausgehen oder uns eher für die Asymptotik von S_n interessieren, reden wir von einem *universellen Portfolio mit festem Horizont* oder einem *horizontfreien universellen Portfolio.*

Wir wollen uns hier auf den einfachsten Fall beschränken, ein von THOMAS COVER vorgeschlagenes universelles Portfolio mit festem Horizont.

THOMAS M. COVER wurde 1938 im kalifornischen San Bernardino geboren. Er studierte zunächst Physik am Massachusetts Institute of Technology; nach seinem Bachelorabschluß 1960 wechselte er zur Elektrotechnik an die Stanford University, wo er 1961 seinen Master und 1964 seinen PhD bekam. Er blieb in Stanford, zunächst als Assistant Professor und Elektrotechnik, dann als Associate Professor, ab 1971 auch für Statistik; seit 1972 hat er einen Lehrstuhl für Elektrotechnik und Statistik, seit 1994 einen *endowed chair.* Er war unter anderem schon Präsident von IEEE, der internationalen Berufsorganisation der Elektro- und Elektronikingenieure, und beratender Statistiker der kalifornischen Staatslotterie. Die meisten kennen ihn wohl als Autor des Buchs *Elements of Information Theory* (zusammen mit JOY THOMAS), das den ersten beiden Kapiteln dieser Vorlesung zu Grunde liegt.

COVERs Ansatz besteht darin, daß er Investoren betrachtet, die mit einem festen Portfolio $b = (b_1, \dots, b_m)$ arbeiten. Ein derartiger Investor realisiert eine Wertentwicklung von

$$S_n(b, x) = \prod_{i=1}^{n} \left\langle b, x^{(i)} \right\rangle = \prod_{i=1}^{n} \sum_{k=1}^{m} b_k x_k^{(i)} ,$$

die natürlich stark davon abhängt, wie gut das Portfolio b an den Börsenverlauf $x = (x^{(1)}, \dots, x^{(n)})$ angepaßt ist. Er möchte sich mit dem erfolgreichsten dieser Investoren vergleichen, also mit einem, der ein Portfolio b^* anwendet, für das $S_n(b^*, x) \ge S_n(b, x)$ ist für alle möglichen Wahlen von b. Ein solches b^* existiert, da die Menge $\mathcal{B}$ aller möglicher Portfolios kompakt ist und $S_n(b, x)$ stetig. Es garantiert zwar keinen Gewinn – bei einem großen Börsencrash wie 1929, 1987 oder 2008 wird man auch mit b^* Verluste machen – aber man kann auch im Verlustfall sicher sein, daß kein anderes konstantes Portfolio einen geringeren Verlust ergeben hätte.

Nun kann allerdings THOMAS COVER genauso wenig in die Zukunft sehen wie der Rest von uns; er weiß also, daß seine Chancen, die nur mit vorheriger Kenntnis aller $x^{(i)}$ realisierbare Strategie b^* zu treffen, äußerst gering sind. Stattdessen versucht er, eine kausale Strategie zu finden, bei der sich die Wertentwicklung möglichst wenig von der des Portfolios b^* unterscheidet.

Diese Wertentwicklung hängt ebenfalls ab von sämtlichen $x^{(i)}$; im Voraus kann man höchstens versuchen, beispielsweise den Erwartungswert des Quotienten $S_n(b, x)/S_n(b^*, x)$ zu optimieren. Dessen Berechnung setzt jedoch voraus, daß wir die Wahrscheinlichkeitsverteilung für die $x^{(i)}$ kennen, und in diesem Paragraphen wollen wir ja ohne deren Kenntnis auskommen. Deshalb wählt COVER ein anderes Kriterium: Selbst bei der aus Sicht unserer Strategie schlimmstmöglichen Börsenentwicklung soll $S_n(b, x)/S_n(b^*, x)$ möglichst klein sein.

Die Grundidee für seinen Ansatz ist einfach: Falls wir im Voraus wüßten, wie sich der Markt entwickelt, würden wir jeden Morgen das gesamte Anlagekapital in die Aktie investieren, die an jeweiligen Tag die beste Wertsteigerung (oder, bei einem Crash, den geringsten Wertverlust) bringt. Da wir diese Aktie aber erst am Abend kennen, betrachten wir stattdessen *alle* Strategien, die jeden Tag das gesamte vorhandene Kapital auf eine einzige Aktie setzen, für jede der n^m Funktionen

$$j \colon \{1, \dots, n\} \to \{1, \dots, m\}$$

also die Strategie, die am i-ten Tag alles auf die Aktie $j(i)$ setzt; die Menge alle dieser Funktionen j bezeichnen wir mit $\mathcal{J}$.

Für sich allein betrachtet kann jede dieser Strategien sehr riskant sein; mit einem „Portfolio“, in dem *alle* diese Strategien vertreten sind, sollten wir aber in der Lage sein, einen vernünftigen Kompromiss zwischen Wachstum und Risiko zu erreichen.

Wir nehmen also an, daß wir für jede der n^m Funktionen j einen gewissen Anteil $w(j)$ des Ausgangskapitals in die entsprechende Anlagestrategie investieren. Dieser Anteil wird während der n-tägigen Laufzeit multipliziert mit $x_{j(1)}^{(1)} \cdots x_{j(n)}^{(n)}$; das Gesamtkapital wird also multipliziert

mit
$$S(w,x) \underset{\text{def}}{=} \sum_{j\in\mathcal{J}} w(j) x^{(1)}_{j(1)} \cdots x^{(n)}_{j(n)} \,.$$
Das Kapital eines Investors, der auf ein festes Portfolio $b = (b_1, \ldots, b_m)$ setzt wird multipliziert mit
$$S(b,x) = \prod_{i=1}^{n} \sum_{k=1}^{m} b_k x_k^{(i)} = \sum_{j\in\mathcal{J}} \prod_{i=1}^{n} \bigl(b_{j(i)} x^{(i)}_{j(i)}\bigr) = \sum_{j\in\mathcal{J}} \prod_{i=1}^{n} b_{j(i)} \prod_{i=1}^{m} x^{(i)}_{j(i)} \,;$$
das Verhältnis zwischen den beiden Wertentwicklungen ist also
$$\frac{S(w,x)}{S(b,x)} = \frac{\displaystyle\sum_{j\in\mathcal{J}} w(j) \prod_{i=1}^{m} x^{(i)}_{j(i)}}{\displaystyle\sum_{j\in\mathcal{J}} \prod_{i=1}^{n} b_{j(i)} \prod_{i=1}^{m} x^{(i)}_{j(i)}} \,.$$
Dieses Verhältnis möchten wir selbst für das beste feste Portfolio b^* möglichst groß werden lassen; wir müssen allerdings realistischerweise davon ausgehen, daß es praktisch immer kleiner als eins sein wird: b^* wird schließlich im Nachhinein berechnet aufgrund der tatsächlichen Wertentwicklung der Aktien, und obwohl wir mit unserem kausalen Portfolio eine größere Flexibilität haben als ein Investor mit einem konstanten Portfolio, wird das *optimale* konstante Portfolio eine Wertsteigerung haben, die wir mit unserer fehlenden Information über die künftige Entwicklung der Aktien nur mit einer verschwindend geringen Wahrscheinlichkeit übertreffen können. Wir reden hier also von der Maximierung einer Größe, die im Allgemeinen deutlich unter eins liegen wird.

Wir könnten uns als eventuell realisierbares Ziel setzen, daß wir Tag für Tag im Durchschnitt wenigstens einen gewissen Prozentsatz der Wertsteigerung des optimalen konstanten Portfolios erreichen können, aber zumindest für große n wäre selbst das ein sehr unbefriedigendes Ergebnis: Auch die Folgen $0{,}9^n$ und $0{,}99^n$ konvergieren schließlich recht schnell gegen null.

Auch die Strategie von COVER führt auf eine Wertentwicklung, die fast sicher langfristig schlechter sein wird als die mit dem im Nachhinein

berechneten optimalen konstanten Portfolio, aber das Verhältnis der Wertentwicklungen geht zumindest langfristig sehr viel langsamer gegen null als jede Folge $(q^n)_{n\in\mathbb{N}}$ für irgendein $q < 1$.

Gemäß der Philosophie von COVER wollen wir auch im schlimmsten Fall noch einen möglichst großen Quotienten haben. Eine Abschätzung für diesen schlimmsten Fall liefert uns das folgende

Lemma: $p_i, q_i \ge 0$ seien reelle Zahlen, und für mindestens ein i sei $(p_i, q_i) \neq (0,0)$. Dann ist
$$\frac{\displaystyle\sum_{i=1}^{n} p_i}{\displaystyle\sum_{i=1}^{n} q_i} \ge \min_i \frac{p_i}{q_i} \,,$$
wobei für die Minimumsbildung nur Indizes i berücksichtigt werden mit $(p_i, q_i) \neq (0,0)$.

Beweis: j sei der Index, für den p_j/q_j minimal ist. Falls q_j verschwindet, ist $p_j/q_j = \infty$, und wir müssen nichts beweisen. Falls p_j verschwindet, haben wir die triviale Aussage, daß der Quotient links nicht negativ sein kann. Somit können wir uns beschränken auf den Fall, daß p_j und q_j beide positiv sind. Dann ist
$$\frac{\displaystyle\sum_{i=1}^{n} p_i}{\displaystyle\sum_{i=1}^{n} q_i} = \frac{p_j}{q_j} \frac{\displaystyle\sum_{i=1}^{n} \frac{p_i}{p_j}}{\displaystyle\sum_{i=1}^{n} \frac{q_i}{q_j}} \,.$$
Nach Wahl von j ist $p_i/q_i \ge p_j/q_j$, also auch $p_i/p_j \ge q_i/q_j$; im zweiten Bruch ist daher der Zähler größer oder gleich dem Nenner, woraus die Behauptung folgt.

■

Auf unsere Situation angewandt, liefert dieses Lemma die Abschätzung

$$\frac{S(w,x)}{S(b,x)} = \frac{\sum\limits_{j\in\mathcal{J}} w(j)\prod\limits_{i=1}^{m} x_{j(i)}^{(i)}}{\sum\limits_{j\in\mathcal{J}}\prod\limits_{i=1}^{n} b_{j(i)}\prod\limits_{i=1}^{m} x_{j(i)}^{(i)}} \geq \min_{j\in\mathcal{J}} \frac{w(j)\prod\limits_{i=1}^{m} x_{j(i)}^{(i)}}{\prod\limits_{i=1}^{n} b_{j(i)}\prod\limits_{i=1}^{m} x_{j(i)}^{(i)}} = \min_{j\in\mathcal{J}} \frac{w(j)}{\prod\limits_{i=1}^{n} b_{j(i)}}\,.$$

Da wir nicht wissen, für welches j das Minimum angenommen wird, müssen wir dafür Sorge tragen, daß die Quotienten, über die wir hier das Minimum bilden, *allesamt* nicht zu klein werden. Dies erreicht COVER dadurch, daß er $w(j)$ proportional zum Maximalwert des Nenners wählt:

$$w(j) \underset{\text{def}}{=} c \max_{\substack{b\in\mathbb{R}^m_{\geq 0}\\ b_1+\cdots+b_m=1}} \prod_{i=1}^{n} b_{j(i)}\,,$$

wobei die Proportionalitätskonstante natürlich so gewählt werden muß, daß die Summe aller $w(j)$ gleich eins ist.

Bezeichnet $n_k = n_k(j)$ die Anzahl der Tage, an denen $j(i) = k$ ist, können wir den Nenner auch schreiben als

$$\prod_{i=1}^{n} b_{j(i)} = \prod_{k=1}^{m} b_k^{n_k}\,;$$

wir müssen also die Funktion $f(b) = \prod_{k=1}^{m} b_k^{n_k}$ maximieren unter der Nebenbedingung $g(b) = \sum_{k=1}^{m} b_k = 1$. Der Gradient von g ist der konstante Vektor mit lauter Einsen; im Maximum müssen daher nach LAGRANGE alle partiellen Ableitungen

$$\frac{\partial f}{\partial b_k} = n_k b_k^{n_k-1}\prod_{\substack{\ell=1\\ \ell\neq k}}^{m} b_k^{n_k} = \frac{n_k}{b_k} f(b)$$

übereinstimmen, d.h. alle Quotienten n_k/b_k müssen gleich sein. Da die Summe der n_k gleich n, die der $b_k = 1$ ist, folgt $b_k = \frac{n_k}{n}$; der Maximalwert von f ist also

$$f\left(\frac{n_1}{n},\ldots,\frac{n_m}{n}\right) = \prod_{k=1}^{m}\left(\frac{n_k}{n}\right)^{n_k} = n^{-n}\prod_{k=1}^{n} n_k^{n_k}\,.$$

Der Logarithmus hiervon ist

$$\sum_{k=1}^{m} n_k \log\frac{n_k}{n} = n\sum_{k=1}^{m}\frac{n_k}{n}\log\frac{n_k}{n}\,;$$

wir können den Maximalwert also auch schreiben als

$$2^{-nH\left(\frac{n_1}{n},\ldots,\frac{n_m}{n}\right)}$$

mit der SHANNONschen Entropiefunktion

$$H(p_1,\ldots.p_m) = -\sum p_k \log_2 p_k\,.$$

Damit setzen wir also

$$w(j) = c\prod_{k=1}^{m}\left(\frac{n_k(j)}{n}\right)^{n_k(j)}\,,$$

und um das auch wirklich berechnen zu können, müssen wir noch die Konstante c bestimmen.

Unter den m^n Funktionen $j\colon\{1,\ldots,n\}\to\{1,\ldots,m\}$ gibt es

$$\binom{n}{n_1,\ldots,n_m} = \frac{n!}{n_1!\cdots n_m!}$$

Funktionen, die n_1-mal den Wert 1, n_2-mal den Wert 2, …, n_m-mal den Wert m annehmen; somit ist

$$1 = \sum_{j\in\mathcal{J}} w(j) = c\sum_{\substack{n_1,\ldots,n_m\\ n_1+\cdots+n_m=n}}\frac{n!}{n_1!\cdots n_m!}$$

und

$$c^{-1} = \sum_{\substack{n_1,\ldots,n_m\\ n_1+\cdots+n_m=n}}\frac{n!}{n_1!\cdots n_m!}\left(\frac{n_1}{n}\right)^{n_1}\cdots\left(\frac{n_m}{n}\right)^{n_m}\,,$$

wobei natürlich alle Summationsindizes $n_j \geq 0$ sein müssen.

Damit können wir die Zahlen $w(j)$ berechnen; tatsächlich arbeiten wir aber natürlich nicht mit m^n verschiedenen Anlagen, von denen wir täglich (fast) jede von Aktie $j(i)$ auf Aktie $j(i+1)$ umschichten; handhabbar wird die Strategie erst, wenn wir wissen, wie wir jeden Tag das *gesamte* vorhandene Kapital auf die m Aktien verteilen.

Der Teil des Anfangskapitals, der nach Strategie $j \in \mathcal{J}$ investiert wird, trägt genau dann am Tag i zum Portfolio für die Aktie k bei, wenn $j(i) = k$ ist. Zu Beginn des i-ten Tages ist der nach dieser Strategie allerdings nicht mehr einfach der Anteil $w(j)$ des Ausgangskapitals, denn der wurde ja an jedem der Vortage multipliziert mit der Wertentwicklung jeder Akte, die j für diesen Tag aussucht. Das Kapital, das zu Beginn des i-ten Tages in Aktie k investiert wird, ist somit das Ausgangskapital mal

$$\sum_{\substack{j \in \mathcal{J} \\ j(i)=k}} w(j) \prod_{\ell=1}^{i-1} x_{j(\ell)}^{(\ell)} .$$

Damit kennen wir den absoluten Betrag des Investments in Aktie k; um das Portfolio für den i-ten Tag zu bekommen, müssen wir noch durch den gesamten zur Verfügung stehenden Betrag dividieren und erhalten

$$\hat{b}_k^{(i)} = \frac{\displaystyle\sum_{\substack{j \in \mathcal{J} \\ j(i)=k}} w(j) \prod_{\ell=1}^{i-1} x_{j(\ell)}^{(\ell)}}{\displaystyle\sum_{j \in \mathcal{J}} w(j) \prod_{\ell=1}^{i-1} x_{j(\ell)}^{(\ell)}} ,$$

wobei $w(j)$ aus der obigen Definition übernommen wird.

Dies definiert ein kausales Portfolio $\hat{b}$, denn um das Portfolio für den i-ten Tag zu berechnen, benötigen wir nur Informationen über die Kursentwicklung der Vortage; außerdem wissen wir, daß für jedes, konstante Portfolio b, insbesondere auch für das Portfolio b^*, das sich nach Ablauf der n Tage als das beste herausstellt, das Verhältnis der Wertentwicklungen $S(\hat{b}, x)$ und $S(b^*, x)$ für jeden Börsenverlauf x größer oder gleich c ist.

Was dieses Resultat wert ist, können wir freilich erst beurteilen, wenn wir wissen, wie sich c als Funktion von n und m verhält. Im Buch von COVER wird zitiert, daß sich c asymptotisch verhält wie ein konstantes Vielfaches von $(n+1)^{-(m-1)/2}$; zumindest für hinreichend große Werte von n ist das besser als jede Folge q^n mit $q < 1$.

Speziell für $m = 2$ ist

$$c^{-1} = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{k}{n}\right)^k \left(\frac{n-k}{n}\right)^{n-k}$$