Die Moment-Ertrag-Relation und optimale Portfolioauswahl

Andreas de Vries

FH Südwestfalen University of Applied Sciences, Hagen

Version: September 2001

Zusammenfassung. In dem vorliegenden Beitrag wird ein Zusammenhang zwischen dem informationstheoretischen Begriff der Information und dem Risikobegriff der Portfoliotheorie hergestellt. Dazu werden Optimierungsprobleme, den zu erwartenden Ertrag eines aus Subportfolios bestehenden Gesamtportfolios unter spezifischen Nebenbedingungen zu maximieren, formuliert und gelöst. Neben einer Anwendung für das zentrale Controlling eines Finanzunternehmens bietet das Resultat die Grundlage für einen Ansatz, den monetären Wert von Information zu bestimmen.

Schlüsselwörter: Information, Risiko, Moment-Ertrag-Relation, Portfolio, Stochastik, Finanzmathematik

Abstract. In the present paper a connection between the notion of information and the notion of risk in portfolio selection theory is shown. For this purpose, optimization problems to maximize, under specific boundary values, the expected return of a portfolio consisting of *n* subportfolios is formulated and solved. Besides an application for the central controlling of a financial institute, the result yields the basis for a new ansatz to determine the monetary value of information.

Key words: information, risk, moment-return relation, portfolio, stochastics, financial mathematics

1 Einleitung

Kürzlich konnte der Autor [19] nachweisen, wie der zu erwartende Ertrag eines Portfolios mit diversifizierter Verteilung statistischer Momente und investierten Kapitals auf Subportfolios informationstheoretisch interpretiert werden kann. Damit wurde ein neuer Zusammenhang zwischen Information und Risiko entdeckt. Kennt man die Verteilung des investierten Kapitals auf die einzelnen Subportfolios sowie diejenige der statistischen Momente, so ist die zu erwartende Rendite des Gesamtportfolios gleich einer einfachen Funktion der Entropien beider Verteilungen. Insbesondere ist

$$\text{Renditeverlust} = \text{Informationsdifferenz}/\text{Zeit}.$$

Im folgenden wird dieser Zusammenhang näher erläutert.

2 Information

Sei $p = (p_1, \dots, p_n)$ eine diskrete Wahrscheinlichkeitsverteilung, d.h. $p_i > 0$ und $\sum_{i=1}^n p_i = 1$ ([2] §4). Oft wird p auch kurz *Wahrscheinlichkeitsvektor* genannt [6] §102. Die *durchschnittliche Information* $H(p)$ ist definiert als

$$H(p) = -\sum_{i=1}^{n} p_i \log p_i, \tag{1}$$

vgl. [5, 11, 16]. Wir werden in diesem Beitrag unter „log" den Logarithmus zur Basis 2 verstehen („log2"), der oft auch mit „ld" bezeichnet wird. $H(p)$ ist somit die Information von p in Bit. Da stets $0 < p_i < 1$, ist die Information $H(p)$ für eine diskrete Wahrscheinlichkeitsverteilung p immer eine nichtnegative Größe, $H(p) \geqq 0$.

Information hängt also lediglich von einer gegebenen Wahrscheinlichkeitsverteilung p ab und bezieht sich nicht auf den Inhalt oder die Bedeutung der zu Grunde liegenden Ereignisse. Die *Wahrscheinlichkeiten*, mit denen die Ereignisse eintreten, sind relevant, nicht die Ereignisse selber. Die Semantik liegt implizit bereits in der Bestimmung der Wahrscheinlichkeiten.

Die Definition von Information kann auf stetige Wahrscheinlichkeitsverteilungen verallgemeinert werden, vgl. [16] §20 oder [11] §5.

Seien nun p und q Wahrscheinlichkeitsvektoren mit der Eigenschaft, dass q_i nur da verschwindet, wo auch p_i gleich Null ist. Dann ist die *Kullback-Leibler-Information* oder die *relative Entropie* von p bezüglich q definiert durch

$$K(p;q) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i}, \tag{2}$$

vgl. [1, 4, 10]. Die Kullback-Leibler-Information ist ein Maß für die Abweichung zweier Verteilungen. Hierbei kann die Wahrscheinlichkeitsverteilung q aufgefasst werden als mathematische Beschreibung des *A-priori*-Wissens oder der anfänglichen Vorkenntnis (*Bayesian prior distribution* [8, 9]), die aufgrund neuer Erkenntnisse, etwa durch eine Messung oder durch Lernen, zu einer geänderten Wahrscheinlichkeitsverteilung p führt; die Kullback-Leibler-Information ist dann der entsprechende Informationsgewinn.

In diesem Beitrag wird die Kullback-Leibler-Information allgemein als die Informationsdifferenz aufgefasst, die das „Wissen“ p von dem „Wissen“ q unterscheidet.

3 Die Moment-Ertrag-Relation

Üblicherweise wird in der Finanzmathematik angenommen, dass die zeitliche Entwicklung des Ertrags eines Wertpapiers (einer Aktie, einer Anleihe, aber auch eines Derivats wie z.B. einer Option) auf einem sogenannten stochastischen Prozess X_τ beruht. Unter der milden Bedingung, dass für diesen Prozess ein stochastisches Integral überhaupt nur definiert ist, dass also

$$x(t) = \int_0^t X_\tau \, \mathrm{d}\tau, \tag{3}$$

existiert, konnte der Autor [19] nachweisen, dass der zu erwartende Ertrag R des in das Wertpapier investierten Kapitals W über den Zeithorizont t durch die Moment-Ertrag-Relation

$$\boxed{R = \frac{W}{t} \, \ln[1+\mu t + \sigma x(t)],} \tag{4}$$

gegeben ist. Hierbei sind die beiden zeitabhängigen Funktionen μ und σ durch die Integrale der beiden ersten Momente des stochastischen Prozesses gegeben:

$$\mu = \mu(t) = \int_0^t \widetilde{\mu}(\tau) \, \mathrm{d}\tau, \quad \sigma = \sigma(t) = \int_0^t \widetilde{\sigma}(\tau) \, \mathrm{d}X_\tau, \tag{5}$$

wo $\widetilde{\mu}$ die erwartete Änderungsrate des Prozesses (die „Drift“) bezeichnet, und $\widetilde{\sigma}$ seine bedingte Standardabweichung (die „Diffusion“); nähere Details zu stochastischen Prozessen findet man z.B. in [2, 14].

Üblicherweise wird für X_τ eine Brownsche Bewegung angenommen, so dass der Wertpapierkurs einen so genannten Ito-Prozess ergibt [3, 7, 13, 17, 20]. Für die Betrachtungen des vorliegenden Beitrags sind diese Annahmen und Zusammenhänge jedoch nicht weiter relevant.

Folgerung. Die Ertragsfunktion genügt dem Gesetz des sinkenden Grenzertrags sowohl bei steigendem erwarteten Ertrag μ als auch bei steigendem Risiko σ, d.h. sie ist konkav bezüglich der stochastischen Momente. Wegen der Eigenschaft des Logarithmus

$$\ln(1+x) = x - x^2/2 + x^3/3 - x^4/4 \pm \ldots, \tag{6}$$

gilt für $u = \mu t + \sigma x(t)$, dass $\log_2(1+u) \approx u/\ln 2$ für $u \ll 1$, und $\log_2(1+u) \approx \log_2 u$ für $u \gg 1$. Das bedeutet, dass für kleine Momente $u \ll 1$ sich der Ertrag etwa proportional zu u verhält (d.h. eine Verdopplung der Momentfunktion u beispielsweise bewirkt in etwa eine Verdopplung des Ertrags), während für extrem hohe Momente $u \gg 1$ eine weitere Erhöhung eine geringer ausfallende Erhöhung des Ertrags nach sich zieht (d.h. eine Verdopplung bewirkt im Wesentlichen nur eine Addition um 1 Bit).

4 Variation der Momente und Investitionsanteile bei *n* Subportfolios

Es wird folgendes Modell zu Grunde gelegt: Ein Portfolio Π bestehe aus n Subportfolios Π_i, $i = 1, \ldots, n$, in denen jeweils der Betrag $p_i W$ mit $p_i \geqq 0$ und $\sum_i p_i = 1$ investiert wird (so dass die Gesamtsumme der investierten Kapitalbeträge also genau W ergibt: $\sum_i p_i W = W$). Die erwarteten Erträge seien nach (4) beziffert mit

$$R_i = \frac{p_i W \ln 2}{t} \log s_i, \quad \text{wo } s_i = [1+\mu_i t + \sigma_i x(t)]. \tag{7}$$

Der zu erwartende Gesamtertrag R des Portfolios beläuft sich auf $R = \sum_i R_i$.

Wir nehmen an, dass das investierte Kapital des i-ten Subportfolios $W_i = p_i W$ nichtnegativ ist. Definieren wir noch den Wahrscheinlichkeitsvektor $q = (q_1, \ldots, q_n)$ mit

$$q_i = s_i/s, \qquad \text{wo } s = \sum_{j=1}^n s_j. \tag{8}$$

Formal können p und q als diskrete Wahrscheinlichkeitsdichten auf dem abstrakten Phasenraum $\Omega = \{\Pi_1, \ldots, \Pi_n\}$ der Subportfolios Π_i aufgefasst werden. Damit können wir die gesamte zu erwartende Rendite $r = R/W$ schreiben als

$$r = \frac{\ln 2}{t} \left[\sum_{i=1}^n p_i \log q_i + \log s \right]. \tag{9}$$

Mit den Gleichungen (1) und (2) kann man dies ausdrücken als

$$r = \frac{\ln 2}{t} \, [\log s - H(p) - K(p;q)]. \tag{10}$$

4.1 Variation der Investitionsanteile *p* und der Momente *s*

Gibt es keine Restriktionen, so wird der zu erwartende Ertrag r maximal, wenn das Investitionskapital und die Momente gleichverteilt sind, wenn also

$$p_i = s_i/s = 1/n \qquad \text{für jedes } i = 1, \ldots, n. \tag{11}$$

Der Beweis wird im Appendix gegeben. Mit Gleichung (10) gilt dann wegen $K(p;q) \geqq 0$ und $H(p) = \log n$:

$$r = \frac{\ln 2}{t} \, [\log s - H(p)] = \frac{\ln 2}{t} \, \log \frac{s}{n}. \tag{12}$$

Dieses Ergebnis, d.h. die Optimalität der Gleichverteilung, deckt sich mit dem Resultat der klassischen Portfoliotheorie, dass ein risikoaverser Investor sein Portfolio auf möglichst viele Risikopositionen diversifiziert („*Don't put all your eggs in one basket*“ [15, 17, 20]).

4.2 Variation der Investitionsanteile *p* bei festem *s*

Sei $s = \text{const}$ gegeben („Gesamtmomentenbedingung“). Mit der elementaren „integrierten Gibbs-Ungleichung“

$$\sum_{i=1}^{n} p_i \log q_i \leqq \sum_{i=1}^{n} p_i \log p_i \tag{13}$$

(z.B. [12]) gilt damit für den Gesamtertrag

$$r \leqq \frac{\ln 2}{t} \, [\log s - H(p)] \,. \tag{14}$$

Der Ausdruck in den eckigen Klammern in (10) zeigt also, dass bei gegebener Momentfunktion s der Ertrag genau dann maximal ist, wenn die Kullback-Leibler-Information verschwindet: $K(p;q) = 0$. Nach [19] ist das für zwei Wahrscheinlichkeitsdichten p und q genau dann der Fall, wenn $p = q$. Der Gesamtertrag ist daher genau dann maximal, wenn

$$p_i = s_i/s \qquad \text{für jedes } i = 1, \dots, n. \tag{15}$$

Geometrisch bedeutet Gleichung (15), dass der „Investmentvektor“ $(p_1, \dots, p_n)$ die gleiche Richtung hat wie der „Momentenvektor“ $(s_1, \dots, s_n)$.

5 Diskussion

Das Optimierungsproblem, den zu erwartenden Ertrag eines aus n Subportfolios bestehenden Gesamtportfolios durch optimale Wahl der Investionskapitalverteilung zu maximieren, wird in dem vorliegenden Beitrag für zwei unterschiedliche Randbedingungen formuliert und gelöst. Durch Gleichung (10) wird der zu erwartende Ertrag r mit der Kullback-Leibler-Information in Beziehung gebracht. Gleichung (10) erlaubt die bemerkenswerte Schlussfolgerung, dass jede Informationsdifferenz $K(p;q)$ zwischen der Kenntnis p der Investitionsverteilung und der Vorkenntnis q der Risiken den zu erwartenden Ertrag mindert. Insbesondere ist bei freier Wahl der Investitionsanteile und der Momentenverteilung die Gleichverteilung optimal.

Nach einer Idee von Stoughton und Zechner [18] lässt sich das Resultat im zentralen Controlling eines aus n Abteilungen bestehenden Finanzinstituts direkt anwenden. Hier ist das Gesamtportfolio das gesamte Unternehmen, während die Abteilungen den Subportfolios entsprechen. Das Controlling muss also bemüht sein, die Verteilung des Investitionskapitals auf die einzelnen Abteilungen der Verteilung der Momente gemäß Gleichung (11) oder, bei gegebener Gesamtmomentbedingung, gemäß (15) anzupassen.

Ein weiterer Aspekt ergibt sich durch den Ausblick auf den Fall unvollständiger Information. Normalerweise wird die Verteilung der Momente nur unvollständig bekannt sein, sei es durch prinzipielles Unwissen oder durch bewusstes *information hiding* der Manager der Subportfolios. Man sieht sofort, dass in diesem Fall die Wahl der optimalen Kapitalverteilung misslingt, denn die Informationsdifferenz drückt den Gewinn: Der zu erwartende Ertrag wird nicht maximal.

Vor allem jedoch liefert das Resultat einen Ansatz zur monetären Bewertung von Information. Denn mit Gleichung (14) sieht man, dass bei gegebener Gesamtmomentenbedingung der maximale zu erwartende Ertrag

$$r_* = \frac{\ln 2}{t} \, [\log s - H(p)]$$

lautet. Ändert sich nun die Momentenverteilung, beispielsweise durch Markt- oder Portfolioveränderungen, so ist $p \neq q$, und der Ertrag r ist durch (10) gegeben. Bei einer Änderung der Momente entspricht die Informationsdifferenz damit

$$\boxed{K(p;q) = \frac{t}{\ln 2} \, (r_* - r).} \tag{16}$$

Information kostet also, nämlich den Preis der relativen Ertragsminderung mal Zeit.

Appendix

Zu lösen ist das folgende Optimierungsproblem: Maximiere den zu erwartenden Ertrag r aus Gleichung (9). Mit den beiden Nebenbedingungen $\sum_{i=1}^{n} p_i = 1$ und $\sum_{i=1}^{n} q_i = 1$ lässt sich das Maximierungsproblem durch die Lagrange-Funktion

$$\mathscr{L} = r + \lambda_1 \left(\sum_{i=1}^{n} p_i - 1 \right) + \lambda_2 \left(\sum_{i=1}^{n} q_i - 1 \right) \tag{17}$$

ausdrücken. Die Parameter λ_1 und λ_2 sind die Lagrange-Multiplikatoren. Bildet man nun den Gradienten der Lagrange-Funktion $\mathscr{L} = \mathscr{L}(p, q, \lambda_1, \lambda_2)$, so erhält man

$$\frac{\partial \mathscr{L}}{\partial p_i} = \ln 2 \, \log q_i + \lambda_1, \tag{18}$$

weiter mit $\partial r / \partial q_i = (\partial r / \partial s_i)(\partial s_i / \partial q_i) = p_i \ln 2/(q_i t)$:

$$\frac{\partial \mathscr{L}}{\partial q_i} = \frac{\ln 2}{t} \frac{p_i}{q_i} + \lambda_2, \tag{19}$$

sowie schließlich

$$\frac{\partial \mathscr{L}}{\partial \lambda_1} = \sum_{i=1}^{n} p_i - 1, \qquad \frac{\partial \mathscr{L}}{\partial \lambda_2} = \sum_{i=1}^{n} q_i - 1. \tag{20}$$

Setzt man den Gradienten gleich Null, $\operatorname{grad} \mathscr{L} = 0$, so erhält man neben den Randbedingungen der beiden Gleichungen (20) durch Gleichung (18)

$$\lambda_1 = -\ln 2 \, \log q_i. \tag{21}$$

Da dies konstant für alle $i = 1, \dots, n$ gilt, folgt sofort, dass $q_1 = \ldots = q_n = 1/n$. Entsprechend folgt dann mit Gleichung (19), dass auch $p_1 = \ldots = p_n = 1/n$. q.e.d.

Literatur

[1] M. Avellaneda, C. Friedman, R. Holmes and D. Samperi: Calibrating volatility surfaces via relative-entropy minimization, *Applied Mathematical Finance* **4** (1997)

[2] H. Bauer: *Wahrscheinlichkeitstheorie.* de Gruyter Berlin New York 1991

[3] D. Duffie: *Dynamic Asset Pricing Theory.* Princeton University Press, Princeton 1996

[4] W. Ebeling, J. Freund and F. Schweizer: *Komplexe Strukturen: Entropie und Information.* B.G. Teubner Stuttgart Leipzig 1998

[5] W. Heise and P. Quattrocchi: *Informations- und Codierungstheorie. 3. Auflage.* Springer-Verlag Berlin Heidelberg 1995

[6] H. Heuser: *Funktionalanalysis.* B.G. Teubner Stuttgart 1986

[7] J.C. Hull: *Options, Futures, and Other Derivatives. Second Edition.* Prentice-Hall, Upper Saddle River 2000

[8] E.T. Jaynes: Information theory and statistical mechanics. *Phys. Rev.* **106**, 620-630 (1957)

[9] E.T. Jaynes: Information theory and statistical mechanics. In K. Ford (Ed.): *1962 Brandeis Summer Institute in Theoretical Physics.* W.A. Benjamin New York 1963

[10] J.N. Kapur: *Measures of Information and Their Applications.* Wiley New Delhi 1994

[11] J.C.A. van der Lubbe: *Information Theory.* Cambridge University Press, Cambridge 1997

[12] M.C. Mackey: *Time's Arrow: The Origins of Thermodynamic Behavior.* Springer-Verlag Berlin 1992

[13] S.N. Neftci: *An Introduction to Mathematics of Financial Derivatives.* Academic Press, San Diego 1996

[14] B. Øksendal: *Stochastic Differential Equations. An Introduction with Applications.* 5th edition. Springer-Verlag Berlin Heidelberg New York 1998

[15] P.A. Samuelson und W.D. Nordhaus: *Economics. 15th Edition*. McGraw-Hill, New York 1995

[16] C.E. Shannon and W. Weaver: *The Mathematical Theory of Communication.* University of Illinois Press Urbana and Chicago 1949

[17] M. Steiner und C. Bruns: *Wertpapiermanagement.* Schäffer-Poeschel Verlag Stuttgart 1994

[18] N.M. Stoughton und J. Zechner: *Optimal Allocation Using RAROC™ and EVA®*. Seminar Paper (1999) http://papers.ssrn.com/sol3/papers.cfm?abstract_id=118208

[19] A. de Vries: How to price information by Kullback-Leibler entropy and a moment-return relation for portfolios. *Int. J. Theor. Appl. Finance* **4** (3), 535–543 (2001)

[20] P. Wilmott: *Derivatives. The Theory and Practice of Financial Engineering.* John Wiley, Chichester 1998

Prof. Dr. Andreas de Vries
Wirtschaftsinformatik
FB Technische Betriebswirtschaft
FH Südwestfalen
Haldener Straße 182, D-58095 Hagen
e-mail: `de-vries@fh-swf.de`