Wissen, Information und Risiko

Andreas de Vries

FH Sdwestfalen University of Applied Sciences, Hagen

Version: 3. Oktober 2001

Zusammenfassung. Aktuell entwickeln sich sehr interessante Ansätze, informationstheoretische Methoden in der Finanzmathematik anzuwenden. In diesem Beitrag wird ein kurzer Überblick über diese Ideen und über die ihnen zu Grunde liegende Definition des Begriffs Information und seine Interpretation im Kontext mit Portfolios gegeben. Neben einer Anwendung für das zentrale Controlling eines Finanzunternehmens bietet das Resultat die Grundlage für einen theoretischen Ansatz, den Geldwert von Information zu bestimmen.

Summary. In recent time there developped some interesting ideas to apply information theoretic methods in financial mathematics. In this article, a survey about these ideas and the underlying definitions of the notion of information and its interpretation in the context of portfolios is given. It contains the information theoretic fundamentals and a mathematical relation between the statistical moments and the expected return of a financial asset developped by the author, as well as a remarkable analogy to the model of the *entropic market* by Les Gulko. Besides an application for the central controlling of a financial institution, the result yields the basis for a theoretical ansatz to evaluate information monetarily. However, the monetary value of information cannot be calculated absolutely, but only relatively with respect to the expected risk.

1 Einleitung

„Was kostet ein Byte?" Information spielt nicht nur in den wirtschaftswissenschaftlichen Theorien eine wichtige Rolle, beispielsweise im Zusammenhang mit der Effizienz von Märkten oder in der Spieltheorie. Aus betriebswirtschaftlicher Sicht erlangt Information als einer der Produktionsfaktoren immer größere Bedeutung. Die Ermittlung eines wissenschaftlich fundierten Preises von Information wäre daher von erheblichem unternehmerischen und theoretischen Interesse.

Voraussetzung dafür sind Anwendungen von Methoden der Informationstheorie in den Wirtschaftswissenschaften, in einem ersten Schritt in der Finanzmathematik. In der letzten Zeit sind es insbesondere zwei internationale Veröffentlichungen, die informationstheoretische Anwendungen in der Finanzmathematik vorschlagen. Es wurden bereits mehrfach informationstheoretische Anwendungen in der Finanzmathematik vorgeschlagen. So zeigte beispielsweise Les Gulko [6] anhand eines idealisierten Modells, wie sich Preise von Wertpapieren durch das aus der Statistischen Physik wohlbekannte „Prinzip der maximalen Entropie", einem Optimierungsproblem mit Randbedingungen, herleiten lassen. Andererseits konnte der Autor [24] mathematisch zeigen, wie der zu erwartende Ertrag eines Portfolios mit diversifizierter Verteilung statistischer Momente und investierten Kapitals auf einzelne Subportfolios informationstheoretisch interpretiert werden kann.

Im folgenden werden die Grundgedanken beider Beiträge kurz dargestellt. Dazu wird zunächst der Begriff der Information erläutert, bevor auf den Ansatz von Gulko und anschließend auf den des Verfassers eingegangen wird.

2 Information und Wissen

Was ist Information? Kaum ein Begriff, der im alltäglichen Leben so oft gebraucht (und gelegentlich strapaziert) wird. Wie mit allen fundamentalen („ersten") Begriffen [27] kann eine Definition nur vage ausfallen: *Information* ist der darstellungs-, sender- und empfängerinvariante Gehalt einer Mitteilung, einer Nachricht oder eines Ereignisses [17].

Nach Weizsäcker [26] muss zunächst zwischen zwei Arten von Information unterschieden werden: Einerseits gibt es die *faktische* Information, „die man bereits weiß", andererseits die *potentielle* oder *virtuelle* Information, „die man (noch) nicht weiß". Faktische Information, also *Wissen*, kann ausschließlich von bereits vergangenen Ereignissen stammen.[1] Im Umkehrschluss liefern zukünftige Ereignisse potentielle Information.

Eine spezifischere Antwort gibt die Informationstheorie. Sie befasst sich speziell mit der potentiellen Information, die in drei Arten zerfällt:

1. Die *syntaktische Information* bezieht sich auf die Symbole oder Zeichen, mit denen Nachrichten übertragen werden. Hieraus leitet sich direkt die Einheit der Information ab, insbesondere das Bit.

[1] Zeit und Wissen hängen also unauflöslich zusammen!

2. Die *semantische Information* bezieht sich auf die Bedeutung von Nachrichten.

3. Die *pragmatische Information* bezieht sich auf die Wirkung und den Nutzen von Nachrichten.

Lyre [16] spricht suggestiv von der „Dreidimensionalität der Information“ um zu verdeutlichen, dass die drei Informationsaspekte lediglich verschiedene Sichtweisen auf ein und dasselbe Phänomen darstellen.

Die syntaktische Information kann am präzisesten beschrieben werden. Die entsprechende mathematische Definition stammt von Shannon [21] aus dem Jahre 1948. Shannons bahnbrechende Idee war es, die Information H als Funktion einer gegebenen Wahrscheinlichkeitsverteilung p zu betrachten, $H = H(p)$, und zwar als Maß für die Unsicherheit des Eintretens eines bestimmten Ereignisses, bzw. als Maß für die entsprechende Überraschung oder den Neuigkeitswert.

Im engeren (und ursprünglich Shannonschen) Sinne der Informationstheorie bezieht sich die Wahrscheinlichkeitsverteilung lediglich auf den Spezialfall einer Quelle, die einen endlichen Vorrat an Nachrichten sendet, wobei die Nachricht Nummer i mit der Wahrscheinlichkeit p_i erzeugt wird. Auf diese Weise wird ganz offensichtlich nur der syntaktische Aspekt der Information berücksichtigt.

Doch vor allem Jaynes [11, 12] vertritt vehement die Auffassung, dass dieser Ansatz auf prinzipiell jede Wahrscheinlichkeitsverteilung verallgemeinert werden kann. Grundlage hierfür ist die Idee, dass eine Wahrscheinlichkeitsverteilung im wesentlichen A-priori-Kenntnis oder Wissen darstellt, also faktische Information. Auf diese Weise kann also auch die semantische Information mathematisch präziser gefasst werden, denn die Wahrscheinlichkeitsverteilung selbst ist formaler Ausdruck des Wissens über die betreffende Situation.

Wenn beispielsweise — wie in den Lehrbüchern üblich — für einen Würfel die Gleichverteilung $p_1 = ... = p_6 = 1/6$ angenommen wird, so wird in Wahrheit das Wissen formal ausgedrückt, dass es sich um einen *idealen* Würfel handelt. Mit anderen Worten: Die Gleichverteilung ist formaler Ausdruck für das Nicht-Wissen, welche Eigenschaften über den betrachteten Würfel vorausgesetzt werden können: also setzt man rationalerweise gar keine voraus. Weiß man dagegen mehr, beispielsweise dass der empirisch beobachtete Mittelwert eines konkreten Würfels 4,5 beträgt (und nicht 3,5!), so darf man keine Gleichverteilung mehr annehmen. Welche Verteilung aber dann? Eine mathematisch eindeutige Antwort gibt das aus der Statistischen Physik bekannte Prinzip der „maximalen Entropie“. Mit ihm lässt sich die Verteilung bei gegebenem Wissen errechnen, so wie es Jaynes [12] in seiner *Brandeis lecture* grandios vorführte.

Allerdings ist dieser Ansatz, obwohl sogar auf die konzeptionellen Ideen der Begründer der Wahrscheinlichkeitstheorie, Laplace und Bernoulli, zurückführbar [13], nicht unumstritten.

3 Information und Wahrscheinlichkeit

Wir werden in diesem Beitrag ganz pragmatisch Information als eine Funktion H einer gegebenen Wahrscheinlichkeitsverteilung *definieren*. Die Interpretation dieser Definition ist dann kontextabhängig.

Ist $p = (p_1, ..., p_n)$ ein Wahrscheinlichkeitsvektor [9], so ist die *durchschnittliche Information* oder auch die *Entropie* $H(p)$ definiert als

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

Wir verstehen in diesem Beitrag unter „log“ den Logarithmus zur Basis 2 („$\log_2$“ oder „ld“). Damit ist $H(p)$ die Information von p in der Einheit Bit.

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

Gibt es nun zwei Wahrscheinlichkeitsvektoren p und q (mit der Eigenschaft, dass q_i nur da verschwindet, wo auch p_i gleich Null ist), so ist die so genannte *Kullback-Leibler-Information* oder die *relative Entropie* von p bezüglich q definiert durch

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

Die Kullback-Leibler-Information ist ein Maß für die Abweichung zweier Verteilungen. Hierbei kann die Wahrscheinlichkeitsverteilung q aufgefasst werden als mathematische Beschreibung des Wissens oder der anfänglichen Vorkenntnis (*Bayesian prior distribution* [12]); auf Grund neuer Erkenntnisse, etwa durch eine Messung oder durch Lernen, wird das Wissen verändert und führt zu einer geänderten Wahrscheinlichkeitsverteilung p. 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 Vorwissen q unterscheidet.

4 Gulkos Entropischer Markt

Gulkos Ansatz erweitert das klassische Arbitrage-Theorem von Arrow [1]. Grob ausgedrückt besagt es, dass dann und nur dann keine Möglichkeit zur Arbitrage besteht, wenn die Marktpreise $\psi_1, \dots, \psi_n$ der n Produkte eines Marktes linear abhängen von den Wahrscheinlichkeiten $p_1, \dots, p_m$ der m möglichen „Zustände der Welt“.

Insbesondere also bestimmen die Preise im Falle von Arbitragefreiheit die Einschätzung risikoneutraler Marktteilnehmer, mit welcher Wahrscheinlichkeit die zukünftigen Weltzustände eintreten. Umgekehrt bestimmen die eingeschätzten Wahrscheinlichkeiten die Preise [4, 18].

Nach Gulkos Hypothese des Entropischen Markts nun ist diese Wahrscheinlichkeitsverteilung durch das Prinzip der maximalen Entropie bestimmbar. Entropie bezeichnet hier die durchschnittliche (potentielle) Information einer Wahrscheinlichkeitsverteilung, wie wir oben gesehen haben. Die Hypothese liefert eine hinreichende Bedingung für das wohlbekannte Black-Scholes-Modell, jedoch *ohne* die Annahme der Normalverteilungshypothese. Noch wichtiger: Unter leichten Veränderungen der Randbedingungen lassen sich sogar ganz andere Wahrscheinlichkeitsverteilungen herleiten, so beispielsweise die γ-Verteilung [7].

5 Information und Rendite

Einen anderen Aspekt beleuchtet die Frage: Was hat Information mit erwarteter Rendite zu tun? Der Zusammenhang offenbart sich erst nach zwei Schritten: Zunächst benötigt man die so genannte Moment-Ertrag-Relation *(moment-return relation)* eines beliebigen Assets oder Wertpapiers, und ferner ein Portfolio mit gegebenen Verteilungen des Investitionskapitals und der stochastischen Momente, jeweils auf einzelne Subportfolios.

5.1 Die Moment-Ertrag-Relation

In der Finanzmathematik wird angenommen, dass die zeitliche Entwicklung des Ertrags eines Wertpapiers (einer Aktie, einer Anleihe, aber auch eines Derivats wie einer Option) auf einem so genannten *stochastischen Prozess* X_τ [3, 19] beruht. (Üblicherweise ist X_τ eine Brownsche Bewegung, und der Wertpapierkurs ein Ito-Prozess [4, 18, 22, 28].) Unter der milden Bedingung, dass für diesen Prozess ein stochastisches Integral überhaupt nur definiert ist, konnte der Autor [24] nachweisen, dass der zu erwartende Ertrag R des in das Wertpapier investierten Kapitals W über den Zeithorizont t durch die *Moment-Ertrag-Relation*

$$\boxed{R = \frac{W \ln 2}{t} \log s, \qquad \text{mit } s = [1 + \mu t + \sigma x(t)]} \tag{3}$$

gegeben ist. Hierbei ist μ die erwartete Rendite, σ die Volatilität des Wertpapierkurses und $x(t)$ das stochastische Integral („nulltes Moment“) des Prozesses. Die Funktion s ist die Momentfunktion des Wertpapiers. Interessant ist folgende elementar beweisbare Folgerung: Die Ertragsfunktion genügt dem Gesetz des sinkenden Grenzertrags sowohl bei steigendem erwarteten Ertrag μ als auch bei steigendem Risiko σ.

5.2 Zusammengesetzte Portfolios

Betrachten wir nun ein Portfolio, das aus n Subportfolios besteht, in denen jeweils der Anteil p_i des zur Verfügung stehenden Gesamtkapitals investiert wird (also $p_i \geqq 0$ und $\sum_i p_i = 1$). Bezeichnen wir ferner die Momentfunktion des i-ten Subportfolios mit s_i, und die Summe der Momentfunktionen des gesamten Portfolios mit $\widehat{s}$. Die Verteilung der Momentfunktionen über die Subportfolios ist dann darstellbar durch den Vektor $q = (q_1, \ldots, q_n)$, wobei für jedes i gilt:

$$q_i = s_i / \widehat{s} \qquad \text{mit } \widehat{s} = \sum_{j=1}^{n} s_j. \tag{4}$$

Es lässt sich zeigen, dass der zu erwartende Ertrag des gesamten Portfolios mit den Informa- tionen der beiden Verteilungen zusammenhängt, und zwar durch die Formel [24]:

$$r = [\log \widehat{s} - H(p) - K(p;q)] \ \ln 2/t. \tag{5}$$

Der Verfasser [25] konnte zudem zeigen, wie der zu erwartende Ertrag bei verschiedenen Randbedingungen maximiert wird:

- Gibt es keine Restriktionen, so wird der erwartete Ertrag r maximal, wenn das Investitionskapital und die Momente gleichverteilt sind. Dieses Ergebnis, also die Optimalität der Gleichverteilung, stimmt vollständig mit dem Resultat der klassischen Portfoliotheorie überein, dass ein risikoaverser Investor sein Portfolio auf möglichst viele Risikopositionen diversifiziert *(„Don't put all your eggs in one basket“) [20, 22, 28].* Der maximale zu erwartende Ertrag ist gegeben durch die Formel

$$r_* = [\log \widehat{s} - H(p)] \ \ln 2/t = [\log \widehat{s}/n] \ \ln 2/t. \tag{6}$$

- Sei $\widehat{s} = \text{const}$ gegeben („Gesamtmomentbedingung“). Dann ist der erwartete Ertrag r maximal, wenn für die beiden Wahrscheinlichkeitsdichten p und q gleich sind: Geometrisch bedeutet das, dass der „Investmentvektor“ $(p_1, \ldots, p_n)$ die gleiche Richtung hat wie der „Momentenvektor“ $(s_1, \ldots, s_n)$.

6 Ausblick und Nutzen

Aktuell werden Anwendungen informationstheoretischer Methoden und Begriffe in der Finanzmathematik mehrfach und oft unabhängig voneinander konkretisiert. Das spricht dafür, dass die Ideen „in der Luft liegen“ und die Zeit reif ist für eine neue Entwicklung.

Gulko löst mit seiner Hypothese des Entropischen Markts das Problem der Preisbildung an Finanzmärkten, indem er den Shannonschen Informationsbegriff der Informationstheorie benutzt und damit gewissermaßen Samuelson „vom Kopf auf die Füße stellt“. Ein anderer Zusammenhang zwischen Information und Risiko ist zu erkennen, wenn man ein aus mehreren Subportfolios bestehendes Portfolio von Risikopositionen betrachtet. 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 (der durchschnittlichen Informationen) beider Verteilungen. Das Optimierungsproblem, den zu erwartenden Ertrag eines aus n Subportfolios bestehenden Gesamtportfolios durch optimale Wahl der Investitionskapitalverteilung zu maximieren, wird formuliert und gelöst. Durch Gleichung (5) der zu erwartende Ertrag r mit der Kullback-Leibler-Information in Beziehung gebracht. Gleichung (5) erlaubt die bemerkenswerte Schlussfolgerung, dass jede Informationsdifferenz $K(p;q)$ zwischen dem Wissen p der Investitionsverteilung und der Vorkenntnis q der Risiken den zu erwartenden Ertrag mindert.

Nach einer Idee von Stoughton und Zechner [23] 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 (5) anzupassen.

Einen weiteren Aspekt erhält man 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 Ertrag wird also nicht maximal.

Schließlich liefert das Resultat einen Ansatz zur monetären Bewertung von Information. Denn mit Gleichung (5) sieht man, dass der maximale zu erwartende Ertrag $r_* = [\log \widehat{s} - H(p)]$ lautet. Ändert sich nun die Risikoverteilung, beispielsweise durch Markt- oder Portfolioveränderungen, so ist $p \neq q$, und der Ertrag r ist durch (5) gegeben. Bei einer Änderung des Risikos entspricht die Informationsdifferenz damit

$$r_* - r = K(p;q) \ln 2/t. \tag{7}$$

Anders ausgedrückt lautet diese Gleichung:

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

Hierbei bemisst sich der Renditeverlust aus der Differenz der optimalen zur tatsächlich erwarteten Rendite, und die Informationsdifferenz der beiden Entropien ist durch die sogenannte Kullback-Leibler-Information gegeben, vgl. Gleichung (7). Information kostet also, nämlich pro Zeiteinheit den Preis der relativen Ertragsminderung.

Danksagung. Ich danke Jean-Christophe Curtillet für wertvolle Hinweise und inspirierende Diskussionen.

Literatur

[1] K.J. Arrow: ‘Le rle des valeurs boursieves pour la repartition la meilleure des risques’, *Economtrie*, 41–48, CNRS (1953)

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

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

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

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

[6] L. Gulko: ‘The Entropic Market Hypothesis’, *International Journal of Theoretical and Applied Finance* **2**(3), 293-329 (1999)

[7] L. Gulko: ‘The entropy theory of stock option pricing’, *International Journal of Theoretical and Applied Finance* **2**(3), 331-355 (1999)

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

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

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

[11] E.T. Jaynes: ‘Information theory and statistical mechanics’. *Phys. Rev.* **106**, 620-630 (1957)

[12] 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

[13] E.T. Jaynes, Edwin T. (1978), ’Where do we stand on maximum entropy?’ In: R.D. Levine & M. Tribus (Eds.), *The Maximum Entropy Formalism*, MIT Press, S. 15 – 118 (1978)

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

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

[16] H. Lyre: *Quantentheorie der Information. Zur Naturphilosophie der Theorie der Ur-Alternativen und einer abstrakten Theorie der Information.* Springer-Verlag, Wien (1998)

[17] P. Mertens (Hrsg.): *Lexikon der Wirtschaftsinformatik.* Springer-Verlag Berlin Heidelberg 1997 (S. 195)

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

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

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

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

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

[23] 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

[24] 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)

[25] A. de Vries: 'Die Moment-Ertrag-Relation und optimale Portfolioauswahl'. http://haegar.fh-swf.de/homepage/publikationen/MERoP.pdf

[26] Carl Friedrich von Weizsäcker: *Aufbau der Physik.* Carl Hanser Verlag, München 1985 (insbes. S. 164ff)

[27] Carl Friedrich von Weizsäcker: *Zeit und Wissen.* Carl Hanser Verlag, München 1992 (S. 154ff)

[28] 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`