Abschätzung von Kursrisiken durch Bootstrapping

von

Dr. rer. nat. Hans Uhlig

Statistische Verfahren zur Abschätzung von Kurszielen und Kursrisiken

Statistische Auswertung der Zeitreihen von Kursen für Wertpapiere, Rohstoffe und Devisen ist kein Selbstzweck, sondern sie soll helfen, künftige Kursverläufe besser abschätzen zu können. Ein Problem mit dem man bei der Analyse dieser Zeitreihen oft zu tun hat, ist der Mangel an Daten. Statistisch betrachtet ist jede Zeitreihe nur eine Stichprobe aus allen möglichen Datenpunkten. Wir hoffen, dass diese Stichprobe repräsentativ für alle möglichen Datenpunkte sei, aber verlassen können wir uns darauf nicht. Bei wenigen Daten ist die Unsicherheit durch Schätzfehler groß. Möchte man den Schätzfehler halbieren, dann muss man die Datenzahl quadrieren. Weil es aber keine beliebig langen Zeitreihen von Kursen für alle möglichen Kapitalanlagen gibt, versucht man, den Datenmangel auf andere Weise zu beheben. Eine davon ist das sogenannte Monte Carlo Verfahren, eine andere, weniger aufwendige Methode, ist das Bootstrapping. Beide Ansätze machen intensiven Gebrauch von Computern und deren Möglichkeit, (Pseudo)-Zufallszahlen mit beliebiger Verteilung zu produzieren.

Monte Carlo Methoden

Wenn man in einer bestimmten Marktsituation wissen möchte, wie sich die Kurse wohl weiterentwickeln werden, dann hat man nicht viele Anhaltspunkte. Eine Möglichkeit besteht darin, ausgehend von der empirisch ermittelten Verteilung der Daten synthetisch zusätzliche Daten mit gleicher Verteilung zu erzeugen und so abzuschätzen, welcher Kurs am wahrscheinlichsten ist und wie wahrscheinlich mögliche Abweichungen von diesem Kurs sind. Es ist möglich, beliebig viele und hohe Kurssprünge mit einzuarbeiten, indem man Poisson-verteilte und log-normal-verteilte Zufallsdaten mischt.

Monte Carlo Methoden wurden zuerst von P.P.Boyle (1977) angewandt, damals mit dem Ziel, Optionspreise zu ermitteln. Mittlerweile ist diese Methode weit verbreitet. Sie findet Anwendung beim Bewerten von Derivaten, aber auch bei Risikoberechnungen, z.B. für das bekannte ‚value at risk'-Kriterium.

Bootstrapping – Monte Carlo Methode des kleinen Mannes

Eine Methode, die keine künstlichen Daten verwendet und dennoch die Anzahl verfügbarer Daten für die Analyse beinahe beliebig erhöht, ist das sogenannte Bootstrapping von B. Efron. Der Name Bootstrapping bezieht sich auf eine Geschichte des Barons Münchhausen, in der er berichtet, wie er sich aus dem Sumpf befreite, indem er sich an den eigenen Schnürsenkeln hinauszog. Die Methode besteht darin, aus einer Datenmenge wiederholt Stichproben zu nehmen und zwar beliebig oft. Dabei werden die Proben immer zufällig aus der gesamten Datenmenge entnommen. So erhält man eine viel größere Datenmenge. Es konnte gezeigt werden, dass diese neue große Datenmenge die zugrunde liegende statistische Datenverteilung gut wiedergibt. Damit ist nicht die Verteilung der ursprünglichen Daten gemeint, sondern die Verteilung, die dieser zugrunde liegt. Denn die Originaldaten, so viele es auch sein mögen, sind, statistisch betrachtet, wiederum nur eine Stichprobe der wahren Verteilung. Zwei wichtige Charakteristika der Daten, nämlich die Korrelation untereinander und die Volatilität, d.h. die Schwankungsbreite der Werte in einem bestimmten Zeitraum, die man mit Monte Carlo Methoden nachbilden kann, gehen beim Bootstrapping verloren. Darum nennt man das Bootstrapping auch die Monte Carlo Methode des kleinen Mannes.

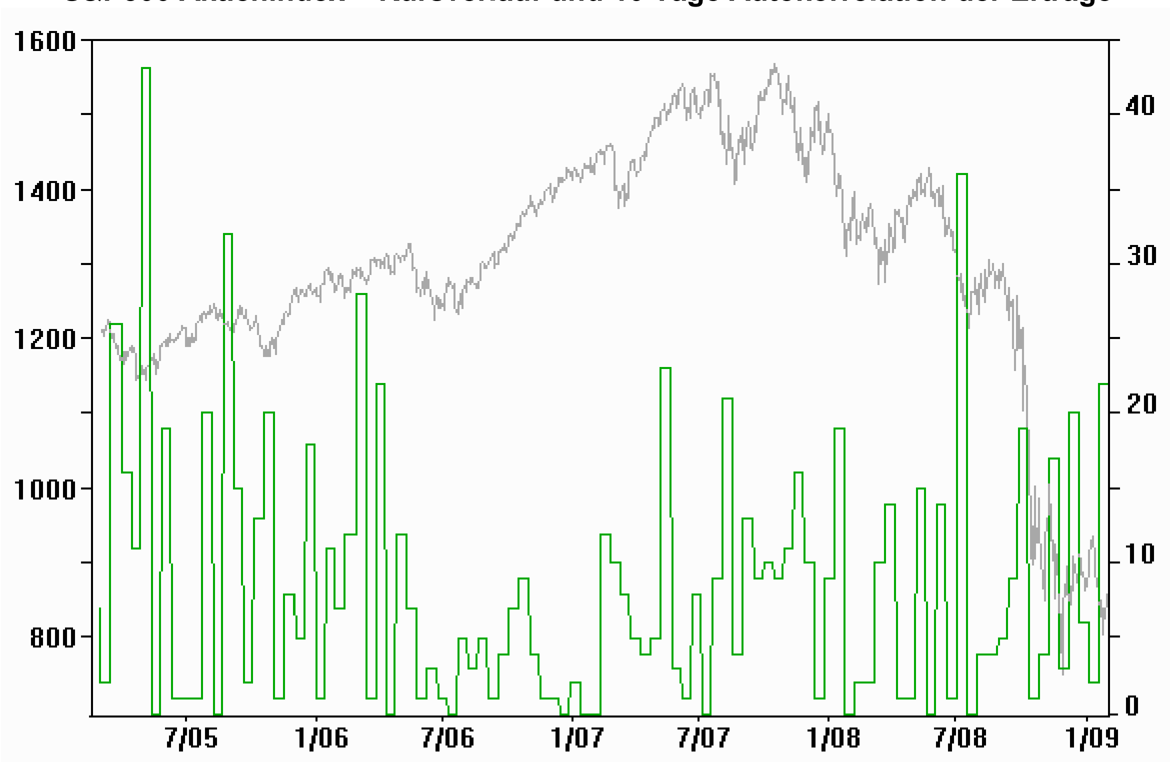
Führt man das Bootstrapping bei, sagen wir, 1000 Datensätzen 1000 Mal auf die soeben beschriebene Weise durch, dann wird im Durchschnitt jeder Datensatz einmal verwendet. Da die Auswahl aber zufällig ist, werden die einzelnen Datensätze der Poisson-Verteilung entsprechend ausgegeben. Wem das nichts sagt, dem hilft vielleicht das folgende Beispiel: Stellen sie sich vor, man würde ein Rosinenbrot in Kastenform backen. Aus dem man nachher zwanzig Scheiben schneiden möchte. Verwendet man für den Teig genau zwanzig Rosinen und knetet den Teig gut durch, dann erwartet man im Durchschnitt eine Rosine pro

Scheibe. Die Poisson-Verteilung sagt uns, welche tatsächliche Verteilung am wahrscheinlichsten wäre, wenn die Rosinen wirklich rein zufällig verteilt wären. Rechnet man angeschnittene Rosinen nur jeweils einer Scheibe zu, dann ist zu erwarten, dass etwa 37% der Scheiben keine Rosine enthalten, wiederum 37% würden eine Rosine enthalten, 18% zwei und 6% drei, der Rest vier oder mehr. Auch wenn man aus den 1000 Datensätzen 2000 zufällige Stichproben nähme, entspräche das im Durchschnitt zwei Zugriffen auf jeden Datensatz. Tatsächlich würde man etwa 13% der Datensätze gar nicht auswählen, etwa 27% würde man einmal auswählen, genau so viele zwei Mal. 18% würde man drei Mal und 9% vier Mal wählen, 3,6% fünf Mal und den Rest noch häufiger. Man kann die Poisson-Verteilung für beliebige Durchschnittswerte mit einem Kalkulationsprogramm berechnen. Die Berechnungsformel dafür ist in jedem Kalkulationsprogramm bereits vorprogrammiert enthalten.

Durch Anwendung des Bootstrapping auf die vorhandenen Marktdaten ließe sich die Aussagekraft der Statistik verbessern. Auch von der Theorie her gibt es keine prinzipiellen Einwände gegen den Einsatz dieses Verfahrens. Es ist nur auf gleichförmig verteilte Daten anwendbar – das ist hier der Fall.

So weit so gut - nur hat die Sache einen Haken. Statistische Daten beziehen sich auf zufällige, ungeordnete Ereignisse. Die statistischen Verfahren, auch das Bootstrapping, setzen eigentlich voraus, dass die Daten, auf die sie angewandt werden, ungeordnet sind. Sie ignorieren jede Ordnung und verschenken damit in den Datenreihen eventuell vorhandene Informationen. Jede Abhängigkeit der Daten untereinander geht beim Bootstrapping verloren.

S&P500 Aktienindex – Kursverlauf und 10 Tage Autokorrelation der Erträge



Tageskurs = grau, linke Skala; Autokorrelation = grün, rechte Skala (%).

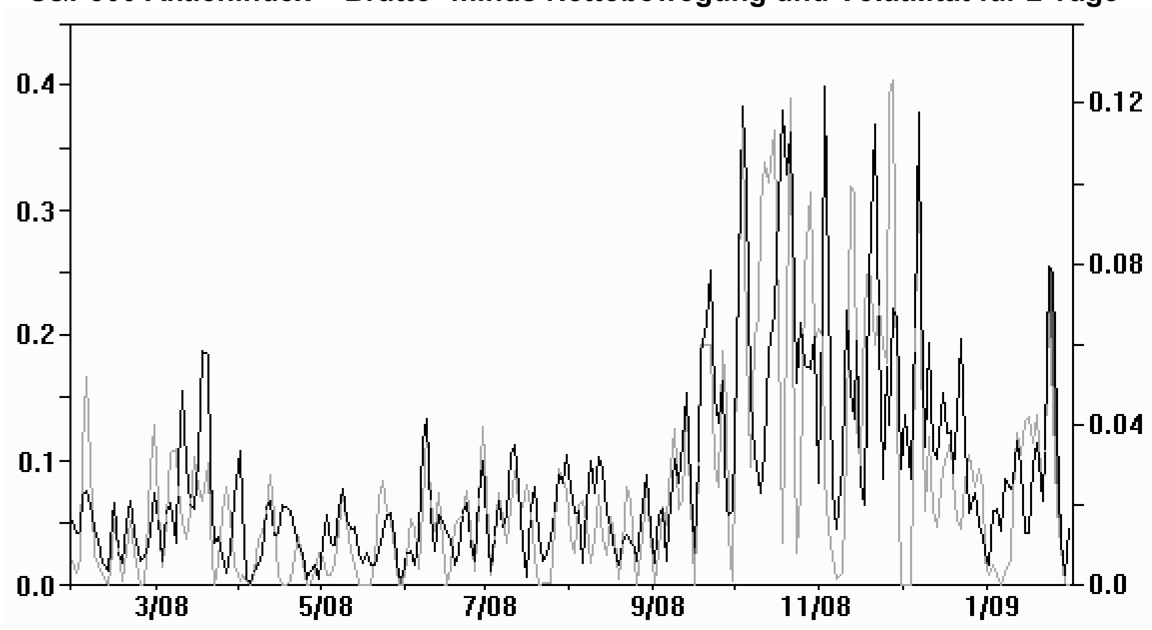
Bei den aufwendigeren Monte Carlo Verfahren hat man dagegen die Möglichkeit, die Autokorrelation der Originaldaten in den Zufallsdaten wiederzugeben, indem man eine sogenannte Cholesky-Zerlegung der Korrelationsmatrix durchführt. Das wäre schon eine große Hilfe für die Risikoabschätzung, doch leider kann die tatsächliche, lokale

Autokorelation von der globalen, der mittleren Autokorrelation stark abweichen, wie man auf der Grafik der vorigen Seite sehen kann. Die Autokorrelationen sind in Zeiten steigender Kurse oft auch geringer als in Zeiten fallender Kurse.

Marktmodelle und Risikoberechnung

Wohin die irrtümlichen Annahmen über scheinbar zufällige Ereignisse führen können, lässt sich an den Banken beobachten, die während der Finanzkrise in Herbst und Winter 2008/2009 in größte Zahlungsschwierigkeiten gerieten. Deren Risikoberechnungen beruhten auf fehlerhaften statistischen Marktmodellen. Die Erkenntnisse der Chaosforschung, obwohl seit 25 Jahren bereits bekannt, hatten die angeblichen Fachleute sträflich vernachlässigt. Wechselseitige Abhängigkeiten unter den Marktpapieren und zeitlich verzögerte Auswirkungen wurden nicht erkannt und/oder berücksichtigt. Es wurden globale Durchschnittswerte für Risikoberechnungen verwendet, anstatt, abhängig von der jeweiligen Marktphase, unterschiedliche Risiken angesetzt. Dadurch wurden die Risiken für aufeinander folgende Verluste während kritischer Marktphasen in ihrer Häufigkeit wie auch in ihrer Höhe, völlig falsch beurteilt, d.h. extrem stark unterschätzt; denn die Zufallswahrscheinlichkeit für eine solche Ereignisfolge war praktisch unendlich klein. So waren auch die für Krisensituationen vorgehaltenen Sicherheiten viel zu gering, denn eine Krise solchen Ausmaßes galt als praktisch unmöglich.

S&P500 Aktienindex – Brutto- minus Nettobewegung und Volatilität für 2 Tage



Brutto- minus Nettobewegung = grau, linke Skala; Volatilität = schwarz, rechte Skala

In der Grafik oben sieht man den Verlauf von Marktschwankungen beim amerikanischen Aktienindex S&P500 vor der Lehman Bros. Pleite, im September 2008 und kurz danach. Man sieht, dass die Stärke der Schwankungen sich vervielfacht und für längere Zeit auf erhöhtem Niveau bleibt. Wie eine Folge von Zufallsereignissen sieht dieser Verlauf nicht aus.

Die Risikoberechnung und damit auch das Risikomanagement beruhten auf dem Verfahren 'Value at risk' (VAR), das gerade zum Branchenstandard erhoben worden war. Dieses Verfahren, nach dem die Wahrscheinlichkeit von Verlusten in Häufigkeit und Höhe berechnet wird, war gegenüber den bis dahin üblichen Berechnungen der Volatilität allein schon ein Fortschritt. Das VAR berücksichtigt endlich die Tatsache, dass nicht nur Einzelverluste entscheidend sind, sondern dass sogar kleine Einzelverluste, wenn sie in kurzen Abständen kommen, zu großen, ja ruinösen Verlusten werden können. Leider ist dieser Ansatz beim VAR nur halbherzig, denn die Berechnungen setzen voraus, dass Marktbewegungen rein zufällig sind. Man hätte es vorher schon besser wissen sollen, denn

Anhaltspunkte dafür gab es viele, doch in der Schönwetterperiode an der Börse lief alles auch so reibungslos. Es gab anscheinend nicht genug Anlass, die Risikoberechnungen zu überarbeiten. So war lange bekannt, dass die Volatilitäten, d.h. die Marktschwankungen nicht zufällig verteilt sind, sondern in unregelmäßigen Abständen Phasen mit starken Schwankungen und solche mit geringen Schwankungen sich abwechseln. Das heißt nicht, die Phasenwechsel kommen jeweils völlig unerwartet. Wie man sie frühzeitig erkennen kann, wird im nächsten Abschnitt besprochen.

Marktschwankungen und die Bedeutung der 200 Tage Durchschnittslinie

Die viel beachtete 200 Tage Durchschnittslinie oder auch 40 Wochen Linie kann in vielen Fällen bei Aktienmärkten die Aufwärtsphase und deren Ende recht gut anzeigen. Die Aufwärtsphase ist dort zuende, wo die aktuelle Kurslinie die 200 Tage Linie von oben nach unten durchbricht. Es gibt verschiedene Möglichkeiten, diese zu berechnen und folglich sind die dabei gefundenen Schnittpunkte nicht immer gleich. Man kann normale gleitende Durchschnitte der Tagesschlusskurse verwenden oder man kann exponentielle gleitende Durchschnitte nehmen. Man kann normale gleitende Durchschnitte von den Logarithmen der Kurse benutzen und man kann statt der Tageswerte auch die Schlusskurse der letzten 40 Wochen verwenden. Ob es sich bei der Aussagekraft dieser Linie um eine sogenannte ‚self fulfilling prophecy' handelt, also um eine Behauptung, die sich selbst erfüllt, kann nicht ausgeschlossen werden, aber das muss uns zunächst nicht kümmern.

Sieht man sich die Marktschwankungen vor und nach dem Durchbruch der 200 Tage Linie an, dann fällt auf, dass die Marktschwankungen vorher deutlich geringer sind als nacher. Es scheint also, dass die 200 Tage Linie recht zuverlässig Marktphasen mit geringen und Marktphasen mit starken Ausschlägen voneinander trennt. Man hat also eigentlich nicht eine Häufigkeitsverteilung, sondern zwei verschiedene.

Marktschwankungen in gößeren Zeiträumen – 1999 bis 2004 im Vergleich mit 2004 bis 2009

Die Frage ist, ob diese nur für die letzten fünf Jahre so war und eigentlich nicht erwartet werden konnte, oder ab dies ein schon länger bekanntes Phänomen ist. Um dies zu prüfen, habe ich zwei aufeinander folgende Fünfjahresabschnitte für verschiedene Aktienmärkte daraufhin untersucht, ob diese Trennung in Marktphasen mit geringen Schwankungen einerseits und solche mit großen Schwankungen andererseits, für beide Zeitabschnitte nachweisbar ist. Dazu habe ich die Tageserträge von Ende Januar 1999 bis Ende Januar 2004 als ersten Zeitabschnitt und von Ende Januar 2004 bis Ende Januar 2009 als zweiten Zeitabschnitt gewählt. Die Märkte waren: der amerikanische Leitindex S&P500, der deutsche Leitindex DAX, der schweizerische Leitindex SMI und der österreichische ATX.

Allgemeine Befunde:
Die Häufigkeitsverteilungen des zweiten Zeitabschnitts streuten deutlich stärker, als die im ersten Zeitabschnitt. Am geringsten war der Unterschied beim SMI, gefolgt von DAX und S&P500. Sehr stark waren die Unterschiede beim ATX. Wenn man sich die Streuung der Ertäge für die Zeiten ansieht, in denen der Index oberhalb der 200 Tage Line lag, dann war die Streuung für alle Indizes bis auf den ATX in den letzten fünf Jahren geringer als in den fünf Jahren davor. Beim ATX ist die Streuung im späteren Zeitabschnitt geringfügig größer. Die Zeiten, in denen der jeweilige Index oberhalb, bzw. unterhalb der 200 Tage Linie verweilte, war im früheren Zeitabschnitt etwa gleich lang, im zweiten Zeitabschnitt war die Verweildauer oberhalb der 200 Tage Linie drei bis vier Mal so lang, wie die Zeit unterhalb dieser Linie. Da aber die starken Schwankungen alle in dieser kurzen Phase auftraten, in der sich der Markt unterhalb der 200 Tage Linie befand, waren diese Ereignisse im letzten Jahr vier bis fünf Mal so häufig, wie nach der Gesamtverteilung zu erwarten wäre.

Nun muss man fairerweise sagen, dass diese Häufigkeitsverteilung ja erst jetzt empirisch festgestellt wurde und daher vorher nicht bekannt sein konnte. Aber in den fünf Jahren davor waren die starken Schwankungen ebenfalls nur aufgetreten, als der Indexkurs unterhalb der 200 Tage Linie lag. Die Schwankungen waren in dieser Zeit also etwa doppelt so häufig, wie nach der empirischen Verteilung für den Gesamtzeitraum angenommen.

Die Tageschwankungen allein lassen keine Aussage darüber zu, wie hoch das Risiko des Ruins tatsächlich ist. Maßgeblich dafür ist, in welcher Folge die jeweiligen negativen und positiven Ereignisse eintreten. Wechseln sich negative und positive Ereignisse immer ab, kann es richtig tiefe Kursstürze nicht geben. Treten aber zuerst überwiegend negative Ereignisse ein und danach erst überwiegend positive, dann ist die Gefahr des Ruins deutlich größer. Wenn die Abfolge zufällig ist, dann kommen lange Serien selten vor, aber wiederholte Wechsel sind auch selten.

Welche Verteilung der Verluste ist innerhalb eines Zeitraumes von zwei Tagen zu erwarten. Ausgehend von den tatsächlichen Tageserträgen habe ich 10 Millionen Mal nacheinander zwei Werte jeweils ‚zufällig' aus den Tageserträgen von Januar 2004 bis Januar 2009 vom S&P500 gewählt und addiert. Dann habe ich überprüft, wie häufig welche Verluste auftraten. Diese, nach der Theorie zu erwartenden Daten habe ich mit den tatsächlich beobachteten Zweitagessummen verglichen. Dabei zeigte sich, dass die Werte um so stärker von der erwarteten Verteilung abwichen, je extremer sie waren, während für moderate Verluste die Abweichungen nur um den Faktor 1,5 bis 3,0 differierten, war die Abweichung in den Extrembereichen Faktor 10 bis Faktor 100. Bei drei Tagen sind die Unterschiede in den Extrembereichen Faktor 5 bis 50. bei vier Tagen sind es Faktoren von 7 bis 100.

Da ich die Originalerträge verwendet habe, kann die hier festgestellte Differenz zwischen den erwarteten und den beobachteten Werten nicht dadurch erklärt werden, dass ich eine fehlerhafte Verteilung verwendet hätte. Wie es scheint, ist die Abfolge der Werte nicht zufällig. Das bestätigt die Befunde mit den Surrogatdaten beim Vergleich der bedingten Entropie und bei Chi²-Test auf Unabhängigkeit.

Wenn man nur die Werte verwendet, die unter der 200 Tage Linie erhalten wurden und damit die Datenfolge für zwei Tage berechnet, dann gibt es eine deutlich bessere Anpassung. Nur beim ganz extremen Wert am Rand der Verteilung ist eine Abweichung von Faktor 10 festzustellen.

In meinem Buch zur Finanzmarktanalyse (Uhlig, 1999) habe ich vor nunmehr zehn Jahren bereits nachgewiesen, dass die Dynamik der Märkte nicht rein zufällig ist. Natürlich war ich weder der erste noch der einzige, der zu dieser Erkenntnis kam. Es gab damals bereits einige entsprechende Untersuchungen anderer Autoren und auch die habe ich zitiert, zusammengefasst sowie für den deutschsprachigen Raum allgemein verständlich aufbereitet. Zwar sind die Märkte, wie auch nach der Theorie zu erwarten, komplizierter geworden, doch hat sich an der Grundaussage, dass die Marktdynamik auch einer gewissen Ordnung folgt, nichts geändert.