**Mannheimer Manuskripte zu Risikotheorie,
Portfolio Management und Versicherungswirtschaft**

Nr. 173

Anwendung des Kalman-Filters zur Identifikation und Projektion von Zinsstrukturmodellen : Modelltheoretische Grundlagen

von
PETER ALBRECHT UND CHRISTOPH MAYER

Mannheim 05/2008

Anwendung des Kalman-Filters zur Identifikation und Projektion von Zinsstrukturmodellen : Modelltheoretische Grundlagen

Peter Albrecht / Christoph Mayer

1. Einführung

Die empirische Identifikation von arbitragefreien Modellen der Zinsstruktur[1] beinhaltet eine spezifische Problematik. Die zentrale die Zinsstruktur treibende Größe ist die Zinsintensität $\{R_t\}$. Die Zinsintensität als Grenzwert kurzfristiger Zinssätze ist nun aber keine am Markt beobachtbare Größe, sondern eine latente Variable. Bei den Standardansätzen zur Identifikation von Zinsstrukturmodellen wird daher die Zinsintensität durch einen kurzfristigen Zinssatz, etwa Zinssätzen auf Monatsbasis[2], approximiert. Diese Vorgehensweise führt notwendigerweise zu Verzerrungen bei der statistischen Identifikation der Prozessparameter.

Eine Alternative hierzu bildet der State Space-Ansatz bzw. Kalman-Filter[3], da dieser explizit die Erfassung latenter Variablen mit stochastischer (linearer) Entwicklungsdynamik erlaubt. Die Verwendung des Kalman-Filters auf die Identifikation von Zinsstrukturmodellen wurde erstmals von PENNACCHI (1991) durchgeführt und hat in jüngerer Zeit[4] verstärkt Beachtung gefunden. Der Kalman-Filter weist dabei eine Reihe von weiteren Vorzügen auf, die ihn – gerade bei höherdimensionalen Modellen – zu einem interessanten und wertvollen generellen Ansatz zur Identifikation von Zinsstrukturmodellen machen:

- Durch seine rekursive Struktur ist ein ständiges Update der Prognosen (Projektionen) der Zinsstrukturkurve möglich, sowohl einstufig als auch mehrstufig.
- Die rekursive Struktur des Kalman-Filters erlaubt auch in einfacher und direkter Weise die Bestimmung der Likelihood-Funktion als Basis einer Maximum Likelihood-Schätzung.

Im ersten Teil der vorliegenden Ausarbeitung soll daher sowohl die allgemeine Systematik des Kalman-Filters als auch dessen Anwendung auf Multifaktormodelle der Zinsstruktur dar-

---

1 Zu einem ersten Überblick hierzu vgl. etwa ALBRECHT/MAURER (2005, Abschnitt 9.3 und Anhang 9E).

2 So verwenden etwa Fischer/May/Walther (2003) in ihrer aktuellen Arbeit zur Anpassung eines eindimensionalen Cox/Ingersoll/Ross-Modells die Monatsdurchschnitte des Geldmarktsatzes für Tagesgeld am Frankfurter Bankenplatz.

3 Vgl. hierzu generell HAMILTON (1994, Kapitel 13) sowie HARVEY (1989, Kapitel 3).

4 Vgl. etwa DUAN/SIMONATO (1999), GEYER/PICHLER (1999), DE JONG (2000), BELETSKI (2003), CHEN/SCOTT (2003) sowie FISCHER/MAY/WALTHER (2004).

gestellt werden, wobei die Intention nicht die größte mathematische Allgemeinheit und Kürze ist, sondern der Focus auf einer nachvollziehbaren und verständlichen Ausarbeitung liegt.[5] In einem zweiten Teil soll dann die empirische Anwendung der hier entwickelten modelltheoretischen Grundlagen im Vordergrund stehen.

2. Allgemeine Systematik des Kalman-Filters

2.1 Ausgangspunkt

Ausgangspunkt der Analyse ist ein zeitdiskretes (lineares) stochastisches dynamisches System, dessen Entwicklung durch den $(r,1)$-Zufallsvektor X_t $(t=1,\ldots,T)$, den Zustandvektor, beschrieben wird. Der Zustandsvektor bestimmt den Zustand, den das System zum Zeitpunkt t einnimmt. Das Besondere hierbei ist der Umstand, dass der Zustandsvektor eine latente Variable ist (in unserer Anwendung wird dies im einfachsten Fall die Zinsintensität sein), d.h. nicht beobachtet werden kann. Damit ist auch nicht beobachtbar, in welchem Zustand sich das System zum jetzigen (oder früheren) Zeitpunkt(en) befindet. Allerdings existiert ein beobachtbares stochastisches dynamisches System, dessen Entwicklung durch den $(n,1)$-Zufallsvektor Y_t $(t=1,\ldots,T)$ beschrieben wird, der wiederum in einer bestimmten Weise von X_t abhängig ist und damit (beobachtbare) Informationen über X_t enthält.

Generell interessiert man sich für den Zustand, den das System im nächsten Zeitpunkt (allgemeiner: in einem künftigen Zeitpunkt) annehmen wird und möchte hierfür eine Vorhersage (Forecast) bzw. Projektion durchführen. In unserer Anwendung geht es darüber hinausgehend darum, die Zufallsgesetzmäßigkeit der latenten Variablen $\{X_t\}$ zu identifizieren, da von dieser wiederum die Zufallsgesetzmäßigkeit unserer Beobachtungen (in unserer Anwendung: die Zinsstruktur) abhängt. Unser Hauptinteresse in der Anwendung gilt der Identifikation der Zufallsgesetzmäßigkeit von $\{Y_t\}$ und der Gewinnung einer Projektion von Y_{t+s}, jeweils gegeben eine Menge $\{y_1,\ldots,y_t\}$ von empirischen Beobachtungen. In der weiteren Darstellung lehnen wir uns weitgehend an HAMILTON (1994, Kapitel 13) an.

---

[5] In den Abschnitten 3 bzw. 4 konzentrieren wir uns dabei auf Modelle mit 1 bzw. 2 bzw. 3 Faktoren des Vasicek-Typus bzw. des Cox/Ingersoll/Ross-Typus. In Anhang A wird der allgemeine Fall dargestellt.

2.2 Systemgleichungen

Die Entwicklung der Zustandsvariablen wird für $t \geq 1$ beschrieben durch die folgende *Zustandsgleichung* (*state equation*), das lineare System

(2.1a) $$X_{t+1} = f + FX_t + V_{t+1}.$$

Der Zustandsprozess ist damit ein autoregressiver Prozess 1. Ordnung (AR(1)-Prozess) und insbesondere ein Markovprozess.

Teilweise wird die Zustandsgleichung in der Literatur auch in der alternativen (äquivalenten) Version

(2.1b) $$X_{t+1} = f + FX_t + WV_{t+1}$$

notiert, vgl. etwa BELETSKI (2003, S. 22) oder FISCHER/MAY/WALTHER (2004, S. 373). Die Varianz-/Kovarianzmatrix des Residuums hat dann die Gestalt $Q^* = WQW^T$ und im Weiteren wäre jeweils Q durch Q^* zu ersetzen.

Die Entwicklung der Beobachtungsgröße Y_t wird für $t \geq 1$ beschrieben durch die folgende *Beobachtungsgleichung* (*observation equation, measurement equation*), das lineare System

(2.2) $$Y_t = h + HX_t + \varepsilon_t .$$

Die Abhängigkeit zwischen der beobachtbaren Variable und der Zustandsvariable ist damit zu jedem Zeitpunkt linear, allerdings überlagert durch einen zufälligen Messfehler (Störterm).
Die Störterme $\{\varepsilon_t\}$ und $\{V_t\}$ sind vom White Noise-Typus, d.h. es gilt

(2.3a) $$E(\varepsilon_t) = 0, \quad E\left(\varepsilon_t \varepsilon_\tau^T\right) = \begin{cases} R & t = \tau \\ 0 & \text{sonst} \end{cases}$$

bzw.

(2.3b) $$E(V_t) = 0, \quad E\left(V_t V_\tau^T\right) = \begin{cases} Q & t = \tau \\ 0 & \text{sonst} \end{cases}.$$

Die Messfehler treten somit im Zeitablauf unabhängig auf und besitzen einen Erwartungswertvektor null sowie eine Varianz-/Kovarianzmatrix R bzw. Q. Auch untereinander sind die Störterme unkorreliert, d.h.

(2.3c) $$E\left(\varepsilon_t V_\tau^T\right) = 0 \quad \text{für alle } t, \tau .$$

Der (nicht-beobachtbare) Startwert des Systems ist X_1. Wir nehmen diesbezüglich des Weiteren an, dass die Störterme unkorreliert mit X_1 sind, d.h.

(2.4a) $$E\left(\varepsilon_t X_1^T\right) = 0 \quad \text{für alle t}$$

(2.4b) $$E\left(V_t X_1^T\right) = 0 \quad \text{für alle t .}$$

Ziel des Kalman-Filters ist, gegeben die Informationsmenge $I_t = \{Y_t, Y_{t-1}, \ldots, Y_1\}$ zum Zeitpunkt t, die Vorhersage von X_{t+1} im Sinne der besten linearen Projektion (d.h., der Minimierung des mittleren quadratischen Fehlers unter allen linearen Projektionsfunktionen $a_0 + a_1 Y_1 + \ldots + a_t Y_t$). Generell notieren wir im Weiteren die beste lineare Projektion einer Zufallsvariablen Z gegeben I_t durch das Symbol $\hat{Z}_t = \hat{E}\left(Z \mid I_t\right)$. $\hat{Z}_t$ ist die orthogonale Projektion von Z auf den durch $\{1, Y_1, \ldots, Y_t\}$ aufgespannten Unterraum und erfüllt die Normalen-Gleichungen ($Y_0 \equiv 1$)

$$E\left[\left(Z - \hat{Z}_t\right) Y_j^T\right] = 0 \quad \text{für alle } j = 0, 1, \ldots, t .$$

Unter Benutzung dieser Notation ist somit das Ziel des Kalman-Filters die Gewinnung der Projektion $\hat{X}_{t+1|t} = \hat{E}\left(X_{t+1} \mid I_t\right)$. Da die X_t nicht beobachtbar sind, muss hierzu "indirekt" unter Ausnutzung der Systemgleichungen (2.1) und (2.2) vorgegangen werden. Die besondere Eigenschaft des Kalman-Filters ist dabei seine rekursive Struktur, d.h. die sukzessive Generierung von $\hat{X}_{1|0}, \hat{X}_{2|1}, \ldots$, die eine einfache Aktualisierung (Update) der Projektion zulässt, wenn neue Informationen eingegangen sind.

Mit diesen Projektionen ist jeweils eine (r, r)-Mean Square Error-Matrix assoziiert:

(2.5) $$P_{t+1} = E\left[\left(X_{t+1} - \hat{X}_{t+1|t}\right)\left(X_{t+1} - \hat{X}_{t+1|t}\right)^T\right].$$

Wir beschreiben im Weiteren zunächst die einzelnen im Rahmen des Kalman-Filters zur Gewinnung der Projektion $\hat{X}_{t+1|t}$ durchzuführenden Schritte.

2.3 Start der Rekursion

Die Rekursion startet mit $\hat{X}_{1|0}$, intuitiv der Prognose von X_1 auf der Grundlage noch keiner vorhandenen Beobachtung. $\hat{X}_{1|0}$ entspricht damit der unbedingten Erwartung

(2.6a) $$\hat{X}_{1|0} = E(X_1)$$

mit assoziierter MSE-Matrix

(2.6b) $$P_1 = E\left[\left(X_1 - E(X_1)\right)\left(X_1 - E(X_1)\right)^T\right] = \Sigma,$$

die mithin der Varianz-/Kovarianzmatrix $\Sigma = \mathrm{Cov}(X_1)$ von X_1 entspricht.

Geht man von der Standardannahme aus, dass $\{X_t\}$ eine stationäre Folge von Zufallsvektoren ist, so sind die folgenden Gleichungen gültig, vgl. etwa HAMILTON (1994, S. 378), deren Lösung auf $E(X_1)$ bzw. Σ führt:

(2.7a) $$(E - F)E(X_1) = f$$

(2.7b) $$\Sigma = F\Sigma F^T + Q.$$

Dabei ist E die (r, r)-Einheitsmatrix.

In Abschnitt 3.1 werden wir jedoch noch einen alternativen Ansatzpunkt darstellen, um $E(X_1)$ bzw. Σ zu gewinnen.

Nachdem der Startwert der Rekursion somit feststeht, dokumentieren wir im Weiteren die für den Schritt von $t-1$ nach t durchzuführenden Einzelprojektionen.

2.4 Projektion von Y_t

Gegeben ist $\hat{X}_{t|t-1}$ und es gilt in diesem Kontext $\hat{E}(X_t \mid I_{t-1}) = \hat{X}_{t|t-1}$. Gesucht wird nun die beste lineare Projektion von Y_t auf der Basis der Informationsmenge I_{t-1}. Aufgrund von (2.2) gilt nun

(2.8a) $$\hat{Y}_{t|t-1} = \hat{E}(Y_t \mid I_{t-1}) = h + H\hat{X}_{t|t-1}.$$

Die zugehörige MSE-Matrix M_t ist dann gegeben durch

(2.8b) $$M_t := E\left[\left(Y_t - \hat{Y}_{t|t-1}\right)\left(Y_t - \hat{Y}_{t|t-1}\right)^T\right] = HP_tH^T + R\,.$$

2.5 Aktualisierung von X_t

Kommt nun noch die Beobachtung Y_t hinzu, d.h. ist die Informationsmenge geben durch I_t, so können wir den Prognosewert für X_t aktualisieren und notieren diesen aktualisierten Wert durch $\hat{X}_{t|t}$. Es gilt dann:

(2.9a) $$\hat{X}_{t|t} := \hat{E}\left(X_t \mid I_t\right) = \hat{X}_{t|t-1} + P_tH^TM_t^{-1}\left[Y_t - h - H\hat{X}_{t|t-1}\right].$$

$\hat{X}_{t|t}$ wird auch als (auf Basis der Beobachtungen) gefilterter Schätzer für X_t bezeichnet. Für die zugehörige (r,r)-MSE-Matrix N_t gilt entsprechend

(2.9b) $$N_t := E\left[\left(X_t - \hat{X}_{t|t}\right)\left(X_t - \hat{X}_{t|t}\right)^T\right] = P_t - P_tH^TM_t^{-1}HP_t\,.$$

2.6 Projektion von X_{t+1}

Damit sind wir nun in der Lage, das Ziel der Prozedur, die Gewinnung der Projektion von X_{t+1}, gegeben die Beobachtungen $\{Y_1,\ldots,Y_t\}$, d.h. der Informationsmenge I_t, zu erreichen. Generell gilt

(2.10a) $$\hat{X}_{t+1|t} = f + F\hat{X}_{t|t}\,.$$

Hieraus folgt nun aus (2.9a)

(2.10b) $$\hat{X}_{t+1|t} = f + F\hat{X}_{t|t-1} + K_t\left(Y_t - h - H\hat{X}_{t|t-1}\right).$$

Dabei bezeichnet die (r,n)-Matrix K_t die sogenannte *Kalman Gain-Matrix* und es gilt

(2.10c) $$K_t = FP_tH^TM_t^{-1}.$$

Die Kalman Gain-Matrix spezifiziert, welches Gewicht dem aktuellen Vorhersagefehler gegeben werden soll. Für die MSE-Matrix P_{t+1} gilt dann die Rekursionsgleichung

(2.10d) $$P_{t+1} = FN_tF^T + Q = F\left[P_t - P_tH^TM_t^{-1}HP_t\right]F^T + Q\,.$$

Die Projektion von Y_{t+1} gegeben die Beobachtungsmenge $\{Y_1,\ldots,Y_t\}$ lautet entsprechend zu (2.8a)

(2.11) $$\hat{Y}_{t+1|t} = h + H\hat{X}_{t+1|t}\,.$$

2.7 Zusammenfassung der Rekursion

Die Rekursion startet mit den unbedingten ersten beiden Momenten $X_{1|0} = E(X_1)$ und $P_1 = \mathrm{Cov}(X_1)$ von X_1. Die Rekursion für die Projektion der Zustandsvariablen ist dann insgesamt gegeben durch (2.10), d.h.

(2.12a) $$\hat{X}_{t+1|t} = f + F\hat{X}_{t|t-1} + FP_tH^TM_t^{-1}\left(Y_t - h - H\hat{X}_{t|t-1}\right)$$

sowie

(2.12b) $$P_{t+1} = F\left[P_t - P_tH^TM_t^{-1}HP_t\right]F^T + Q\,.$$

2.8 Maximum Likelihood-Schätzung

Die Projektionen $\hat{X}_{t|t-1}$ bzw. $\hat{Y}_{t|t-1}$ sind per Konstruktion die besten linearen Projektionen für X_t bzw. Y_t gegeben die Beobachtungsmenge $I_{t-1} = \{Y_1,\ldots,Y_{t-1}\}$.

Nimmt man darüber hinaus für X_1 sowie die Störterme $\{\varepsilon_t, V_t; t = 1,\ldots,T\}$ eine multivariate Normalverteilung an, so liegt sogar die beste Projektion unter allen messbaren Funktionen von $Y_1,\ldots,Y_{t-1}$ vor. Im betrachteten Fall ist auch Y_t bedingt auf die Informationsmenge I_{t-1} multivariat normalverteilt und es gilt

(2.13a) $$Y_t \mid I_{t-1} \sim N_n\left(\mu_t, \Sigma_t\right)$$

wobei gemäß (2.8) gilt

(2.13b) $$\mu_t = E\left(Y_t \mid I_{t-1}\right) = h + H\hat{X}_{t|t-1}$$

(2.13c) $$\Sigma_t = \mathrm{Cov}\left(Y_t \mid I_{t-1}\right) = M_t = HP_tH^T + R\,.$$

Für den Logarithmus der bedingten Dichte gilt damit ($t = 1,\ldots,T$)

(2.14a) $$\ln f_{Y_t|I_{t-1}}(y_t) = -\frac{1}{2}\left\{n\ln(2\pi) + \ln\left[\det(\Sigma_t)\right] + (y_t - \mu_t)^T \Sigma_t^{-1}(y_t - \mu_t)\right\}.$$

Da $n\ln(2\pi)$ nur ein additiver Niveauterm bzw. $1/2$ ein multiplikativer Niveauterm ist, und beide keinen Einfluss auf die Lage des Optimums besitzen, reduziert sich (2.14a) auf

(2.14b) $$\ln f_{Y_t|I_{t-1}}(y_t) = -\left\{\ln\left[\det(\Sigma_t)\right] + (y_t - \mu_t)^T \Sigma_t^{-1}(y_t - \mu_t)\right\}.$$

Numerisch aufwendig bei der Auswertung von (2.14) ist dabei insbesondere die Invertierung der (n,n)-Matrix $\Sigma_t = M_t$ und die Berechnung ihrer Determinante. Da n in der Regel deutlich größer als r ist, empfiehlt es sich, die folgenden Berechnungsformeln – vgl. LUND (1997, S. 6), HARVEY (1989, S. 108) – zu verwenden, um die Auswertungseffizienz zu erhöhen:

(2.15a) $$\Sigma_t^{-1} = R^{-1} - R^{-1}H\left[P_t^{-1} + H^T R^{-1} H\right]^{-1} H^T R^{-1}$$

(2.15b) $$\det(\Sigma_t) = \det(R)\cdot\det(P_t)\cdot\det\left(P_t^{-1} + H^T R^{-1} H\right).$$

Bei Vorliegen einer Menge $\{y_1,\ldots,y_T\}$ von Beobachtungen ergibt sich aufgrund der generellen Faktorisierung $f_{Y_1,\ldots,Y_T} = f_{Y_1}\cdot f_{Y_2|Y_1}\cdot\ldots\cdot f_{Y_T|Y_{T-1},\ldots,Y_1}$ die Log-Likelihood-Funktion zu

(2.16) $$\begin{aligned} L_\theta(y_1,\ldots,y_T) &= \sum_{t=1}^{T} \ln f_{Y_t|I_{t-1}}(y_t) \\ &= -\sum_{t=1}^{T}\left\{\ln\left[\det(\Sigma_t)\right] + (y_t - \mu_t)^T \Sigma_t^{-1}(y_t - \mu_t)\right\}, \end{aligned}$$

wobei θ den Vektor der offenen Parameter (in den Systemmatrizen) bezeichne. Zur Schätzung von θ ist $L_\theta(y_1,\ldots,y_T)$ numerisch zu optimieren. Damit werden Algorithmen zur numerischen Optimierung relevant für die Maximum Likelihood-Schätzung. In der Literatur wird ferner ausgeführt, dass hierfür insbesondere der EM-Algorithmus geeignet ist.

2.9 Mehrstufige Projektion

Neben der Gewinnung einer Projektion $\hat{Y}_{t+1|t}$ von Y_{t+1} gegeben die Informationsmenge $\{Y_1,\ldots,Y_t\}$ sind wir darüber hinausgehend interessiert an der Gewinnung einer Projektion $Y_{t+s|t}$. Dies gelingt wie folgt, zur Vorgehensweise vgl. HAMILTON (1994, S. 384 f.).

Aufgrund von (2.1a) erhalten wir durch Auflösung der Rekursion den Zusammenhang

$$X_{t+s} = \left(E + F + F^2 + \ldots + F^{s-1}\right) f + F^s X_t + F^{s-1} V_{t+1} + F^{s-2} V_{t+2} + \ldots + F V_{t+s-1} + V_{t+s} .$$

Hieraus ergibt sich

(2.17)
$$\begin{aligned} \hat{X}_{t+s|t} &= \hat{E}\left(X_{t+s} \mid I_t\right) \\ &= \left(E + F + \ldots + F^{s-1}\right) f + F^s \hat{X}_{t|t} . \end{aligned}$$

Dies ist die Verallgemeinerung der Beziehung (2.10a).

Die entsprechende Projektion für Y_{t+s} lautet dann

(2.18)
$$\hat{Y}_{t+s|t} = h + H\hat{X}_{t+s|t} .$$

Der Kalman-Filter ermöglicht damit eine Alternative zur üblichen Vorgehensweise, zunächst die aktuelle Zinsstruktur (möglichst perfekt[6]) zu identifizieren und dann die weitere Prozessentwicklung auf der Basis einer Monte Carlo-Simulation zu generieren[7].

3. Zinsstrukturmodelle des Vasicek-Typus

3.1 Vorüberlegungen zur Gewinnung der Zustandsgleichung

Unter Ausnutzung der Eigenschaften der bedingten Erwartung lässt sich die folgende generelle Darstellung von X_{t+1} nachweisen, man vgl. etwa DUAN/SIMONATO (1999),

(3.1a)
$$X_{t+1} = E\left(X_{t+1} \mid X_t\right) + V_{t+1}^* ,$$

wobei

(3.1b)
$$E\left(V_{t+1}^*\right) = 0 , \quad \mathrm{Cov}\left(V_{t+1}^*\right) = \mathrm{Cov}\left(X_{t+1} \mid X_t\right) .$$

Diese generelle Darstellung gilt zunächst für eine beliebige Folge $\{X_t\}$ von Zufallsgrößen. In äquivalenter Form lautet sie

(3.2)
$$X_{t+1} = E\left(X_{t+1} \mid X_t\right) + \mathrm{Cov}\left(X_{t+1} \mid X_t\right)^{1/2} Z_{t+1} ,$$

wobei $Z_{t+1} \sim N(0, E)$ ein multivariat standardnomalverteilter Zufallsvektor ist und $\mathrm{Cov}\left(X_{t+1} \mid X_t\right)^{1/2}$ die Cholesky-Zerlegung der positiv definiten Matrix $\mathrm{Cov}\left(X_{t+1} \mid X_t\right)$.

---

6 Vgl. hierzu etwa ALBRECHT/MAURER (2005, S. 473 f.).

7 Die im Wege der Simulation ermittelte weitere Prozessentwicklung besitzt zudem nicht die Optimalitätseigenschaften einer (linearen) Prozessprojektion.

Unter Zugrundelegung der Zustandsgleichung (2.1a) gilt ferner $E(X_{t+1} \mid X_t) = E(f + X_t + V_{t+1} \mid X_t) = E(f + X_t \mid X_t) + E(V_{t+1} \mid X_t) = f + FX_t$. Dabei nutzen wir die generelle Eigenschaft $E[h(X) \mid X] = h(X)$ der bedingten Erwartung, und, dass V_{t+1} unkorreliert ist mit X_t, vgl. zu Letzterem etwa HAMILTON (1994, S. 373).

Wir halten somit fest, dass bei Gültigkeit von (2.1a)

(3.3) $$E(X_{t+1} \mid X_t) = f + FX_t$$

gilt. Für die Residualvariable V_{t+1}^* gilt bereits $E(V_{t+1}^*) = 0$, damit liegt gemäß (3.1b) insgesamt der White Noise-Fall vor, wenn $Cov(X_{t+1} \mid X_t)$ eine deterministische Matrix ist. In diesem Falle sind V_{t+1}^* gemäß (3.1a) und V_{t+1} gemäß (2.1a) offenbar identisch, d.h. (3.1a) ist eine äquivalente Darstellung der Zustandsgleichung (2.1a).

Allerdings ist $Cov(X_{t+1} \mid X_t)$ nicht generell eine deterministische Matrix, dies gilt nur in Spezialfällen. Ein zentraler Spezialfall ist der Gauß-Fall, d.h. X_1 und die Residuen V_t^* (und damit alle X_t) sind jeweils multivariat normalverteilt. Im Gauß-Fall ist somit (3.1a) eine äquivalente Darstellung von (2.1a) und (3.1a) kann als Ansatzpunkt genommen werden, um die Zustandsgleichung zu spezifizieren.

Die in diesem Abschnitt betrachteten Zinsstrukturmodelle vom Vasicek-Typus sind jeweils unter den Gauß-Fall subsumierbar, so dass auf dieser Grundlage die Zustandsgleichung in einfacher Form abgeleitet werden kann. Die im nächsten Abschnitt betrachteten Zinsstrukturmodelle vom Cox/Ingersoll/Ross-Typus sind hingegen nicht mehr unter den Gauß-Fall subsumierbar und in der Tat ist dann $Cov(X_{t+1} \mid X_t)$ stochastisch, so dass eine alternative Vorgehensweise gefunden werden muss, auf die wir zurückkommen werden.

Im Gauß-Fall eröffnet sich darüber hinaus ein einfacher Weg zur Identifikation der Startwerte $E(X_1)$ bzw. $Cov(X_1)$. Aus (3.1a) folgt mit (3.3) zunächst $E(X_{t+1}) = E[E(X_{t+1} \mid X_t)]$ $= E(f + FX_t) = f + FE(X_t)$. Unter der Annahme der Stationarität folgt hieraus insgesamt

(3.4a) $$(E - F)E(X_t) = f \ .$$

Dies entspricht der Gleichung (2.7a). Weiter gilt $\mathrm{Cov}(X_{t+1}) = E\left[\mathrm{Cov}(X_{t+1} \mid X_t)\right] + \mathrm{Cov}\left[E(X_{t+1} \mid X_t)\right] = E\left[\mathrm{Cov}(X_{t+1} \mid X_t)\right] + F\mathrm{Cov}(X_t)$. Aus der Annahme der Stationarität folgt hieraus die Gleichung

(3.4b) $$(E - F)\mathrm{Cov}(X_t) = E\left[\mathrm{Cov}(X_{t+1} \mid X_t)\right].$$

Die Gleichungen (3.4a) und (3.4b) werden in Anhang B benutzt, um die Startwerte der Kalman-Rekursion zu bestimmen.

3.2 Einfaktormodell von Vasicek

Das Einfaktormodell von Vasicek beruht[8] auf der Annahme eines Ornstein/Uhlenbeck-Prozesses für die Entwicklung der Zinsintensität $\{R_t\}$, üblicherweise geschrieben als stochastische Differentialgleichung der Form

(3.5) $$dR_t = \alpha(\mu - R_t)dt + \sigma dW_t .$$

Dabei gilt $\mu = \mu_P$, d.h. der Prozess (3.5) wird im Weiteren unter dem physischen (beobachtbaren) Wahrscheinlichkeitsmaß P spezifiziert, nicht unter der risikoneutralen Wahrscheinlichkeitsbelegung. Zur Gewinnung der Zustandsgleichung knüpfen wir an deren äquivalenter allgemeiner Form (3.1) an und benutzen die Ergebnisse in ALBRECHT/MAURER (2005, S. 168) für $E(R_t \mid R_s)$ bzw. $\mathrm{Var}(R_t \mid R_s)$. Wir erhalten hieraus[9]

(3.6a) $$E(R_{t+1} \mid R_t) = \mu\left(1 - e^{-\alpha}\right) + e^{-\alpha}R_t$$

sowie

(3.6b) $$\mathrm{Var}(R_{t+1} \mid R_t) = \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha}\right).$$

Die standardisierte Form (2.1a) der Zustandsgleichung ergibt sich damit im vorliegenden Spezialfall zu

(3.7a) $$R_{t+1} = f + FR_t + V_{t+1},$$

wobei

(3.7b) $$f = \mu\left(1 - e^{-\alpha}\right)$$

(3.7c) $$F = e^{-\alpha}.$$

---

8 Vgl. etwa ALBRECHT/MAURER (2005, S. 167 ff.).
9 Vgl. auch Anhang B.1.

Die Varianz-/Kovarianzmatrix Q aus (2.3b) reduziert sich im eindimensionalen Fall auf $\mathrm{Var}\left(V_{t+1}\right) = \mathrm{Var}\left(R_{t+1} \mid R_t\right)$, d.h.

(3.7d) $$Q = \frac{\sigma^2\left(1-e^{-2\alpha}\right)}{2\alpha}.$$

Die beobachteten Zufallsvektoren Y_t entsprechen im Anwendungsfall von Zinsstrukturmodellen den Spot Rates $R\left(t, t+\tau_i\right)$ zum Zeitpunkt t für ausgewählte Restlaufzeiten $0 < \tau_1 < \ldots < \tau_n$, d.h. den internen Zinsfüßen der entsprechenden Einheitszerobonds. Nach dem Vasicek-Zinsstrukturmodell gilt zunächst allgemein:[10, 11]

(3.8a) $$R\left(t, t+\tau\right) = -\frac{1}{\tau}A\left(\tau\right) - \frac{1}{\tau}B\left(\tau\right)R_t,$$

wobei

(3.8b) $$B\left(\tau\right) = \frac{e^{-\alpha\tau}-1}{\alpha},$$

(3.8c) $$A\left(\tau_i\right) = -q\left[B\left(\tau\right)+\tau\right] - \frac{\sigma^2}{4\alpha}B\left(\tau\right)^2,$$

(3.8d) $$q = \mu + \frac{\lambda\sigma}{\alpha} - \frac{\sigma^2}{2\alpha^2}.$$

Der Term $\mu + \frac{\lambda\sigma}{\alpha}$, wobei λ den "Marktpreis des Risikos" bezeichne, entspricht dabei der Drift des Prozesses (3.2) unter der risikoneutralen Wahrscheinlichkeitsbelegung[12, 13].

In Vektorform gilt somit $Y_t = \left(R_{t1}, \ldots, R_{tn}\right)^T$, wobei $R_{ti} = R\left(t, t+\tau_i\right)$. Da bei einer empirischen Identifikation üblicherweise die Anzahl der Beobachtungen (deutlich) größer als die Anzahl der Parameter ist, wird eine exakte empirische Erfüllung von (3.8a) nicht möglich sein, d.h. die theoretische Zinsstrukturkurve kann die Datenlage nicht vollständig erklären.

---

[10] Vgl. etwa ALBRECHT/MAURER (2005, S. 490).

[11] Die Darstellung der Spot Rates gemäß (3.8d) beruht auf der Spezifikation $B\left(t, t+\tau\right) = \exp\left[A\left(\tau\right)+B\left(\tau\right)R_t\right]$ der Zerobondpreise einer affinen Zinsstruktur. Daneben sind in der Literatur auch die Spezifikationen $B\left(t, t+\tau\right) = \exp\left[A\left(\tau\right)-B\left(\tau\right)R_t\right]$ sowie $B\left(t, t+\tau\right) = \exp\left[-A\left(\tau\right)-B\left(\tau\right)R_t\right]$ zu finden.

[12] Vgl. allgemein in Anhang A.2.

[13] Dabei wurde, im Unterschied etwa zu KWOK (1998, S. 323) der überwiegenden Vorgehensweise in der Literatur gefolgt, vgl. etwa BELETSKI (2003, S. 87), und λ durch $-\lambda$ ersetzt. Dies sichert eine positive Risikoprämie bei positivem λ.

Für die empirische Identifizierung ist es daher vorteilhaft, von einer gestörten (mit Messfehlern beobachteten) Version von (3.8a) auszugehen, nämlich

(3.8d) $$R(t,t+\tau_i) = -\frac{1}{\tau_i}A(\tau_i) - \frac{1}{\tau_i}B(\tau_i)R_t + \varepsilon_{ti}\,.$$

Dies ist die "statistische" Modellversion der Zinsstrukturkurve (3.8a). In Vektor-/Matrixform ergibt sich daher

(3.9a) $$\begin{pmatrix} R(t,t+\tau_1) \\ \vdots \\ R(t,t+\tau_n) \end{pmatrix} = \begin{pmatrix} -A(\tau_1)/\tau_1 \\ \vdots \\ -A(\tau_n)/\tau_n \end{pmatrix} + \begin{pmatrix} -B(\tau_1)/\tau_1 \\ \vdots \\ -B(\tau_n)/\tau_n \end{pmatrix} R_t + \begin{pmatrix} \varepsilon_{t1} \\ \vdots \\ \varepsilon_{tn} \end{pmatrix}$$

bzw. in kompakter Form

(3.9b) $$Y_t = h + HR_t + \varepsilon_t\,,$$

d.h. es gilt $h = \left(-A(\tau_1)/\tau_1,\ldots,-A(\tau_n)/\tau_n\right)^T$, $H = \left(-B(\tau_1)/\tau_1,\ldots,-B(\tau_n)/\tau_n\right)^T$ sowie $\varepsilon_t = \left(\varepsilon_{t1},\ldots,\varepsilon_{tn}\right)^T$. Damit ist auch die Beobachtungsgleichung gemäß Abschnitt 2 in ihrer standardisierten Form festgelegt. Offen ist nur noch die Spezifikation der Varianz-/Kovarianzmatrix der Störterme, d.h. $R = \mathrm{Cov}(\varepsilon)$. In der Literatur zur Anwendung des Kalman-Filters auf die Identifikation von Zinsstrukturmodellen werden hierzu unterschiedliche Spezifikationen vorgeschlagen. Die einfachste Variante ist[14]

(3.10a) $$R = \sigma_\varepsilon^2 E_n\,,$$

wobei E_n die (n,n)-Einheitsmatrix bezeichne. Die Messfehler haben hier eine identische Varianz für alle Laufzeiten und sind untereinander unkorreliert (cross-sectionally uncorrelated). Eine Alternative wäre[15]

(3.10b) $$R = DE_n\,,$$

wobei $D = \mathrm{diag}(d_1,\ldots,d_n)$ eine Diagonalmatrix mit $d_i = \mathrm{Var}(\varepsilon_i)$ ist. Die Messfehler wären dann nach wie vor untereinander unkorreliert.

Ein solcher Ansatz ist allerdings nicht invariant gegenüber linearen Datentransformationen. DE JONG (2000) empfiehlt daher die Spezifikation

(3.10c) $$R = LDL^T\,.$$

---

[14] Vgl. zu einer entsprechenden Spezifikation etwa JEGADEESH/PENNACCHI (1996, S. 432), LUND (1997, S. 18) sowie (in einem Cox/Ingersoll/Ross-Kontext) FISCHER/MAY/WALTHER (2004, S. 378).

[15] Zu einer entsprechenden Vorgehensweise vgl. BABBS/NOWMAN (1999, S. 124), BELETSKI (2003, S. 69) und DUAN/SIMONATO (1999).

Dabei ist $D = \mathrm{diag}\left(d_1, \ldots, d_n\right)$ eine Diagonalmatrix (die d_i entsprechen den Eigenwerten von R) und L ist eine untere Dreiecksmatrix der Form

(3.10d)
$$L = \begin{pmatrix} 1 & & & & \\ l_{21} & 1 & & 0 & \\ l_{31} & l_{32} & 1 & & \\ \vdots & \vdots & & \ddots & \\ l_{n1} & l_{n2} & & l_{nn-1} & 1 \end{pmatrix}.$$

Durch diese Parametrisierung wird insbesondere gewährleistet, dass R positiv definit ist.

In den folgenden Auswertungen beschränken wir uns allerdings zur Reduktion der traktablen Parameteranzahl auf die Variante (3.10a), um die Probleme bei der praktischen Implementierung des Modells zu verringern. Zur Durchführung der Rekursionen des Kalman-Filters gemäß Abschnitt 2 sind nun nur noch die Startwerte $\hat{R}_{1|0} = \hat{X}_{1|0}$ und P_1 festzulegen. Gemäß den Ausführungen in Abschnitt 2.3 gilt $R_{1|0} = E\left(R_1\right)$ und $P_1 = \mathrm{Var}\left(R_1\right)$. Auf der Basis der Überlegungen in Abschnitt 3.1 ergeben sich die Startwerte – man vgl. hierzu Anhang B.1 – zu

(3.11a)
$$\hat{R}_{1|0} = \mu$$

und

(3.11b)
$$P_1 = \frac{\sigma^2}{2\alpha}.$$

Auf dieser Grundlage können die linearen Projektionen $\hat{R}_{t+1|t} = \hat{X}_{t+1|t}$ des Kalman-Filters gemäß (2.12a) sowie die damit assoziierten MSE-Matrizen P_{t+1} gemäß (2.12b) berechnet werden. Der Parametervektor, bezüglich dessen die Likelihood-Funktion (2.16) gemäß Abschnitt 2.8 zu maximieren ist, lautet $\theta = \left(\mu, \alpha, \sigma, \lambda, \sigma_\varepsilon\right)$, es sind damit fünf Parameter zu bestimmen.

Darüber hinaus können, gegeben eine Zinsstrukturgeschichte $\left\{Y_1, \ldots, Y_t\right\}$ eine einstufige Projektion $\hat{Y}_{t+1|t}$ gemäß (2.11) sowie mehrstufige Projektionen $\hat{Y}_{t+s|t}$ gemäß (2.18) der Zinsstrukturkurve vorgenommen werden.

3.3 Zweifaktormodell des Vasicek-Typus

Wir gehen aus von dem folgenden Modellansatz für die Zinsintensität ($r_0 \in IR$):

(3.12a) $$R_t = r_0 + X_1(t) + X_2(t).$$

Die Zinsintensität ergibt sich damit als lineare Funktion der Faktoren $X_1(t)$ und $X_2(t)$, wobei $X_t = (X_1(t), X_2(t))^T$ ein zweidimensionaler Ornstein/Uhlenbeck-Prozess mit der folgenden P-Dynamik ist:

(3.12b) $$dX_t = -\alpha X_t dt + C dW_t.$$

$W_t = (W_1(t), W_2(t))^T$ ist dabei ein zweidimensionaler Standard-Wienerprozess, insbesondere sind $W_1(t)$ und $W_2(t)$ unabhängige Prozesse. Weiter gilt

(3.12c) $$\alpha = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}$$

und

(3.12d) $$C = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sqrt{1-\rho^2}\sigma_2 \end{pmatrix}.$$

Dabei ist C die Cholesky-Zerlegung der Varianz-/Kovarianzmatrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

d.h. $\Sigma = CC^T$. Diese Spezifikation[16] führt zu einer Korrelation der beiden Faktoren $X_1(t)$ und $X_2(t)$. Im Fall $\rho = 0$ führt dies zu unabhängigen Faktoren und damit zu einer direkten Verallgemeinerung der Ergebnisse des Abschnitts 3.1. BELETSKI (2003) kommt aufgrund seiner empirischen Analyse allerdings zum Schluss, dass im Kontext von Zweifaktormodellen des Vasicek-Typus Modelle mit unabhängigen Faktoren den empirischen Daten (Varianz- und Korrelationsstruktur) nicht hinreichend entsprechen. Auch DE JONG (2000, S. 308) konstatiert eine signifikante Korrelation zwischen den Faktoren. Daher ist die allgemeine Faktorstruktur (3.12) vorzuziehen. Die Einzelgleichungen der Faktoren lauten dann:

(3.13a) $$dX_1(t) = -\alpha_1 X_1(t) dt + \sigma_1 dW_1(t)$$

(3.13b) $$dX_2(t) = -\alpha_2 X_2(t) dt + \rho\sigma_2 dW_1(t) + \sqrt{1-\rho^2}\sigma_2 dW_2(t)$$

bzw.

[16] Zur grundsätzlichen Konstruktion (allerdings im Kontext der geometrischen Brownschen Bewegung) vgl. etwa ALBRECHT/MAURER (2005, S. 184 f.).

(3.13c) $$dX_2(t) = -\alpha_2 X_2(t)dt + \sigma_2 dW_1^*(t),$$

wobei $W_1^*(t) = \rho W_1(t) + \sqrt{1-\rho^2} W_2(t)$ wiederum ein Standard-Wienerprozess ist, allerdings nun korreliert mit $W_1(t)$.

Im Unterschied zur Spezifikation (3.2) fällt dabei auf, dass in (3.13) die beiden Prozesse jeweils ein Mean Reversion-Niveau von null besitzen. Dies liegt an der linearen Struktur von (3.12a). Solange r_0 nicht restringiert wird, kann die Summe der Langfristmittel von $X_1(t)$ und $X_2(t)$ auf den Parameter r_0 reduziert werden[17]. Der Parameter r_0 entspricht dann dem Langfristmittel bzw. dem Mean Reversion-Niveau der Zinsintensität R_t.

Die Zustandsgleichung für die Entwicklung des Vektors X_t lässt sich nun auf der Basis von (3.1) ableiten. Es gilt zunächst[18]:

(3.14a) $$\begin{aligned} E(X_{t+1} \mid X_t) &= \begin{pmatrix} e^{-\alpha_1} & 0 \\ 0 & e^{-\alpha_2} \end{pmatrix} X_t \\ &= \begin{pmatrix} e^{-\alpha_1} X_1(t) \\ e^{-\alpha_2} X_2(t) \end{pmatrix} \end{aligned}$$

(3.14b) $$\begin{aligned} Q &:= \mathrm{Cov}(X_{t+1} \mid X_t) \\ &= \begin{pmatrix} \frac{\sigma_1^2}{2\alpha_1}\left(1-e^{-2\alpha_1}\right) & \frac{\rho\sigma_1\sigma_2}{\alpha_1+\alpha_2}\left(1-e^{-(\alpha_1+\alpha_2)}\right) \\ \frac{\rho\sigma_1\sigma_2}{\alpha_1+\alpha_2}\left(1-e^{-(\alpha_1+\alpha_2)}\right) & \frac{\sigma_2^2}{2\alpha_2}\left(1-e^{-2\alpha_2}\right) \end{pmatrix}. \end{aligned}$$

Die Zustandsgleichung lautet somit

(3.15a) $$\begin{pmatrix} X_1(t+1) \\ X_2(t+1) \end{pmatrix} = \begin{pmatrix} e^{-\alpha_1} X_1(t) \\ e^{-\alpha_2} X_2(t) \end{pmatrix} + V_{t+1}$$

mit

(3.15b) $$\mathrm{Cov}(V_{t+1}) = Q.$$

Für den Start der Rekursion werden die unbedingten Momente $E(X_1)$ und $\mathrm{Cov}(X_1)$ benötigt. Aus Anhang B.2 ergibt sich[19],

---

[17] Vgl. hierzu BELETSKI (2003, S. 58).
[18] Man vgl. hierzu BELETSKI (2003, S. 66), DE JONG (2000, S. 312) und beachte die Normierung $\mu = 0$.
[19] Vgl. auch DE JONG (2000, S. 312).

(3.16a) $$X_{1|0} = E(X_1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

(3.16b) $$P_1 = \mathrm{Cov}(X_1) = \begin{pmatrix} \frac{\sigma_1^2}{2\alpha_1} & \frac{\rho\sigma_1\sigma_2}{\alpha_1+\alpha_2} \\ \frac{\rho\sigma_1\sigma_2}{\alpha_1+\alpha_2} & \frac{\sigma_2^2}{2\alpha_2} \end{pmatrix}.$$

Kommen wir damit zur Beobachtungsgleichung. Generell gilt hier

(3.17a) $$\begin{pmatrix} R(t,t+\tau_1) \\ \vdots \\ R(t,t+\tau_n) \end{pmatrix} = \begin{pmatrix} -A(\tau_1)/\tau_1 \\ \vdots \\ -A(\tau_n)/\tau_n \end{pmatrix} + \begin{pmatrix} -B_1(\tau_1)/\tau_1 & -B_2(\tau_1)/\tau_1 \\ \vdots & \vdots \\ -B_1(\tau_n)/\tau_n & -B_2(\tau_n)/\tau_n \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} + \begin{pmatrix} \varepsilon_{t1} \\ \vdots \\ \varepsilon_{tn} \end{pmatrix}$$

bzw. in kompakter Form

(3.17b) $$Y_t = h + HX_t + \varepsilon_t .$$

Weiter gilt[20]:

(3.18) $$B(\tau) = \left(B_1(\tau), B_2(\tau)\right)^T = \left(\frac{e^{-\alpha_1\tau}-1}{\alpha_1}, \frac{e^{-\alpha_2\tau}-1}{\alpha_2}\right)^T$$

(3.19a) $$q_1 = -\frac{\lambda_1\sigma_1}{\alpha_1} - \frac{\sigma_1^2}{2\alpha_1^2}$$

(3.19b) $$q_2 = -\frac{\left(\rho\lambda_1 + \sqrt{1-\rho^2}\lambda_2\right)\sigma_2}{\alpha_2} - \frac{\sigma_2^2}{2\alpha_2^2}$$

(3.19c) $$A_\rho(\tau) = \frac{\rho\sigma_1\sigma_2}{\alpha_1+\alpha_2}\left[\frac{1}{\alpha_1}\left(B_1(\tau)+\tau\right) + \frac{1}{\alpha_2}\left(B_2(\tau)+\tau\right) - B_1(\tau)B_2(\tau)\right]$$

(3.19d) $$A(\tau) = \sum_{i=1}^{2}\left\{-q_i\left[B_i(\tau)+\tau\right] - \frac{\sigma_i^2}{4\alpha_i}B_i(\tau)^2\right\} - r_0\tau + A_\rho(\tau).$$

Bei der Spezifikation von $R = \mathrm{Cov}(\varepsilon)$ gehen wir unverändert von (3.10a) aus. Damit sind die Ausgangsdaten zur Anwendung der Kalman-Filter-Prozedur sämtlich festgelegt. Der Parametervektor bezüglich dessen die Likelihood-Funktion (2.16) gemäß Abschnitt 2.8 zu maximieren ist, lautet hierbei $\theta = (r_0, \alpha_1, \alpha_2, \sigma_1, \sigma_2, \rho, \lambda_1, \lambda_2, \sigma_\varepsilon)$, es sind damit neun Parameter zu bestimmen.

[20] Vgl. BELETSKI 2003, S. 57.

3.4 Dreifaktormodell des Vasicek-Typus

Die Zinsintensität sei gegeben durch

(3.20a) $$R_t = r_0 + X_1(t) + X_2(t) + X_3(t).$$

Dabei ist $X(t) = \left(X_1(t), X_2(t), X_3(t)\right)^T$ ein dreidimensionaler Ornstein/Uhlenbeck-Prozess mit der P-Dynamik

(3.20b) $$dX_t = -\alpha X_t dt + C dW_t .$$

Dabei ist $W(t) = \left(W_1(t), W_2(t), W_3(t)\right)^T$ ein dreidimensionaler Standard-Wienerprozess. Weiter gilt:

(3.20c) $$\alpha = \begin{pmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{pmatrix}$$

sowie

(3.20d) $$C = \begin{pmatrix} \sigma_1 & 0 & 0 \\ \sigma_2\rho_{12} & \sigma_2\sqrt{1-\rho_{12}^2} & 0 \\ \sigma_3\rho_{13} & \sigma_3\dfrac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}} & \sigma_3\dfrac{\sqrt{1-\rho_{12}^2-\rho_{13}^2-\rho_{23}^2+2\rho_{12}\rho_{13}\rho_{23}}}{\sqrt{1-\rho_{12}^2}} \end{pmatrix}.$$

Dabei ist wiederum C eine Cholesky-Zerlegung, hier der Varianz-/Kovarianz-Matrix Σ mit

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 \end{pmatrix}.$$

Wie im Falle $n = 2$ wurde ferner die Normierung $\mu = 0$ gewählt.

Die Zustandsgleichung lässt sich nun wiederum auf der Basis von (3.1) ableiten. In Verallgemeinerung von (3.14a) gilt zunächst:

(3.21a) $$E\left(X_{t+1} \mid X_t\right) = \begin{pmatrix} e^{-\alpha_1} X_1(t) \\ e^{-\alpha_2} X_2(t) \\ e^{-\alpha_3} X_3(t) \end{pmatrix}$$

In Verallgemeinerung von (3.14b) gilt des Weiteren

(3.21b) $$Q := \mathrm{Cov}\left(X_{t+1} \mid X_t\right) = q_{ij}$$

mit ($\rho_{ii} = 1$ für $i = 1,2,3$)

(3.21c) $$q_{ij} = \frac{\rho_{ij}\sigma_i\sigma_j}{\alpha_i + \alpha_j}\left[1 - e^{-\left(\alpha_i + \alpha_j\right)}\right].$$

Die Zustandsgleichung lautet somit

(3.22a) $$\begin{pmatrix} X_1(t+1) \\ X_2(t+1) \\ X_3(t+1) \end{pmatrix} = \begin{pmatrix} e^{-\alpha_1} X_1(t) \\ e^{-\alpha_2} X_2(t) \\ e^{-\alpha_3} X_3(t) \end{pmatrix} + V_{t+1}$$

mit

(3.22b) $$\mathrm{Cov}\left(V_{t+1}\right) = Q\,.$$

Für den Start der Rekursion ergibt sich in Verallgemeinerung von (3.16):

(3.23a) $$X_{1|0} = E\left(X_1\right) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

(3.23b) $$P_1 = \mathrm{Cov}\left(X_1\right) = \mathrm{Cov}\left(X_1\right)_{ij}$$

mit ($\rho_{ii} = 1$ für $i = 1,2,3$)

(3.23c) $$\mathrm{Cov}\left(X_1\right)_{ij} = \frac{\rho_{ij}\sigma_i\sigma_j}{\alpha_i + \alpha_j}\,.$$

Kommen wir damit zur Beobachtungsgleichung. Generell gilt hier

(3.24a) $$\begin{pmatrix} R(t, t+\tau_1) \\ \vdots \\ R(t, t+\tau_n) \end{pmatrix} = \begin{pmatrix} -A(\tau_1)/\tau_1 \\ \vdots \\ -A(\tau_n)/\tau_n \end{pmatrix} + \begin{pmatrix} -B_1(\tau_1)/\tau_1 & -B_2(\tau_1)/\tau_1 & -B_3(\tau_1)/\tau_1 \\ \vdots & \vdots & \vdots \\ -B_1(\tau_n)/\tau_n & -B_2(\tau_n)/\tau_n & -B_3(\tau_n)/\tau_n \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} + \begin{pmatrix} \varepsilon_{t1} \\ \vdots \\ \varepsilon_{tn} \end{pmatrix}$$

bzw. in kompakter Form

(3.24b) $$Y_t = h + HX_t + \varepsilon_t\,.$$

Für die beteiligten Funktionen gilt nun[21] ($i = 1,2,3,\ \rho_{ii} = 1$ für $i = 1,2,3$),

[21] Vgl. hierzu BELETSKI (2003, S. 59; Appendix 1) sowie Anhang A.2.

(3.25a) $$B_i(\tau) = \frac{e^{-\alpha_i \tau} - 1}{\alpha_i}$$

(3.25b) $$A_{ij}(\tau) = \frac{\rho_{ij}\sigma_i\sigma_j}{\alpha_i + \alpha_j}\left\{\frac{1}{\alpha_i}\left[B_i(\tau)+\tau\right] + \frac{1}{\alpha_j}\left[B_j(\tau)+\tau\right] - B_i(\tau)B_j(\tau)\right\},$$

für $1 \le i, j \le 3$

(3.25c) $$A(\tau) = -\lambda^T C^T \alpha^{-1}\left[B(\tau) + \tau E_r\right] - r_0\tau + \frac{1}{2}\sum_{1 \le i,j \le 3} A_{ij}(\tau),$$

wobei E_r die (r,r)-Einheitsmatrix bezeichne.

Bei der Spezifikation von $R = \mathrm{Cov}(\varepsilon)$ gehen wir unverändert von (3.10a) aus. Damit sind die Ausgangsdaten zur Anwendung der Kalman-Filter-Prozedur sämtlich festgelegt. Der Parametervektor, bezüglich dessen die Likelihood-Funktion (2.16) gemäß Abschnitt 2.8 zu maximieren ist, lautet hierbei $\theta = (r_0, \alpha_1, \alpha_2, \alpha_3, \sigma_1, \sigma_2, \sigma_3, \rho_{12}, \rho_{13}, \rho_{23}, \lambda_1, \lambda_2, \lambda_3, \sigma_\varepsilon)$, es sind damit vierzehn Parameter zu bestimmen.

4. Zinsstrukturmodelle des Cox/Ingersoll/Ross (CIR)-Typus

4.1 Vorüberlegungen zur Gewinnung der Zustandsgleichung

Ausgangspunkt der weiteren Überlegungen ist wiederum die generelle Faktorisierung (3.1). Im Unterschied zum Gauß-Fall ist in den folgenden Fällen $\mathrm{Cov}(X_{t+1} \mid X_t)$ aber nicht mehr deterministisch, sondern eine Funktion von X_t, d.h. $\mathrm{Cov}(X_{t+1} \mid X_t) = h(X_t)$. Aus dieser Tatsache ergeben sich mehrere Probleme. Zunächst ist (3.1) nun keine äquivalente Darstellung von (2.1a) mehr. Es liegt nicht mehr der Fall des Standard-Kalman-Filters, sondern eines verallgemeinerten Kalman-Filters vor, vgl. hierzu etwa HARVEY (1990, S. 156 f.). Zum anderen ist X_t nach wie vor nicht beobachtbar. Ein auf DUAN/SIMONATO (1999) zurückgehender Vorschlag besteht nun darin, X_t durch $\hat{X}_{t|t}$ zu ersetzen, die lineare Projektion von X_t auf $\{1, Y_t, \ldots, Y_t\}$, diese wird ja durch den Kalman-Filter generiert.

Die Problematik pflanzt sich des Weiteren fort in die Thematik der Maximum Likelihood-Schätzung. Das Residuum V^*_{t+1} aus (3.1a) ist nun nicht mehr normalverteilt, die Likelihood-

Funktionen aus (2.14b) bzw. (2.16) sind nicht mehr korrekt. Die Beibehaltung dieser Likelihood-Funktionen (d.h., der Approximation der wahren Dichte durch eine normalverteilte Dichte) führt auf die sogenannte Quasi-Likelihood-Methode. Da im Weiteren aber X_t durch $\hat{X}_{t|t}$ ersetzt wird, resultiert aus dieser Vorgehensweise insgesamt nur eine approximative Variante der Quasi-Likelihood-Methode. Dieser Problemkreis wird in der Literatur ausführlich diskutiert, insbesondere in DE JONG (2002), DUAN/SIMONATO (1999) und LUND (1997). Die in diesem Zusammenhang durchgeführten empirischen Simulationen deuten darauf hin, dass der aus der dargestellten Vorgehensweise resultierende Bias gering ist. Im Weiteren schließen wir uns daher der Vorgehensweise von DUAN/SIMONATO (1999) an, wie zuvor schon BELETSKI (2003), DE JONG (2000) und FISCHER/MAY/WALTHER (2004).

4.2 Einfaktormodell von CIR

Das Einfaktormodell von Cox/Ingersoll/Ross (CIR) beruht[22] auf der Annahme eines Quadratwurzelprozesses für die Entwicklung der Zinsintensität $\{R_t\}$, üblicherweise geschrieben als stochastische Diffferentialgleichung der Form

(4.1) $$dR_t = \alpha(\mu - R_t)dt + \sigma\sqrt{R_t}dW_t .$$

Dabei gilt $\alpha = \alpha_P$ und $\mu = \mu_P$, d.h. der Prozess wird unter dem physischen Wahrscheinlichkeitsmaß P spezifiziert. Dabei ist $2\alpha\mu > \sigma^2$ voraus zu setzen, damit der Prozess im positiven Bereich bleibt.

Zur Gewinnung der Zustandsgleichung knüpfen wir unter Berücksichtigung der Diskussion in Abschnitt 4.1 an (3.1) an und benutzen die Ergebnisse in ALBRECHT/MAURER (2005, S. 170) für $E(R_t \mid R_s)$ bzw. $Var(R_t \mid R_s)$. Wir erhalten hieraus

(4.2a) $$E(R_{t+1} \mid R_t) = \mu\left(1 - e^{-\alpha}\right) + e^{-\alpha}R_t$$

(4.2b) $$Var(R_{t+1} \mid R_t) = \mu\frac{\sigma^2}{2\alpha}\left(1 - e^{-\alpha}\right)^2 + \frac{\sigma^2}{\alpha}\left(e^{-\alpha} - e^{-2\alpha}\right)R_t .$$

Aufgrund der Abhängigkeit von R_t ist $Var(R_{t+1} \mid R_t)$ damit insbesondere nicht mehr – im Gegensatz zum Vasicek-Fall – deterministisch.

[22] Vgl. etwa ALBRECHT/MAURER (2005, S. 169 f.).

Die entsprechende Version der Zustandsgleichung (2.1a) lautet somit:

(4.3a) $$R_{t+1} = f + FR_t + V_{t+1},$$

wobei

(4.3b) $$f = \mu\left(1 - e^{-\alpha}\right)$$

(4.3c) $$F = e^{-\alpha}.$$

Die Parameter f und F sind somit identisch im Vergleich zum Vasicek-Fall. Die Varianz-/Kovarianz-Matrix von V_{t+1} lautet hingegen

(4.3d) $$Q = Q\left(R_t\right) = \mu\frac{\sigma^2}{2\alpha}\left(1 - e^{-\alpha}\right)^2 + \frac{\sigma^2}{\alpha}\left(e^{-\alpha} - e^{-2\alpha}\right)R_t.$$

Bei der Implementierung des Kalman-Filters ist dabei für $t = 1, 2, \ldots$ die Zinsintensität R_t durch $\hat{R}_{t|t}$ gemäß (2.9a) zu ersetzen. $\hat{R}_{1|1}$ greift dabei insbesondere auf den Startwert $\hat{R}_{1|0}$ der Kalman-Rekursion zurück.

Die Abhängigkeit der Varianz-/Kovarianzmatrix Q von R_t zieht ein weiteres Problem nach sich, das – beim Einsatz des Kalman-Filters – bei allen Zinsstrukturmodellen vom CIR-Typus auftritt. Ein negativer Projektionswert von $\hat{R}_{t|t}$ (bzw. allgemein $\hat{X}_{t|t}$) ist nicht zulässig und kann des Weiteren zu einer Matrix Q führen, die nicht mehr positiv definit ist. CHEN/SCOTT (2003, S. 149) schlagen in diesem Zusammenhang vor, einen negativen Projektionswert $\hat{R}_{t|t}$ (bzw. negative Komponenten von $\hat{X}_{t|t}$) durch null zu ersetzen, um die Restriktion $R_t \geq 0$ (bzw. $X_i\left(t\right) \geq 0$ für alle i) zu wahren.[23]

Der Startwert $\hat{R}_{1|0} = E\left(R_1\right)$ und die assoziierte Größe $P_1 = Var\left(R_1\right)$ ergeben sich[24] zu

(4.4a) $$\hat{R}_{1|0} = \mu$$

und

(4.4b) $$P_1 = \mu\frac{\sigma^2}{2\alpha}.$$

[23] Dieses Censoring der Projektionswerte der Zustandsvariablen ist zwar notwendig, führt aber möglicherweise zu zusätzlichen Verzerrungen beim Schätzprozess, vgl. etwa LUND (1997, S. 10).

[24] Vgl. Anhang C.1 oder DE JONG (2000, S. 312).

Die Spot Rates $R(t, t+\tau_i)$ sind wie im Vasicek-Fall strukturell gegeben durch (3.9a), allerdings mit unterschiedlichen Werten für die Funktionen $B(\tau)$ sowie $A(\tau)$. Es gilt[25] mit $\gamma := \sqrt{(\alpha+\lambda)^2 + 2\sigma^2}$:

(4.5a) $$A(\tau) = \frac{2\alpha\mu}{\sigma^2} \ln\left[\frac{2\gamma e^{(\alpha+\lambda+\gamma)\tau/2}}{(\alpha+\lambda+\gamma)\left(e^{\gamma\tau}-1\right)+2\gamma}\right]$$

(4.5b) $$B(\tau) = \frac{2\left(e^{\gamma\tau}-1\right)}{(\alpha+\lambda+\gamma)\left(e^{\gamma\tau}-1\right)+2\gamma}.$$

Bei der Spezifikation von $R = \mathrm{Cov}(\varepsilon)$ gehen wir unverändert von (3.10a) aus. Damit sind die Ausgangsdaten zur Anwendung der Kalman-Filter-Prozedur sämtlich festgelegt. Der Parametervektor, bezüglich dessen die (Quasi-)Likelihood-Funktion (2.16) gemäß Abschnitt 2.8 zu maximieren ist, lautet hierbei $\theta = (\mu, \alpha, \sigma, \lambda, \sigma_\varepsilon)$, es sind damit fünf Parameter zu bestimmen.

4.3 Zweifaktormodell des CIR-Typus: Unabhängige Faktoren

Das Zweifaktormodell des CIR-Typus, wobei die beiden Faktoren als unabhängig vorausgesetzt werden, wird wie im Weiteren dargestellt spezifiziert. Wir gehen zunächst aus von dem folgenden Modell für die Zinsintensität:

(4.6) $$R_t = X_1(t) + X_2(t).$$

Dabei ist $X_t = \left(X_1(t), X_2(t)\right)^T$ ein zweidimensionaler Quadratwurzelprozess mit der P-Dynamik

(4.7a) $$dX_t = \alpha(\mu - X_t)dt + C\sqrt{X_t}dW_t.$$

Dabei ist die symbolisch mit $\sqrt{X_t}$ bezeichnete Matrix gegeben durch

(4.7b) $$\sqrt{X_t} = \begin{pmatrix} \sqrt{X_1(t)} & 0 \\ 0 & \sqrt{X_2(t)} \end{pmatrix}.$$

Aufgrund der Unabhängigkeit der Faktoren $X_1(t)$ und $X_2(t)$ reduziert sich die Cholesky-Matrix C gemäß (3.12d) zu

[25] Dabei wurde der überwiegenden Vorgehensweise in der Literatur gefolgt und die Risikoprämie pro Einheit Risiko ausgedrückt, d.h. λ durch λ/σ ersetzt. Im Unterschied hierzu vgl. etwa KWOK (1998, S. 325).

(4.7c) $$C = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}$$

und mit

(4.7d) $$\alpha = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}$$

sowie

(4.7e) $$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

lässt sich (4.7a) wie folgt in die äquivalente Dynamik der eindimensionalen Prozesse $X_1(t)$ und $X_2(t)$ zerlegen:

(4.8a) $$dX_1(t) = \alpha_1\left(\mu_1 - X_1(t)\right)dt + \sigma_1\sqrt{X_1(t)}dW_1(t)$$

(4.8b) $$dX_2(t) = \alpha_2\left(\mu_2 - X_2(t)\right)dt + \sigma_2\sqrt{X_2(t)}dW_2(t)$$

Um die Positivität von $\{R_t\}$ zu garantieren, muss dabei $2\alpha_i\mu_i > \sigma_i^2$ für $i = 1,2$ gelten.

Anzumerken ist schließlich, dass es – im Unterschied zum Vasicek-Fall des Kapitels 3 – im Fall eines mehrdimensionalen Quadratwurzelprozesses nicht zulässig ist, unter der P-Dynamik $\mu = 0$ zu setzen und in einem einzigen Parameter r_0 zu kondensieren[26].

Aufgrund der angenommenen Unabhängigkeit der Faktoren können die Resultate des Abschnitts 4.2 direkt auf den zweidimensionalen Fall übertragen werden.

Die Zustandsgleichung lautet

(4.9a) $$\begin{pmatrix} X_1(t+1) \\ X_2(t+1) \end{pmatrix} = \begin{pmatrix} \mu_1\left(1-e^{-\alpha_1}\right) \\ \mu_2\left(1-e^{-\alpha_2}\right) \end{pmatrix} + \begin{pmatrix} e^{-\alpha_1} & 0 \\ 0 & e^{-\alpha_2} \end{pmatrix}\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} + V_{t+1}.$$

Dabei ist $Q = \mathrm{Cov}\left(X_{t+1} \mid X_t\right)$ eine Diagonalmatrix mit ($i = 1,2$)

(4.9b) $$[Q]_{ii} = \mu_i \frac{\sigma_i^2}{2\alpha_i}\left(1-e^{-\alpha_i}\right)^2 + \frac{\sigma_i^2}{\alpha_i}\left(e^{-\alpha_i} - e^{-2\alpha_i}\right)X_i(t).$$

---

[26] Vgl. hierzu BELETSKI (2003, S. 61) sowie DAI/SINGLETON (2000, S. 1949 und S. 1973). Abweichend hiervon arbeitet DE JONG (2000) generell mit der Annahme $\mu_P = 0$.

Im Rahmen der Kalman-Rekursion ist gemäß der Vorbemerkungen in Abschnitt 4.1 dabei $X_t = \left(X_1(t), X_2(t)\right)^T$ zu ersetzen durch $\hat{X}_{t|t}$.

Die Startwerte der Rekursion ergeben sich durch

(4.10a) $$\hat{X}_{1|0} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

und

(4.10b) $$P_1 = \begin{pmatrix} \mu_1 \dfrac{\sigma_1^2}{2\alpha_1} & 0 \\ 0 & \mu_2 \dfrac{\sigma_2^2}{2\alpha_2} \end{pmatrix}.$$

Kommen wir damit zur Spezifikation der Beobachtungsgleichung. Grundsätzlich gilt hier unverändert die Darstellung (3.17a) des zweidimensionalen Vasicek-Falles. Die Funktionen A und B ergeben sich gemäß

(4.11a) $$A(\tau) = A_1(\tau) + A_2(\tau)$$

(4.11b) $$B(\tau) = \left(B_1(\tau), B_2(\tau)\right)^T,$$

wobei $(i = 1, 2)$

(4.11c) $$\gamma_i = \sqrt{\left(\alpha_i + \lambda_i\right)^2 + 2\sigma_i^2}$$

(4.11d) $$A_i(\tau) = \frac{2\alpha_i\mu_i}{\sigma_i^2} \ln\left[\frac{2\gamma_i e^{(\alpha_i+\lambda_i+\gamma_i)\tau/2}}{\left(\alpha_i + \lambda_i + \gamma_i\right)\left(e^{\gamma_i\tau} - 1\right)}\right]$$

(4.11e) $$B_i(\tau) = \frac{2\left(e^{\gamma_i\tau} - 1\right)}{\left(\alpha_i + \lambda_i + \gamma_i\right)\left(e^{\gamma_i\tau} - 1\right) + 2\gamma_i}.$$

Bei der Spezifikation von $R = \mathrm{Cov}(\varepsilon)$ gehen wir unverändert von (3.10a) aus. Damit sind die Ausgangsdaten zur Anwendung der Kalman-Filter-Prozedur sämtlich festgelegt. Der Parametervektor, bezüglich dessen die (Quasi-)Likelihood-Funktion zu maximieren ist, lautet $\theta = \left(\mu_1, \mu_2, \alpha_1, \alpha_2, \sigma_1, \sigma_2, \lambda_1, \lambda_2, \sigma_\varepsilon\right)$, es sind damit neun Parameter zu bestimmen.

4.4 Dreifaktormodell des CIR-Typus: Unabhängige Faktoren

Die Ergebnisse des Abschnitts 4.3 können aufgrund der Unabhängigkeit der Faktoren direkt auf den dreidimensionalen Fall verallgemeinert werden. Es gilt:

(4.12) $$R_t = X_1(t) + X_2(t) + X_3(t).$$

Die P-Dynamik von $X_t = \left(X_1(t), X_2(t),, X_3(t)\right)^T$ ist gegeben durch

(4.13) $$dX_t = \alpha\left(\mu - X_t\right)dt + C\sqrt{X_t}\,dW_t.$$

Dabei sind α, C und $\sqrt{X_t}$ Diagonalmatrizen mit $\alpha = \mathrm{diag}\left(\alpha_1, \alpha_2, \alpha_3\right)$, $C = \mathrm{diag}\left(\sigma_1, \sigma_2, \sigma_3\right)$ und $\sqrt{X_t} = \mathrm{diag}\left(X_1(t), X_2(t), X_3(t)\right)$ und es gilt $\mu = \left(\mu_1, \mu_2, \mu_3\right)^T$. Dabei ist $2\alpha_i\mu_i > \sigma_i^2$ ($i = 1, 2, 3$) zu fordern, damit die Positivität von $\{R_t\}$ gewährleistet ist.

Die Zustandsgleichung lautet

(4.14a) $$\begin{pmatrix} X_1(t+1) \\ X_2(t+1) \\ X_3(t+1) \end{pmatrix} = \begin{pmatrix} \mu_1\left(1 - e^{-\alpha_1}\right) \\ \mu_2\left(1 - e^{-\alpha_2}\right) \\ \mu_3\left(1 - e^{-\alpha_3}\right) \end{pmatrix} + \begin{pmatrix} e^{-\alpha_1} & 0 & 0 \\ 0 & e^{-\alpha_2} & 0 \\ 0 & 0 & e^{-\alpha_3} \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} + V_{t+1}.$$

Dabei ist $Q = \mathrm{Cov}\left(X_{t+1} \mid X_t\right)$ eine Diagonalmatrix mit ($i = 1, 2, 3$)

(4.14b) $$\left[Q\right]_{ii} = \mu_i \frac{\sigma_i^2}{2\alpha_i}\left(1 - e^{-\alpha_i}\right)^2 + \frac{\sigma_i^2}{\alpha_i}\left(e^{-\alpha_i} - e^{-2\alpha_i}\right)X_i(t).$$

Im Rahmen der Kalman-Rekursion ist dabei gemäß den Vorüberlegungen in Abschnitt 4.1 $X_t = \left(X_1(t), X_2(t), X_3(t)\right)^T$ zu ersetzen durch $\hat{X}_{t|t}$.

Die Startwerte der Rekursion ergeben sich durch

(4.15a) $$\hat{X}_{1|0} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}$$

und

(4.15b) $$\left[P_1\right]_{ii} = \mu_i \frac{\sigma_i^2}{2\alpha_i}.$$

Für die Beobachtungsgleichung gilt wie im Vasicek-Fall grundsätzlich die Darstellung (3.24 a). Die Funktionen A und B ergeben sich gemäß

(4.16a) $$A(\tau) = A_1(\tau) + A_2(\tau) + A_3(\tau)$$

(4.16b) $$B(\tau) = \left(B_1(\tau), B_2(\tau), B_3(\tau)\right)^T,$$

wobei $(i = 1, 2, 3)$

(4.16c) $$\gamma_i = \sqrt{(\alpha_i + \lambda_i)^2 + 2\sigma_i^2}$$

(4.16d) $$A_i(\tau) = \frac{2\alpha_i\mu_i}{\sigma_i^2} \ln\left[\frac{2\gamma_i e^{(\alpha_i+\lambda_i+\gamma_i)\tau/2}}{(\alpha_i + \lambda_i + \gamma_i)\left(e^{\gamma_i\tau} - 1\right)}\right]$$

(4.16e) $$B_i(\tau) = \frac{2\left(e^{\gamma_i\tau} - 1\right)}{(\alpha_i + \lambda_i + \gamma_i)\left(e^{\gamma_i\tau} - 1\right) + 2\gamma_i}.$$

Bei der Spezifikation von $R = \mathrm{Cov}(\varepsilon)$ gehen wir unverändert von (3.10a) aus. Damit sind die Ausgangsdaten zur Anwendung der Kalman-Filter-Prozedur sämtlich festgelegt. Der Parametervektor, bezüglich dessen die (Quasi-)Likelihood-Funktion zu maximieren ist, lautet $\theta = (\mu_1, \mu_2, \mu_3, \alpha_1, \alpha_2, \alpha_3, \sigma_1, \sigma_2, \sigma_3, \lambda_1, \lambda_2, \lambda_3, \sigma_\varepsilon)$, es sind damit dreizehn Parameter zu bestimmen.

4.5 Mehrfaktormodell des CIR-Typus: Korrelierte Faktoren

Im Falle abhängiger Faktoren lautet, ausgehend von einer Zinsintensität der Form

(4.17) $$R_t = \sum_{i=1}^{r} X_i(t),$$

die P-Dynamik der Faktoren $X_t = \left(X_1(t), \ldots, X_r(t)\right)^T$

(4.18) $$dX_t = \alpha(\mu - X_t)dt + C\sqrt{X_t}dW_t.$$

Dabei sind α und $\sqrt{X_t}$ Diagonalmatrizen[27] mit $\alpha = \mathrm{diag}(\alpha_1, \ldots, \alpha_r)$ und $\sqrt{X_t} = \mathrm{diag}\left(\sqrt{X_1(t)}, \ldots, \sqrt{X_r(t)}\right)$. Für μ gilt $\mu = (\mu_1, \ldots, \mu_r)^T$ und C ist die Cholesky-Zerlegung einer Varianz-/Kovarianzmatrix.

[27] Wie in Anhang A.1 angemerkt, beinhaltet die Wahl von α als Diagonalmatrix im Grunde eine Überidentifizierung der Parameter und eine untere Dreiecksmatrix wäre vorzuziehen. Anhang A.1 enthält ebenfalls eine allgemeinere Version von $\sqrt{X_t}$.

Mehrfaktormodelle des CIR-Typus sind im Falle korrelierter Faktoren mit einer Reihe von Problemen verbunden. Zunächst ergeben sich, wie in Anhang A.1 ausgeführt, eine Reihe von Restriktionen, um die Positivität von $\{R_t\}$ zu gewährleisten sowie eine Reihe von Normierungen, um die eindeutige Identifizierbarkeit der Parameter sicherzustellen. Eine weitere Problematik besteht darin, dass das Differentialgleichungssystem (A.9) bzw. (A.10) zur Gewinnung der Funktionen $A(\tau)$ und $B(\tau)$ im Gegensatz zu den bisher betrachteten Fällen nur noch numerisch zu lösen ist.

Eine Anwendung des Kalman-Filters auf Mehrfaktormodelle des CIR-Typus mit korrelierten Faktoren erfolgt unseres Wissens bisher nur in DE JONG (2000). DE JONG arbeitet allerdings einheitlich mit der Restriktion $\mu_P = 0$, die – wie bereits in Abschnitt 4.3 angesprochen - problematisch ist.

DAI/SINGLETON (2000, S. 1970) kommen darüber hinaus zum Schluss, dass auf der einen Seite nur nicht-negative Korrelationen zwischen den Faktoren zu einem zulässigen Modell führen und auf der anderen Seite die empirischen Daten eher für eine negative Korrelation der Faktoren sprechen. Durch eine Zulassung (notwendigerweise positiver) Korrelation wäre damit im Vergleich zum unkorrelierten Fall nichts gewonnen. Wie schon BELETSKI (2003) verzichten wir daher auf die empirische Identifizierung von multifaktoriellen CIR-Modellen mit korrelierten Faktoren.

LITERATUR

Albrecht, P., R. Maurer (2005): Investment- und Risikomanagement, 2. Aufl., Stuttgart.

Babbs, S.H., K.B. Nowman (1999): Kalman Filtering of Generalized Vasicek Term Structure Models, Journal of Financial and Quantitative Analysis 34, S. 115-130.

Beletski, T. (2003): Forecasting the Term Structure of Interest Rates with Stochastic Models, Helsinki, University of Technology.

Cairns, A.J.G. (2004): Interest Rate Models, Princeton, Oxford.

Chen, R.-R., L. Scott (2003): Multi-Factor CIR Models of the Term Structure: Estimates and Tests from a State-Space Model Using a Kalman Filter, Journal of Real Estate Finance and Economics 27, S. 143-172.

Dai, Q., K.J. Singleton (2000): Specification Analysis of Affine Term Structure Models, Journal of Finance 55, S. 1943-1978.

De Jong, F. (2000): Time Series and Cross-section Information in Affine Term-Structure Models, Journal of Business & Economic Statistics 18, S. 300-314.

Duan, J.C., J.G. Simonato (1999): Estimating Exponential-Affine Term Structure Models by Kalman Filter, Review of Quantitative Finance and Accounting 13, S. 111-135.

Duffie, D., R. Kan (1996): A Yield Factor Model of Interest Rates, Mathematical Finance 16, S. 379-406.

Fischer, T., A. May, B. Walther (2003): Anpassung eines CIR-1-Modells zur Simulation der Zinsstrukturkurve, Blätter der Deutschen Gesellschaft für Versicherungs- und Finanzmathematik, Band XXVI, Heft 2, November 2003, S. 193-206.

Fischer, T., A. May, B. Walther (2004): Anpassung eines CIR-k-Modells zur Simulation der Zinsstrukturkurve, Blätter der Deutschen Gesellschaft für Versicherungs- und Finanzmathematik, Band XXVI, Heft 3, Mai 2004, S. 309-387.

Geyer, A., S. Pichler (1999): A state-space approach to estimate and test multifactor Cox-Ingersoll-Ross models of the term structure, Journal of Financial Research 22, S. 167-130.

Hamilton, J.D. (1994): Time Series Analysis, Princeton, N.J.

Harvey, A.C. (1989): Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge.

Jegadeesh, N., G.G. Pennacchi (1996): The Behaviour of Interest Rates Implied by the Term Structure of Eurodollar Futures, Journal of Money, Credit and Banking 28, S. 426-446.

Kwok, Y.K. (1998): Mathematical Models of Financial Derivatives, Singapur.

Lund, J. (1997): Econometric Analysis of Continuous-Time Arbitrage-Free Models of the Term Structure of Interest Rates, Aarhus School of Business.

Pennacchi, G. (1991): Identifying the Dynamics of Real Interest Rates and Inflation: Evidence Using Survey Data, Review of Financial Studies 4, S. 53-86.

Anhang A: Allgemeine Zinsstrukturmodelle

A.1 Das allgemeine affine Modell

Im Weiteren gehen wir auf zentrale strukturelle Resultate für den Fall eines allgemeinen mehrdimensionalen Zinsstrukturmodells der affinen Form ein, man vgl. hierzu DUFFIE/KAN (1996) sowie CAIRNS (2004, S. 102 ff.).

Die Zinsintensität R_t ist gegeben durch die Spezifikation

(A.1) $$R_t = r_0 + \sum_{i=1}^{r} X_i(t) = r_0 + e^T X_t .$$

Dabei ist $r_0 \in IR$, $e = (1,\ldots,1)^T$ und $X_t = \left(X_1(t),\ldots,X_r(t)\right)^T$ ein r-dimensionaler Diffusionsprozess.

Ziel ist die Ableitung eines Ausdrucks für die Zerobondpreise der Form

(A.2) $$\begin{aligned} B(t,t+\tau) &= E_t^Q\left[\exp\left(-\int_t^{t+\tau} R(s)\,ds\right)\right] \\ &= \exp\left[A(\tau) + \sum_{j=1}^{r} B_j(\tau) X_j(t)\right] = \exp\left[A(\tau) + B(\tau)^T X_t\right], \end{aligned}$$

einer sogenannten *zeithomogenen affinen Zinsstruktur*.

Die Spot Rates sind in diesem Falle gegeben durch ($\tau > 0$)

(A.3) $$R(t,t+\tau) = -\frac{1}{\tau}\left[A(\tau) + B(\tau)^T X_t\right].$$

Der folgende allgemeine r-dimensionale Diffusionsprozess führt auf eine zeithomogene affine Zinsstruktur und umfasst zahlreiche in der Literatur vertretene Spezialfälle:

(A.4a) $$dX_t = \mu_P(X_t)\,dt + \sigma(X_t)\,dW_P(t),$$

wobei

(A.4b) $$\mu_P(X_t) = \alpha_P(\mu_P - X_t)$$

und

(A.4c) $$\sigma(X_t) = C\sqrt{S_t} .$$

Dabei ist anzumerken, dass der Diffusionsprozess unter dem physischen (beobachtbaren) Wahrscheinlichkeitsmaß P spezifiziert wurde. Auf die Repräsentation unter dem risikoneutralen Wahrscheinlichkeitsmaß wird im Weiteren noch eingegangen.

Die Matrix α_P ist eine reellwertige (r,r)-Diagonalmatrix[28], $\alpha_P = \mathrm{diag}\left(\alpha_1^P,\ldots,\alpha_r^P\right)$. Der Vektor $\mu_P = \left(\mu_P(1),\ldots,\mu_P(r)\right)^T$ ein reellwertiger $(r,1)$-Vektor. Die Matrix $C = \Sigma^{1/2}$ ist die Cholesky-Zerlegung einer Varianz-/Kovarianzmatrix $\Sigma = \left(\sigma_{ij}\right)$. Insbesondere ist C eine untere Dreiecksmatrix. Die symbolisch mit $\sqrt{S_t}$ bezeichnete Matrix ist eine (r,r)-Diagonalmatrix, deren i-tes Diagonalelement gegeben ist durch

(A4.d) $$\left[\sqrt{S_t}\right]_{ii} = a_i + b_i^T X_t,$$

die Skalare a_i und die $(r,1)$-Vektoren b_i sind dabei jeweils reellwertig.

Für weitergehende Zwecke fassen wir noch die Skalare a_i in einem $(r,1)$-Vektor $a = \left(a_1,\ldots,a_r\right)^T$ zusammen und die Vektoren b_i in eine (r,r)-Matrix $B = \left(b_1,\ldots,b_r\right)$.

Ist $B = 0$, so führt die Spezifikation (A.4) auf einen r-dimensionalen Gauß-Prozess und es treten keine Zulässigkeitsprobleme auf. Im Allgemeinen ist der Modellansatz (A.4) nur zulässig (admissible), wenn $\left[\sqrt{S_t}\right]_{ii} > 0$ für $i = 1,\ldots,r$. Im Allgemeinen sind daher Restriktionen an die Parameter notwendig, um die Zulässigkeit zu sichern. Im Allgemeinen existiert darüber hinaus ein Identifizierungsproblem, da nicht alle Parameter isoliert zu identifizieren sind, beispielsweise ist nur die Produktmatrix $C^{-1}\alpha C$ identifizierbar. Dies macht eine Reihe von Normierungen der in das Modell (A.4) eingehenden Matrizen bzw. Vektoren notwendig. Zur ausführlichen Diskussion der Problemkreise "Zulässigkeit" und "Identifizierung" bzw. "Normierung" verweisen wir auf DUFFIE/KAN (1996), DAI/SINGLETON (2000) und DE JONG (2000).

Kommen wir nun auf die Zusammenhänge zwischen der Spezifikation des Prozesses $\{X_t\}$ unter Q auf der einen Seite und unter dem physischen (beobachtbaren) Wahrscheinlichkeitsmaß P auf der anderen. Hierzu betrachten wir einen Vektor $\lambda(t) = \left(\lambda_1(t),\ldots,\lambda_r(t)\right)^T$ von

[28] Wie DAI/SINGLETON (2000, S. 1995) anmerken, führt die Annahme einer Diagonalmatrix zu einer Überidentifizierung der Parameter. Im allgemeinen ist α_P als untere Dreiecksmatrix zu wählen, deren Werte außerhalb der Diagonale nur kleiner oder gleich null sind. Wir folgen trotzdem der üblichen Vorgehensweise, α_P als Diagonalmatrix zu postulieren, um die Anzahl der offenen – und damit empirisch zu identifizierenden Parameter – unter Kontrolle zu halten.

Marktpreisen des Risikos. Unter Q gelten dann für Drift und Diffusion des Prozesses die allgemeinen Beziehungen

(A.5a) $$\mu_Q(X_t) = \mu_P(X_t) - \sigma(X_t)\lambda(t)$$

(A.5b) $$\sigma_Q(X_t) = \sigma(X_t),$$

d.h. nur die Drift ändert sich bei Betrachtung des Prozesses unter Q.

Im Kontext der für diese Ausarbeitung relevanten Modelle wird nun die folgende Annahme für die Struktur der Marktpreise des Risikos getroffen, wobei $\lambda := (\lambda_1, \ldots, \lambda_r)^T$:

(A.6) $$\lambda(t) = \sqrt{S_t}\lambda .$$

In diesem Falle geht (A.5a) zunächst über in

(A.7a) $$\mu_Q(X_t) = \mu_P(X_t) - CS_t\lambda ,$$

wobei S_t eine Diagonalmatrix mit $\left[\sqrt{S_t}\right]_{ii} = a_i + b_i^T X_t$ ist. Die Beziehung (A.7a) lässt sich daher in disaggregierter Form schreiben als

(A.7b) $$\mu_Q(X_t) = \mu_P(X_t) - C\psi - C\Phi X_t$$

wobei ψ ein $(r,1)$-Vektor $\psi = (\psi_1, \ldots, \psi_r)^T$ mit $\psi_i = \lambda_i a_i$ ist und Φ eine (r,r)-Matrix, deren i-te Zeile gegen ist durch $\lambda_i b_i^T$, d.h.

$$\Phi = \begin{pmatrix} \lambda_1 b_1^T \\ \vdots \\ \lambda_r b_r^T \end{pmatrix}.$$

Notieren wir nun noch $\mu_Q(X_t)$ in der Form

(A.7c) $$\mu_Q(X_t) = \alpha_Q(\mu_Q - X_t),$$

so gelten die Zusammenhänge

(A.8a) $$\alpha_P = \alpha_Q - C\Phi$$

und

(A.8b) $$\alpha_P\mu_P = \alpha_Q\mu_Q + C\psi$$

bzw.

(A.8c) $$\mu_P = \alpha_P^{-1}\left[\alpha_Q\mu_Q + C\psi\right].$$

Die eindimensionale reelle Funktion $A(\tau)$ und die vektorwertige reelle Funktion $B(\tau) = \left(B_1(\tau), \ldots, B_r(\tau)\right)^T$ genügen nun dem folgenden System von (gewöhnlichen) Differentialgleichungen, vgl. etwa DAI/SINGLETON (2000, S. 1947) oder BELETSKI (2003, S. 55):

(A.9a) $$dA(\tau)/d\tau = \left(\alpha_Q \mu_Q\right)^T B(\tau) + \frac{1}{2}\sum_{i=1}^{r}\left[C^T B(\tau)\right]_i^2 a_i - r_0$$

(A.9b) $$dB(\tau)/d\tau = -\alpha_Q B(\tau) + \frac{1}{2}\sum_{i=1}^{r}\left[C^T B(\tau)\right]_i^2 b_i - e\,.$$

Dabei bedeute generell $[A]_i$ die i-te Zeile der Matrix A.

Die folgende alternative Darstellung bietet DE JONG (2000, S. 301):

(A.10a) $$dA(\tau)/d\tau = \left(\alpha_Q \mu_Q\right)^T B(\tau) + \frac{1}{2}\sum_i \sum_j B_i(\tau) B_j(\tau) a_{ij} - r_0$$

(A.10b) $$dB(\tau)/d\tau = -\alpha_Q B(\tau) + \frac{1}{2}\sum_i \sum_j B_i(\tau) B_j(\tau) b_{ij} - e\,,$$

wobei die Skalare a_{ij} und die Vektoren b_{ij} der Gleichung

(A.10c) $$a_{ij} + b_{ij}^T x = \left[C\,\mathrm{diag}\left(a + B^T x\right) C^T\right]_{ij}$$

genügen.

Das Differentialgleichungssystem (A.9) bzw. (A.10) ist in einigen zentralen Fällen explizit lösbar, im Allgemeinen – vgl. etwa DAI/SINGLETON (2000, S. 1947) – jedoch durch numerische Integration unter Zugrundelegung der Startbedingungen $A(0) = 0$ und $B(0) = 0$.

A.2 Das verallgemeinerte Vasicek-Modell

Das verallgemeinerte Vasicek-Modell gewinnt man, indem man im Rahmen von (A.4) jeweils $b_i = 0$ (bzw. $B = 0$ bzw. $\Phi = 0$) und $a_i = 1$ (bzw. $a = e$) setzt.

Die allgemeine stochastische Differentialgleichung (A.4) besitzt dann die Form

(A.11) $$dX_t = \alpha_P \left(\mu_P - X_t\right) dt + C dW_P\left(t\right).$$

Die Zusammenhänge (A.8) hinsichtlich des Zusammenhangs der Parameter α und μ lauten dann (man beachte $C\psi = C\lambda$)

(A.12a) $$\alpha_P = \alpha_Q = \alpha$$

(A.12b) $$\mu_P = \mu_Q + \alpha^{-1} C\lambda .$$

In dem im Haupttext relevanten Fall $\mu_P = 0$ folgt hieraus

(A.12c) $$\mu_Q = \alpha^{-1} C\lambda ,$$

wobei – der Majorität der Literatur über Gaußsche affine Zinsstrukturmodelle folgend[29], λ durch $-\lambda$ ersetzt wird.

Das Differentialgleichungssystem (A.9) lautet hiermit (man beachte $CC^T = \Sigma$)

(A.13a) $$\frac{dA\left(\tau\right)}{d\tau} = \lambda^T C^T B\left(\tau\right) + \frac{1}{2} B\left(\tau\right)^T \Sigma B\left(\tau\right) - r_0$$

(A.13b) $$\frac{dB\left(\tau\right)}{\tau} = -\alpha B\left(\tau\right) - e .$$

Da α eine Diagonalmatrix ist, kann (A.13b) relativ einfach gelöst werden und man erhält insbesondere

(A.14) $$B\left(\tau\right) = \left(B_1\left(\tau\right), \ldots, B_r\left(\tau\right)\right)^T = \left(\frac{e^{-\alpha_1 \tau} - 1}{\alpha_1}, \ldots, \frac{e^{-\alpha_r \tau} - 1}{\alpha_r}\right).$$

Zur Bestimmung von (A.13a) vgl. im Detail BELETSKI (2003, S. 87).

[29] Vgl. etwa BELETSKI (2003, S. 87).

A.3 Das verallgemeinerte CIR-Modell

Das verallgemeinerte CIR-Modell gewinnt man, indem man generell $a_i = 0$ (bzw. $a = 0$) sowie $b_{ij} = 1$ für $j = i$ und $b_{ij} = 0$ für $j \neq i$ setzt. Die allgemeine stochastische Differentialgleichung (A.4) besitzt dann die Form

(A.15) $$dX_t = \alpha_P \left(\mu_P - X_t\right) dt + C \begin{pmatrix} \sqrt{X_1(t)} & & 0 \\ & \ddots & \\ 0 & & \sqrt{X_r(t)} \end{pmatrix} dW_P(t).$$

Die Zusammenhänge (A.8) hinsichtlich der Parameter α und μ lauten dann (man beachte $\psi = 0$ und $\Phi = \mathrm{diag}\left(\lambda_1, \ldots, \lambda_r\right)$)

(A.16a) $$\alpha_P = \alpha_Q - C\lambda$$

(A.16b) $$\mu_P = \alpha_P^{-1} \alpha_Q \mu_Q .$$

Anhang B: Ermittlung der Startwerte bei Modellen des Vasicek-Typus

B.1 Einfaktormodell

Gemäß ALBRECHT/MAURER (2005, S. 168) gilt zunächst

(B.1) $$E\left(R_{t+h} \mid R_t\right) = \mu + e^{-\alpha h}\left(R_t - \mu\right)$$

sowie

(B.2) $$\mathrm{Var}\left(R_{t+h} \mid R_t\right) = \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha h}\right).$$

Hieraus folgt zunächst

(B.3) $$E\left(R_{t+h}\right) = E\left[E\left(R_{t+h} \mid R_t\right)\right] = \mu + e^{-\alpha h}\left[E\left(R_t\right) - \mu\right]$$

und damit unter Annahme der Stationarität, d.h. $E\left(R_{t+h}\right) = E\left(R_t\right)$

(B.4) $$E\left(R_t\right) - \mu = e^{-\alpha h}\left[E\left(R_t\right) - \mu\right].$$

Da $\exp\left(-\alpha h\right) > 0$ folgt hieraus

(B.5) $$E\left(R_t\right) = \mu\,.$$

Des Weiteren folgt

(B.6) $$\begin{aligned}\mathrm{Var}\left(R_{t+h}\right) &= E\left[\mathrm{Var}\left(R_{t+h} \mid R_t\right)\right] + \mathrm{Var}\left[E\left(R_{t+h} \mid R_t\right)\right] \\ &= \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha h}\right) + e^{-2\alpha h}\mathrm{Var}\left(R_t\right)\end{aligned}$$

und damit unter Annahme der Stationarität

(B.7) $$\mathrm{Var}\left(R_t\right)\left(1 - e^{-2\alpha h}\right) = \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha h}\right).$$

Insgesamt folgt damit

(B.8) $$\mathrm{Var}\left(R_t\right) = \frac{\sigma^2}{2\alpha}\,.$$

Ein entsprechendes Ergebnis erhält man aus der AR(1)-Eigenschaft von R_t, wenn man von einem "eingeschwungenen" Prozess ausgeht. Ansonsten sind (B.5) und (B.8) als approximative Beziehungen anzusehen.

B.2 Zweifaktormodell

Aus (3.14a) folgt aus $E\left(X_{t+1}\right)=E\left[E\left(X_{t+1}\mid X_t\right)\right]$ und der Beachtung der Stationarität zunächst komponentenweise ($i=1,2$) $E\left[X_i\left(t\right)\right]=e^{-\alpha_i}E\left[X_i\left(t\right)\right]$ und damit $E\left[X_i\left(t\right)\right]=0$.

Hinsichtlich der Kovarianzmatrix gilt $Cov\left(X_{t+1}\right)=E\left[Cov\left(X_{t+1}\mid X_t\right)\right]+Cov\left[E\left(X_{t+1}\mid X_t\right)\right]$. Es gilt dann unter Annahme der Stationarität komponentenweise ($\rho_{11}=\rho_{22}=1$)

(B.9) $$Cov_{ij}=\frac{\rho_{ij}\sigma_i\sigma_j}{\alpha_i+\alpha_j}e^{-\left(\alpha_i+\alpha_j\right)}+e^{-\left(\alpha_i+\alpha_j\right)}Cov_{ij}$$

und damit

(B.10) $$Cov_{ij}=\frac{\rho_{ij}\sigma_i\sigma_j}{\alpha_i+\alpha_j}.$$

Anhang C: Ermittlung der Startwerte bei Modellen des Cox/Ingersoll/Ross-Typus

Gemäß ALBRECHT/MAURER (2005, S. 170) gilt zunächst

(C.1) $$E\left(R_{t+h} \mid R_t\right) = R_t e^{-\alpha h} + \mu\left(1 - e^{-\alpha h}\right)$$

sowie

(C.2) $$Var\left(R_{t+h} \mid R_t\right) = R_t \frac{\sigma^2}{2\alpha}\left(e^{-\alpha h} - e^{-2\alpha h}\right) + \mu \frac{\sigma^2}{2\alpha}\left(1 - e^{-\alpha h}\right)^2 .$$

Hieraus folgt zunächst

(C.3) $$E\left(R_{t+h}\right) = E\left[E\left(R_{t+h} \mid R_t\right)\right] = e^{-\alpha h} E\left(R_t\right) + \mu\left(1 - e^{-\alpha h}\right)$$

und damit unter Annahme der Stationarität

(C.4) $$E\left(R_t\right)\left(1 - e^{-\alpha h}\right) = \mu\left(1 - e^{-\alpha h}\right),$$

d.h., insgesamt

(C.5) $$E\left(R_t\right) = \mu .$$

Des Weiteren folgt

(C.6) $$\begin{aligned} Var\left(R_{t+h}\right) &= E\left[Var\left(R_{t+h} \mid R_t\right)\right] + Var\left[E\left(R_{t+h} \mid R_t\right)\right] \\ &= E\left(R_t\right)\frac{\sigma^2}{2\alpha}\left(e^{-\alpha h} - e^{-2\alpha h}\right) + \mu \frac{\sigma^2}{2\alpha}\left(1 - e^{-\alpha h}\right)^2 + Var\left(R_t\right) e^{-2\alpha h} . \end{aligned}$$

Unter der Annahme der Stationariät folgt hieraus

(C.7) $$Var\left(R_t\right)\left(1 - e^{-2\alpha h}\right) = \mu \frac{\sigma^2}{\alpha}\left(e^{-\alpha h} - e^{-2\alpha h}\right) + \mu \frac{\sigma^2}{2\alpha}\left(1 - e^{-\alpha h}\right)^2 .$$

Insgesamt ergibt sich hieraus:

(C.8) $$Var\left(R_t\right) = \mu \frac{\sigma^2}{2\alpha} .$$