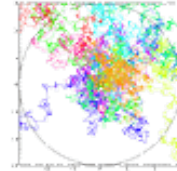


Lehrstuhl für Angewandte Statistik
und Methoden der empirischen Sozialforschung



Zustandsraummodelle und Kalman-Filter

Seminararbeit

von

Corvin Idler

c/o European External Action Service, Delegation Australia, Diplomatic Pouch,
1049 Brussels, BELGIUM
+61 41 471 288, idler@uni-koblenz.de

Studiengang	Bachelor Wirtschaftsinformatik
Betreuer	Dipl.-Kffr. Marina Lorenz
Matrikelnummer	7529953
Arbeit vorgelegt am:	13.04.2012

Inhaltsverzeichnis

1	Einleitung	3
2	Grundlagen	5
2.1	Ausgangsproblemstellung	5
2.2	Mathematische Beschreibung	5
2.2.1	Systemmodell	5
2.2.2	Messmodell	6
2.3	Zustandsschätzung	6
2.3.1	Kalman-Filter-Algorithmus	9
2.3.2	Bezug zur Bayesschen Schätztheorie	9
2.4	Likelihoodfunktion im Kalman-Filter Kontext	11
3	Ökonometrisches Anwendungsbeispiel	12
3.1	Datenbasis	12
3.2	Zinsstrukturmodell	13
3.2.1	Ein-Faktor-Fall	13
3.2.2	Multi-Faktor-Fall	14
3.3	Parameterschätzung	16
	Literaturverzeichnis	17
A	Anhang	21

1. Einleitung

Die vorliegende Arbeit hat den nach seinem Erfinder Rudolf E. Kálmán benannten Kalman-Filter[Kál60, KB61] zum Inhalt. In seiner Urform handelt es sich dabei um eine rekursive Berechnungsvorschrift der Methode der kleinsten Quadrate¹[Leg05, Gau09] im Kontext der messwertbasierten Schätzung linearer dynamischer endlichdimensionaler stochastischer Systeme in Zustandsraumdarstellung. Seinen Durchbruch in der Praxis erfuhr der Filter durch Verwendung einer weiterentwickelten Version² im Rahmen der Mondlandung der Apollo 11 Mission[SM85]. Eine interessante historische Abhandlung über seine Rolle in der Raumfahrt von Stanley F. Schmidt, dem die erste erfolgreiche Computerimplementierung des Algorithmus zugerechnet wird, findet sich in [SM85]. Es ist anzumerken, dass der Filter schon vor R. Kálmán in nahezu identischer Form durch Peter Swerling veröffentlicht wurde [Swe58, Swe59], der dafür jedoch nie wirklich die ihm gebührende wissenschaftliche Anerkennung erhielt. Der Kalman-Filter gehört heute zum Standardrepertoire der Praktiker vieler wissenschaftlicher Disziplinen und ist seit langem nicht mehr nur auf die Aeronautik als sein Ursprungsanwendungsgebiet beschränkt. So lassen sich viele praktische Problemstellungen durch eine Menge interessierender möglicherweise stochastischer Größen beschreiben, deren zeitliche Entwicklung durch Differential-³ bzw. Differenzgleichungen⁴ modelliert wird. Im Paradigma der sogenannten Zustandsraumdarstellung wird dabei die jeweils vorliegende Problemstellung als ein dynamisches System verstanden, dessen zeitliche Entwicklung sich aus den oben genannten systembeschreibenden Differential- bzw. Differenzgleichungen ableiten lässt und dessen intrinsischer Zustand zu jedem Zeitpunkt durch eine Menge sogenannter Zustandsvariablen repräsentiert wird. Diese Größen sind jedoch per definitionem unbekannt, oft stochastischer Natur und lassen sich nur über möglicherweise verrauschte Messungen schätzen. Die Problematik kann man sich exemplarisch anhand der Umfeldbeobachtung, wie in Abb. 1.1 illustriert, veranschaulichen. Die beiden Kernkomponenten stellen die Umgebung (The World) und derjenige, der die Umgebung beobachtet und eine Beschreibung von ihr erlangen möchte (The Estimator) dar. Die Zustände (hidden states) der externen Welt sind nicht direkt erfassbar oder bekannt, es kann nur auf Basis der Sensordaten, die ein Zustand hervorruft, auf selbigen zurückgeschlossen werden. Bei der Beobachtung einer Umgebung ist also immer mindestens ein Sensor involviert, der durch eine bestimmte Funktion (mapping function) die Umweltzustände auf Sensor- oder Messdaten abbildet. Diese Funktion bezeichnet man auch als Sensor-, Mess- oder Beobachtungsmodell. Um nun Rückschlüsse auf den Zustand der beobachteten Umgebung ziehen zu können, braucht man neben dem Sensormodell auch ein Modell der Umgebung. Dieses abstrahiert bzw. reduziert die Umgebung auf eine Menge von interessierenden Größen und modelliert deren zeitliche Entwicklung. Hat man sowohl ein Systemmodell als auch ein Sensormodell aufgestellt, dann kann man versuchen, auf Basis beider Modelle und unter Verwendung der Sensordaten den Zustand des Systems zu bestimmen (inverse mapping), um somit eine abstrakte Beschreibung der aktuellen Umgebung zu erlangen. In ganz ähnlicher Weise lassen sich eine Vielzahl von Problemstellungen wie beispielsweise radarbasierte Objektverfolgung⁵ oder auch ökonometrische Fragestellungen als

¹least squares algorithm

²extended Kalman-Filter

³im zeitkontinuierlichen Fall

⁴im zeitdiskreten Fall

⁵tracking

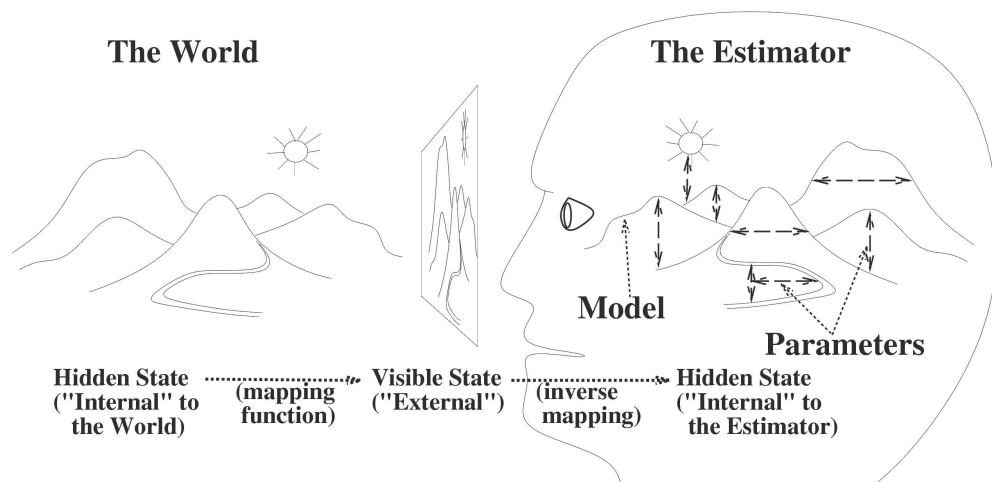


Abbildung 1.1.: Das Zustandsschätzproblem im Kontext der Umfeldbeobachtung.
Entnommen aus [O'R96, S. 72]

ein Schätzproblem unbekannter Zustände eines dynamischen Systems verstehen. Dem Kalman-Filter-Algorithmus kommt dabei die Aufgabe zu, in möglichst optimaler Art und Weise das inverse mapping vorzunehmen, also das Schätzen der Systemzustände auf Basis der Messdaten sowie des System- und Messmodells [AM91].

Der Aufbau der weiteren Ausführungen dieser Arbeit gestaltet sich wie folgt: In Kapitel 2 erfolgt eine mathematische Beschreibung der Problemstellung und des Filters. In Kapitel 3 wird dann an einem praktischen Anwendungsbeispiel aus der Ökonometrie die Arbeitsweise des Kalman-Filter zur Parameterschätzung im Kontext der Zinsstrukturmodellierung vorgeführt.

2. Grundlagen

2.1. Ausgangsproblemstellung

Von dem konkreten Beispiel der Umfeldbeobachtung aus Kapitel 1 abstrahiert, ist das Ausgangsproblem nochmal in Abbildung 2.1 schematisch visualisiert. Es gilt die Zustände ei-

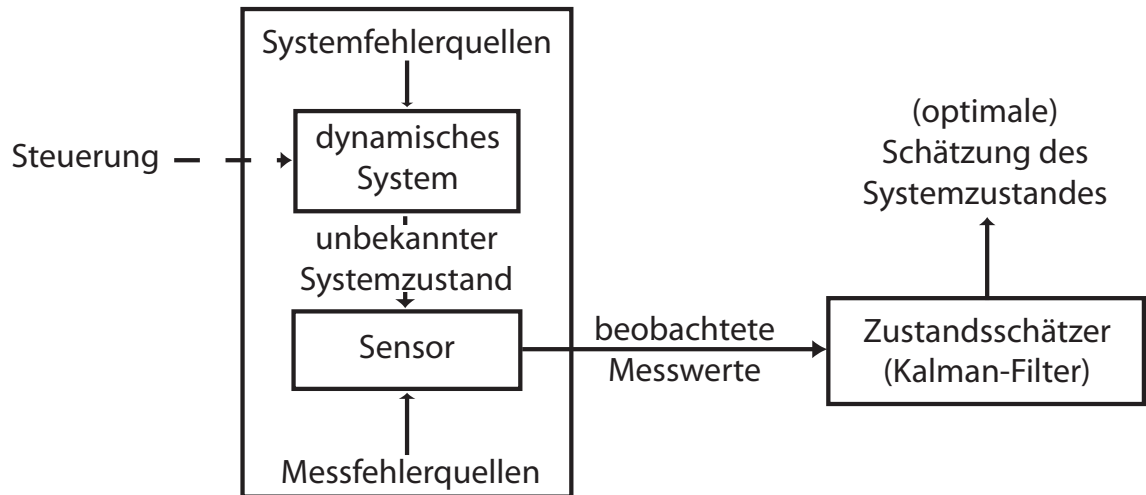


Abbildung 2.1.: Schema des typischen Einsatzkontexts eines Kalman-Filters.

In Anlehnung an [May79, S. 5]

nes linearen stochastischen dynamischen Systems auf Basis verrauschter Beobachtungswerte möglichst optimal zu schätzen. Hierzu muss man, wie schon in Kapitel 1 erwähnt, ein System- und ein Messmodell aufstellen und zusätzlich eine Rechenvorschrift zum Zustandsschätzen formulieren. Dies wird in mathematischer Form im folgenden Abschnitt beschrieben.

2.2. Mathematische Beschreibung

Die folgenden Ausführungen beschränken sich aus Platzgründen auf die zeitdiskrete Version des Kalman-Filters.

2.2.1. Systemmodell

Das System- bzw. Zustandsmodell sei im Allgemeinen [Sor70, S. 10][Kál60, S. 38ff] durch folgendes lineares rekursives Gleichungssystem ausgedrückt

$$\underbrace{x_{t+1}}_{\text{Zustandsvektor}} = \underbrace{\Phi_{t+1}}_{\text{Übergangsmatrix}} \cdot \underbrace{x_t}_{\text{Zustandsvektor}} + \underbrace{\Xi_{t+1}}_{\text{Steuermatrix}} \cdot \underbrace{u_t}_{\substack{\text{exogener} \\ \text{Steuervektor}}} + \underbrace{w_t}_{\text{Rauschgröße}} ; \quad (2.1)$$

$$x_t \in \mathbb{R}^n, \Phi_t \in \mathbb{R}^{n \times n}, u_t \in \mathbb{R}^l, \Xi_t \in \mathbb{R}^{n \times l}, w_t \sim N(0, Q_t), Q_t \in \mathbb{R}^{n \times n}, \forall t \in \mathbb{N}_0$$

Dabei ist der n-dimensionale Zustandsvektor x_{t+1} zum Zeitpunkt $t+1$ eine durch die Übergangsmatrix Φ_{t+1} festgelegte Linearkombination des Zustandsvektors des vorherigen Zeitpunktes,

dem zusätzlich ein weißer Rauschvektor¹ w_t mit Kovarianzmatrix Q_t additiv überlagert wird. In manchen Fällen ist in der Systemmodellierung zusätzlich noch die Möglichkeit eines exogenen Kontrollinputs über die Steuermatrix Ξ_{t+1} und den Inputvektor u_t vorgesehen. Sofern in der vorliegenden Problemausprägung kein externer Kontrollinput vorgesehen ist, entfällt dieser Term, sodass sich Gleichung 2.1 vereinfacht zu:

$$\underbrace{x_{t+1}}_{\text{Zustandsvektor}} = \underbrace{\Phi_{t+1}}_{\text{Übergangsmatrix}} \cdot \underbrace{x_t}_{\text{Zustandsvektor}} + \underbrace{w_t}_{\text{Rauschgröße}} ; \quad (2.2)$$

$$x_t \in \mathbb{R}^n, \Phi_t \in \mathbb{R}^{n \times n}, w_t \sim N(0, Q_t), Q_t \in \mathbb{R}^{n \times n}, \forall t \in \mathbb{N}_0$$

Wenn man sich vergegenwärtigt, dass Φ_t und Q_t zwar mathematisch zeitvariant modelliert werden können, in der Praxis aber eine solche Modellierung bzw. Kalibrierung meist alles andere als trivial ist, weil man dafür ja eine sehr genaue Kenntnis der System- und Rauschdynamik haben müsste, so wird schnell klar, dass diese beiden Matrizen meist als zeitinvariant angenommen werden. Gleichung 2.2 lautet dann:

$$x_{t+1} = \Phi x_t + w_t; \quad (2.3)$$

$$x_t \in \mathbb{R}^n, \Phi \in \mathbb{R}^{n \times n}, w_t \sim N(0, Q), Q \in \mathbb{R}^{n \times n}, \forall t \in \mathbb{N}_0$$

2.2.2. Messmodell

Das Beobachtungs- bzw. Messmodell ist in Anlehnung an [Sor70, S. 10] durch folgendes Gleichungssystem gegeben:

$$\underbrace{z_t}_{\text{Messvektor}} = \underbrace{H_t}_{\text{Sensormatrix}} \cdot \underbrace{x_t}_{\text{Zustandsvektor}} + \underbrace{v_t}_{\text{Rauschgröße}} ; \quad (2.4)$$

$$z_t \in \mathbb{R}^m, x_t \in \mathbb{R}^n, H_t \in \mathbb{R}^{m \times n}, v_t \sim N(0, R_t), R_t \in \mathbb{R}^{m \times m}, \forall t \in \mathbb{N}_0$$

Auch hierbei gilt, dass die Sensormatrix H_t und die Rauschkovarianzmatrix R_t in der Praxis meist als zeitinvariant angenommen werden, weil sonst die genaue zeitliche Dynamik des Sensors und seines Rauschens bekannt sein müsste. Daraus ergibt sich folgendes vereinfachte Sensormodell:

$$z_t = H_t x_t + v_t; \quad (2.5)$$

$$z_t \in \mathbb{R}^m, x_t \in \mathbb{R}^n, H \in \mathbb{R}^{m \times n}, v_t \sim N(0, R), R \in \mathbb{R}^{m \times m}, \forall t \in \mathbb{N}_0$$

2.3. Zustandsschätzung

Die zu lösende Aufgabe, welche in Kapitel 2.1 bereits nicht-mathematisch veranschaulicht wurde, besteht nun darin, in rekursiver Weise zukünftige Zustände x_{t+1} auf Basis des vergangenen geschätzten Zustandes \hat{x}_t in Form von \hat{x}_{t+1} so zu schätzen, dass die dabei auftretenden Diskrepanzen mit den Messwerten z_t im Sinne der Methode der kleinsten Quadrate minimal werden, also eine Minimierung der Schätzfehlervarianzen vorgenommen wird [AM91]. Da dies, wie später noch zu sehen ist, ein sich periodisch wiederholender zweigeteilter Vorgang ist, gebietet es sich für die weiteren Ausführungen in der Notation zu reflektieren, ob man sich auf den Zeitpunkt vor Eintreffen eines Messwertes (a priori) oder nach Eintreffen zusätzlicher Information in Form

¹Prozess- bzw. Systemrauschen

eines Messwertes (a posteriori) bezieht [WB06, S.2]. In der Literatur (z.B. [Sor70]) wird dies oft durch tiefgestellte Zeitindizes der Form t/t gelöst, von denen sich der erste Index auf den Betrachtungszeitpunkt bezieht und der zweite Index auf den Zeitpunkt, von dem das zur Verfügung stehende Wissen stammt. Parallelen zur Bayesschen Schätztheorie sind hier offensichtlich und nicht rein zufällig, wie im letzten Teil dieses Kapitels aufgezeigt wird.

Zustandsschätzungen erfolgen in der Regel in zwei Schritten: 1) Vorhersage des nächsten Zustandes und 2) Korrektur der Schätzung auf Basis neu eingetroffener Messdaten. In Schritt 1) trifft man eine Vorhersage über den aktuellen Systemzustand in folgender Form:

$$\underbrace{\hat{x}_{t/t-1}}_{\substack{\text{a priori Schätzung} \\ \text{aktueller Zeitschritt}}} = \Phi \underbrace{\hat{x}_{t-1/t-1}}_{\substack{\text{a posteriori Schätzung} \\ \text{vorheriger Zeitschritt}}} \quad (2.6)$$

Interessant wird es nun im zweiten Schritt bei der Frage, wie man bei Eintreffen neuer Messwerte eine möglichst optimale Schätzung des Zustandes aus einer Kombination der alten Schätzung und der neu eingetroffenen Messwerte vornehmen kann. Mathematisch lässt sich die wie folgt formulieren [Sor70]:

$$\underbrace{\hat{x}_{t/t}}_{\substack{\text{a posteriori Schätzung} \\ \text{aktueller Zeitschritt}}} = \underbrace{\Phi}_{\substack{\text{Zeitinnovation}=\hat{x}_{t/t-1}}} \underbrace{\hat{x}_{t-1/t-1}}_{\substack{\text{a posteriori Schätzung} \\ \text{vorheriger Zeitschritt}}} + K_t \left(\overbrace{\left(\underbrace{\underbrace{z_t}_{\substack{\text{reale Messwerte} \\ \text{Messwert} \\ \text{aktueller Zeitschritt}}} - \underbrace{H \hat{x}_{t/t-1}}_{\substack{\text{Messwertvorhersage}=\hat{z}_t \\ \text{a priori Schätzung} \\ \text{aktueller Zeitschritt}}} \right)}^{\text{Messinnovation} = \text{Vorhersagefehler}} \right) \quad (2.7)$$

Die a posteriori Schätzung wird also dargestellt als Linearkombination der a priori Schätzung und der Diskrepanz zwischen hypothetischen Messwerten, basierend auf der a priori Schätzung und den neu eingetroffenen wirklichen Messwerten. Man kombiniert also die Schätzung des Zustandes, in den sich das System in der Zwischenzeit laut Systemmodell hinbewegt haben müsste (Zeitinnovation), mit der Diskrepanz zwischen dem, was man in diesem Fall hätte messen müssen und dem, was man für den aktuellen Zeitpunkt wirklich gemessen hat (Messinnovation). Die sogenannte Gain- oder Kalman-Matrix $K_t \in \mathbb{R}^{n \times m}$ ist dabei offensichtlich von zentraler Bedeutung. Die Spezifizierung einer optimalen Wahl für K_t ist einer der wissenschaftlichen Verdienste von R. Kalman. Optimal heißt in diesem Zusammenhang, dass der a posteriori Schätzfehler zwischen dem wahren und dem geschätzten Zustand $e_{t/t} = x_t - \hat{x}_{t/t}$ minimiert werden soll. Von Interesse sind dabei die a posteriori Fehlerkovarianzmatrix $P_{t/t}$ mit dem Erwartungswert

$$P_{t/t} = E[e_{t/t} e_{t/t}^T] \Leftrightarrow \text{cov}(x_t - \hat{x}_{t/t}) \quad (2.8)$$

und die a priori Fehlerkovarianzmatrix $P_{t/t-1}$ mit dem a priori Schätzfehler $e_{t/t-1} = x_t - \hat{x}_{t/t-1}$ und dem Erwartungswert

$$P_{t/t-1} = E[e_{t/t-1} e_{t/t-1}^T] \Leftrightarrow \text{cov}(x_t - \hat{x}_{t/t-1}) \quad (2.9)$$

Setzt man Gleichung (2.4) für z_t in Gleichung (2.7) ein, wendet Gleichung (2.6) auf den ersten Summanden in Gleichung (2.7) an und setzt den dadurch resultierenden Term für $\hat{x}_{t/t}$ in Gleichung (2.8) ein so erhält man

$$P_{t/t} = \text{cov}(x_t - (\hat{x}_{t/t-1} + K_t (H x_t + v_t - H \hat{x}_{t/t-1}))) \quad (2.10)$$

Diese Gleichung lässt sich unter Verwendung matrix-algebraischer Rechenregeln [Sor70, S. 10][Kál60, S. 39ff.][WB06, S. 3] weiterentwickeln zu:

$$\begin{aligned} P_{t/t} &= \text{cov} \left(x_t - \left(\hat{x}_{t/t-1} + K_t (Hx_t + v_t - H\hat{x}_{t/t-1}) \right) \right) \\ \Leftrightarrow P_{t/t} &= \text{cov} \left((I - K_t H) (x_t - \hat{x}_{t/t-1}) \right) + \text{cov} (K_t v_t) \\ \Leftrightarrow P_{t/t} &= (I - K_t H) \text{cov} (x_t - \hat{x}_{t/t-1}) (I - K_t H)^T + K_t \text{cov} (v_t) K_t^T \end{aligned} \quad (2.11)$$

Durch Anwendung von Gleichung (2.9) erhält man schlussendlich folgende sogenannte Joseph-Form² [SGA08, S. 136]:

$$P_{t/t} = (I - K_t H) P_{t/t-1} (I - K_t H)^T + K_t R K_t^T \quad (2.12)$$

Es lässt sich zeigen, dass die Minimierung des Erwartungswertes der mittleren quadratischen Abweichung des a posteriori Schätzfehlers $e_{t/t} = x_t - \hat{x}_{t/t}$ von der optimalen Wahl von K_t abhängt [Sor70, S. 10][Kál60, S. 39ff.][WB06, S. 3]. Unter Bezug auf Gleichung (2.12) gilt es dabei die Spur der Matrix $P_{t/t}$ zu minimieren. Durch ausmultiplizieren der Terme in Gleichung (2.12) und Bildung der Spur durch $\text{tr}(\cdot)$ erhält man:

$$\text{tr}(P_{t/t}) = \text{tr}(P_{t/t-1} - K_t H P_{t/t-1} - P_{t/t-1} K_t^T H^T + K_t (H P_{t/t-1} H^T + R) K_t^T) \quad (2.13)$$

Um diese Spur in Bezug auf die Wahl von K_t zu minimieren, wird die Matrixfunktion partiell nach K_t abgeleitet und gleich Null gesetzt. Es ergibt sich somit mittels der Ableitungsgesetze für Matrixfunktionen und der Rechenregeln bezüglich der Spurbildung von Matrizen folgender Zusammenhang:

$$\frac{\partial \text{tr}(P_{t/t})}{\partial K_t} = -2(H P_{t/t-1})^T + 2K_t (H P_{t/t-1} H^T + R) = 0 \quad (2.14)$$

Nach K_t aufgelöst folgt dann als optimale Wahl für K_t :

$$K_t = P_{t/t-1} H^T \underbrace{(H P_{t/t-1} H^T + R)^{-1}}_{\text{Residualkovarianz}} \quad (2.15)$$

mit

$$P_{t/t} = P_{t/t-1} - K_t H P_{t/t-1} \quad (2.16)$$

und

$$P_{t/t-1} = \Phi P_{t-1/t-1} \Phi^T + Q \quad (2.17)$$

Die Gleichungen 2.7 und 2.15-2.17 bezeichnet man in aller Regel als den Kalman-Filter. Interessant an Gleichung 2.15 ist, dass für eine gegen Null strebende Messfehlerkovarianzmatrix R , K_t gegen H^{-1} strebt und somit die Messinnovation gegenüber der Zeitinnovation übergewichtet wird. Umgekehrt gilt, dass wenn die a priori Fehlerkovarianzmatrix $P_{t/t-1}$ gegen Null strebt, K_t ebenfalls gegen Null strebt und somit die Messinnovation gegenüber der Zeitinnovation untergewichtet wird. Der Filter passt sich also sozusagen automatisch in seinem Gewichtungsverhalten an, je nachdem wie (un)zuverlässig seine vorherigen Schätzungen waren und je nachdem wie (un)zuverlässig das Sensorverhalten ist. Der Filter besitzt also einen eingebauten Rückkopplungssteuermechanismus.

²benannt nach Peter. D. Joseph

2.3.1. Kalman-Filter-Algorithmus

Der Kalman-Filter in seiner Form als mathematisches Gleichungssystem kann in eine praktische Rechenvorschrift als zweigeteilter rekursiv-zyklischer Algorithmus umgewandelt werden. Mit jedem Zeitschritt wird ein neuer Rekursionsschritt des Algorithmus vorgenommen. Wie in Abbildung 2.2 ersichtlich, wird in dem Zeit-Innovationsschritt eine a priori Vorhersage (Prädiktion) für den Systemzustand und die Fehlerkovarianzmatrix vorgenommen. In dem darauf folgenden Mess-Innovationsschritt wird nach Eintreffen neuer Messdaten eine a posteriori beobachtungs-basierte Bestimmung des Kalman-Gain sowie eine Korrektur der Zustands- und Fehlerkovarianzschätzung vorgenommen.

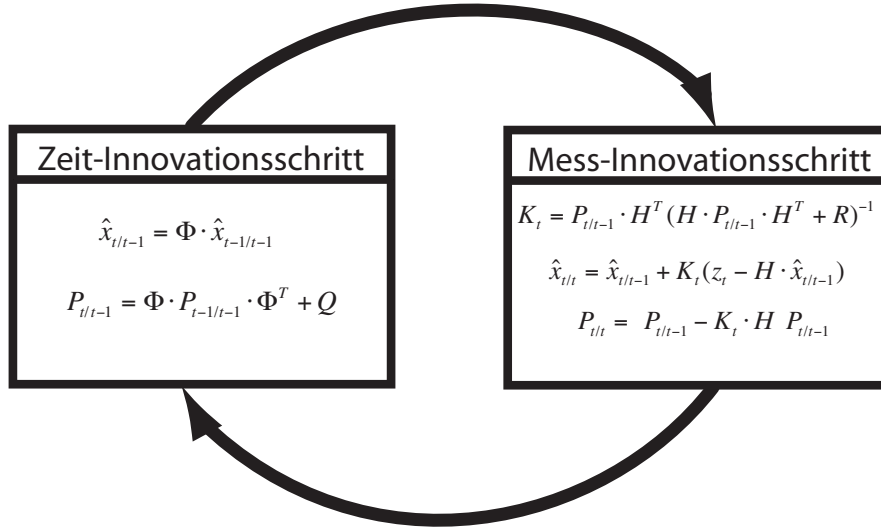


Abbildung 2.2.: Ablaufschema des Kalman-Filter-Algorithmus
In Anlehnung an [WB06, S. 24]

2.3.2. Bezug zur Bayesschen Schätztheorie

Der Kalman-Filter kann auch im Paradigma der Bayesschen Schätztheorie betrachtet werden und als Gaußscher Bayes-Filter zur Zustandsschätzung von Markovprozessen verstanden werden, wie im Folgenden aufgezeigt wird. Die Ausführungen basieren dabei auf [Idl05]. Die zu schätzende Wahrscheinlichkeitsdichte besteht aus

$$p(x_t \mid z_1, \dots, z_t),$$

wobei z_1, \dots, z_t die Sequenz aller bisherigen Sensordaten, ausgehend vom Zeitpunkt 1 bis zum aktuellen Zeitpunkt t , repräsentiere. Unter Anwendung der bayesschen Regel, lässt sich dieser Term ausdrücken durch

$$p(x_t \mid z_1, \dots, z_t) = \frac{p(z_t \mid x_t, z_1, \dots, z_{t-1}) p(x_t \mid z_1, \dots, z_{t-1})}{p(z_t \mid z_1, \dots, z_{t-1})}. \quad (2.18)$$

Durch die für den zu beschreibenden Prozess angenommene Markov-Umgebung erster Ordnung kann (2.18) umgeformt werden zu

$$p(x_t \mid z_1, \dots, z_t) = \frac{p(z_t \mid x_t) p(x_t \mid z_1, \dots, z_{t-1})}{p(z_t \mid z_1, \dots, z_{t-1})}. \quad (2.19)$$

und durch Anwendung des Satzes der totalen Wahrscheinlichkeit kann (2.19) in die eine rekursive Berechnung ermöglichende Gleichung

$$p(x_t | z_1, \dots, z_t) = \frac{p(z_t | x_t)}{p(z_t | z_1, \dots, z_{t-1})} \int p(x_t | x_{t-1}) p(x_{t-1} | z_1, \dots, z_{t-1}) dx_{t-1} \quad (2.20)$$

überführt werden. Sei mit η_t folgende in Abhängigkeit zum Sensormodell stehende Normalisierungskonstante bezeichnet [AMGC02, S. 175],

$$\eta_t^{-1} = p(z_t | z_1, \dots, z_{t-1}) = \int p(z_t | x_t) p(x_t | z_1, \dots, z_{t-1}) dx_t, \quad (2.21)$$

kann (2.20) umgeformt werden zu

$$p(x_t | z_1, \dots, z_t) = \eta_t \underbrace{p(z_t | x_t)}_{\text{Sensormodell}} \int \underbrace{p(x_t | x_{t-1})}_{\text{Systemmodell}} p(x_{t-1} | z_1, \dots, z_{t-1}) dx_{t-1}. \quad (2.22)$$

Dies ist die im Allgemeinen als **Bayesfilter** bezeichnete Gleichung zur rekursiven Abschätzung des aktuellen Systemzustandes, basierend auf allen bisherigen Messdaten, in Form einer bedingten Wahrscheinlichkeitsdichte über dem Zustandsraum. Der Bezug zum Kalman-Filter wird klar, wenn man sich verdeutlicht, dass eine algorithmische Realisierung der Bayesfiltergleichung (2.22) in der Regel ebenfalls in zwei Schritten abläuft:

1. Vorhersage: $p(x_t | z_1, \dots, z_{t-1}) = \int p(x_t | x_{t-1}) p(x_{t-1} | z_1, \dots, z_{t-1}) dx_{t-1}$
2. Aktualisierung: $p(x_t | z_1, \dots, z_t) = \eta_t p(z_t | x_t) p(x_t | z_1, \dots, z_{t-1})$

In einem Prädiktionsschritt (Vorhersage) wird das Systemmodell verwendet, um basierend auf der sogenannten Chapman-Kolmogorov-Gleichung (2.23) eine Voraussage darüber zu treffen, in welchem Zustand sich das System im nächsten Zeitschritt vermutlich befinden wird. Unter der erwähnten Chapman-Kolmogorov-Gleichung versteht man dabei folgende Integralgleichung

$$p(x_t | z_1, \dots, z_{t-1}) = \int p(x_t | x_{t-1}) p(x_{t-1} | z_1, \dots, z_{t-1}) dx_{t-1}, \quad (2.23)$$

die auf Basis des Systemmodells die **a posteriori** Wahrscheinlichkeitsdichte $p(x_{t-1} | z_1, \dots, z_{t-1})$ aus dem Zeitschritt $t-1$ in eine Randverteilung überführt, welche gleich der **a priori** Wahrscheinlichkeitsdichte $p(x_t | z_1, \dots, z_{t-1})$ zum Zeitpunkt t ist. Die Gleichung entspricht somit genau dem Integralterm des Bayesfilters aus (2.22). Der Term $p(x_t | x_{t-1})$ repräsentiert das System- bzw. Prozessmodell [Sch04, S. 35]. Beim Eintreffen neuer Sensordaten wird im zweiten Schritt des Bayesfilters (Aktualisierung) auf Basis der prädiktierten a priori Wahrscheinlichkeitsdichte $p(x_t | z_1, \dots, z_{t-1})$ zum Zeitpunkt t , die a posteriori Wahrscheinlichkeit $p(x_t | z_1, \dots, z_t)$ zum Zeitpunkt t durch folgende Gleichung bestimmt

$$p(x_t | z_1, \dots, z_t) = \eta_t p(z_t | x_t) p(x_t | z_1, \dots, z_{t-1}). \quad (2.24)$$

Der Term $p(z_t | x_t)$ modelliert das Sensor- bzw. Beobachtungsmodell und wird als Beobachtungswahrscheinlichkeit bezeichnet. Die Interpretation des Kalman-Filters als Gaußscher Bayesfilter lässt sich prägnant nochmal durch folgendes Gleichungssystem aufzeigen:

$$\begin{aligned} p(x_t | x_{t-1}) &\sim N(\Phi x_{t-1}, Q) \\ p(z_t | x_t) &\sim N(H x_t, R) \\ p(x_{t-1} | z_1, \dots, z_{t-1}) &\sim N(\hat{x}_{t-1/t-1}, P_{t-1/t-1}) \end{aligned} \quad (2.25)$$

2.4. Likelihoodfunktion im Kalman-Filter Kontext

Im Hinblick auf den im nächsten Kapitel folgenden praktischen Teil der Arbeit wird in diesem Abschnitt kurz darauf eingegangen, wie eine Likelihoodfunktion als Beiprodukt aus den in dem Kalmanfilter berechneten Matrizen definiert werden kann. Von besonderem Interesse dafür sind die Messinnovation, die im Folgenden mit ζ_t bezeichnet sei und die Residualkovarianzmatrix, welche durch Σ_t symbolisiert sei:

$$\begin{aligned}\zeta_t &= z_t - H\hat{x}_{t/t-1} \\ \Sigma_t &= HP_{t/t-1}H^T + R\end{aligned}\tag{2.26}$$

Mit Hilfe dieser beiden Terme lässt sich nach [Bol01, S. 34ff] und [DS04, S. 9ff] folgende Likelihoodfunktion aufstellen [Maz07, S. 83ff.]:

$$\Lambda_t = \Lambda_{t-1} - n \log(2\pi)2^{-1} - \frac{1}{2} (\log |\Sigma_t| + \zeta_t^T \Sigma_t^{-1} \zeta_t)\tag{2.27}$$

und daraus ein aggregierter Likelihoodwert für den Zeitraum $t = 1$ bis T ableiten durch:

$$\Lambda_T = \sum_{t=1}^T \Lambda_t\tag{2.28}$$

Wie in Kapitel 3.3 noch gezeigt wird, spielt diese Likelihoodfunktion eine bedeutende Rolle, wenn man neben den Systemzuständen auch noch eventuelle Modellparameter abschätzen muss bzw. von verschiedenen zur Auswahl stehenden Modellen das am besten geeignete aussuchen muss.

3. Ökonometrisches Anwendungsbeispiel

Im Folgenden wird der Einsatz des Kalmanfilters im Kontext der Parameterbestimmung ausgewählter affiner Zinsstrukturmodelle vorgestellt [DK96, CIR85, Vas77]. Die Ausführungen basieren überwiegend auf [Bol01, Idl11]. Es werden fundierte Kenntnisse der allgemeinen Problematik der Zinsstrukturmodellierung vorausgesetzt und aus Platzgründen auf eine Einführung in das Problemfeld an dieser Stelle verzichtet und stattdessen auf die einschlägige Literatur [BM07, Rud00, Wil07, AP10b, AP10a, Hug03, Reb04] und [Idl11] verwiesen.

3.1. Datenbasis

Als Datenbasis für die folgenden Betrachtungen wird auf historische LIBOR-Geldmarktzinsdaten¹ zurückgegriffen, welche täglich von der British Bankers' Association (BBA)² veröffentlicht werden und im Internet-Archiv³ abrufbar sind. Die Untersuchungen in dieser Arbeit wurden auf den Zeitraum 2000 bis 2008 und die Laufzeiten 1-12 Monate eingeschränkt. Bei den LIBOR-

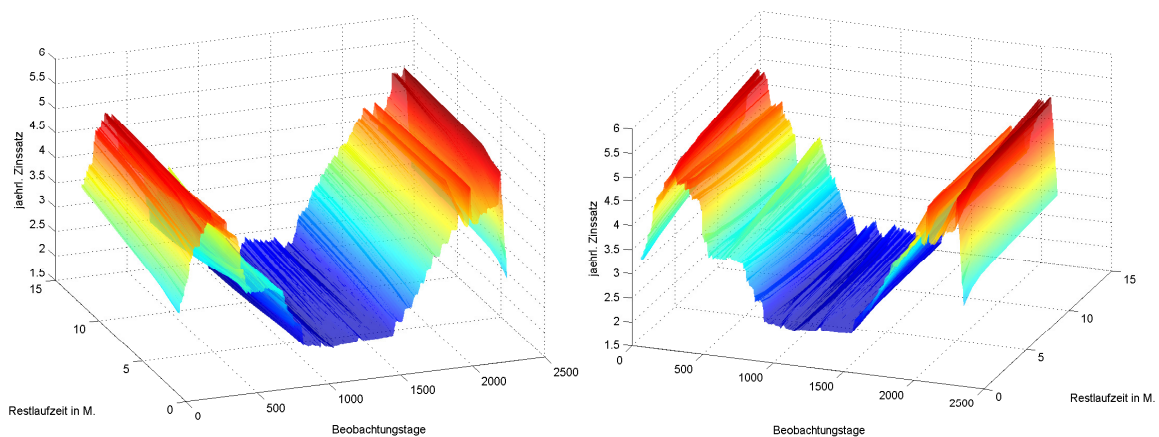


Abbildung 3.1.: geglättete Zinsstruktur €-LIBOR 2000-2008, Tageswerte, 1-12 Monate Laufzeit

Zinsdaten aus Abbildung 3.1 handelt es sich um jährliche Zinssätze für feste diskrete Laufzeiten τ . Symbolisieren $L(\tau)$; $\tau \in \frac{1}{12}, \dots, \frac{12}{12}$ die diskreten LIBOR-Sätze für die Laufzeiten τ von einem bis 12 Monaten⁴, dann können daraus implizierte Kurse $P(\tau)$ für hypothetische ausfallrisikofreie Nullkuponanleihen mit korrespondierenden Laufzeiten wie folgt berechnet werden [Cha04, S. 16ff.]:

$$P(\tau) = (1 + L(\tau)\tau)^{-1}; \forall \tau \in \frac{1}{12}, \dots, \frac{12}{12} \quad (3.1)$$

¹London Interbank Offered Rate; Berechnung: <http://www.bbalibor.com/bbalibor-explained/the-basics>

²Pinnars Hall, 105-108 Old Broad Street, London EC2N 1EX, United Kingdom

³<http://web.archive.org/web/20090413091623/http://www.bba.org.uk/bba/jsp/polopoly.jsp?d=141&a=627>

⁴ausgedrückt in Jahresbruchteilen

Für obige Berechnung sei ein Jahr, in Übereinstimmung mit der LIBOR-Berechnungskonvention⁵, auf 360 Tage normiert. Ein Monat sei für die Zwecke dieser Arbeit aus Vereinfachungsgründen auf 30 Tage normiert. Die Berechnungsergebnisse aus Gleichung (3.1) stellen die Zinsstruktur in Form von Diskontfaktoren dar. Aus ihnen lassen sich die laufzeitabhängigen Kassazinsraten⁶ $z(\tau)$ wie folgt berechnen:

$$z(\tau) = -\frac{\ln P(\tau)}{\tau} = \frac{\ln(1 + L(\tau)\tau)}{\tau} \quad (3.2)$$

Die instantane Kassazinsrate⁷ $r(t)$ zum Zeitpunkt t lässt sich aus $z(\tau)$ durch Grenzwertbildung ableiten:

$$r(t) = \lim_{\tau \rightarrow 0} z(\tau) = \lim_{\tau \rightarrow 0} -\frac{\ln P(\tau)}{\tau} \quad (3.3)$$

In dem Paradigma der affinen Zinsstrukturmodelle wird ihr zeitlicher Verlauf im Einfaktorfall durch folgende stochastische Differentialgleichung beschrieben:

$$dr(t) = \underbrace{A_0 dt}_{\text{Drift}} + \underbrace{A_1 dW(t)}_{\text{Diffusion}} \quad (3.4)$$

Zwei populäre Ausprägungen dieser Modellklasse [Bol01, S. 16], welche den short-rate-Prozess als Ornstein-Uhlenbeck-Prozess modellieren, bestehen in dem Zinsstrukturmodell nach [Vas77] mit

$$\begin{aligned} A_0 &= \kappa(\bar{\theta} - r(t)); \\ A_1 &= \sigma; \\ \bar{\theta} &= \theta - \frac{\sigma\lambda}{\kappa} \end{aligned} \quad (3.5)$$

und dem nach [CIR85] mit

$$\begin{aligned} A_0 &= \kappa(\bar{\theta} - r(t)); \\ A_1 &= \sigma\sqrt{r(t)}; \\ \bar{\theta} &= \theta - \frac{\sigma\lambda}{\kappa} \end{aligned} \quad (3.6)$$

Dabei bezeichnet θ einen langfristigen mittleren Zinssatz, κ den Mittelwertrückkehrparameter, σ die Volatilität und λ den Marktpreis des Risikos der ungewissen zukünftigen Zinsentwicklung.

3.2. Zinsstrukturmodell

3.2.1. Ein-Faktor-Fall

Dem hier betrachteten Paradigma der affinen Zinsstrukturmodelle nach [DK96] liegt die Annahme zu Grunde, dass die Nullkuponanleihenpreisfunktion eine lineare Funktion der ihr zugrunde liegenden Zustandsvariablen ist, welche die Unsicherheit bzgl. der zukünftigen Zinsentwicklung

⁵<http://www.bbalibor.com/technical-aspects/calculating-interest>

⁶(continuously compounded) spot rates

⁷short rate

bzw. -dynamik repräsentiert [Bol01, S. 4]. Die Preisfunktion wird dabei allgemein wie folgt ausgedrückt:

$$P(\tau) = e^{A(\tau) - B(\tau)r} \quad (3.7)$$

Bei $A(t)$ und $B(t)$ muss es sich dabei um Funktionen handeln, die (wenn überhaupt) nur deterministisch von der Zeit abhängen. Bei dem Vasicek-Modell [Vas77] ergeben sich diese Funktionen laut Herleitung in [Bol01, S. 16ff] aus Gleichung 3.5 wie folgt:

$$\begin{aligned} B(\tau) &= \frac{1}{\kappa}(1 - e^{-\kappa\tau}); \\ A(\tau) &= \frac{\gamma(B(\tau) - \tau)}{\kappa^2} - \frac{\sigma^2 B^2(\tau)}{4\kappa}; \\ \gamma &= \kappa^2 \left(\theta - \frac{\sigma\lambda}{\kappa} \right) - \frac{\sigma^2}{2} \end{aligned} \quad (3.8)$$

Für das CIR-Modell nach [CIR85] ergibt sich laut [Bol01, S. 42ff] unter Verwendung von Gleichung 3.6 folgender Zusammenhang:

$$\begin{aligned} B(\tau) &= \frac{2(e^{\gamma\tau} - 1)}{(\gamma + \kappa + \lambda)(e^{\gamma\tau} - 1) + 2\gamma}; \\ A(\tau) &= \ln \left(\frac{2\gamma e^{\frac{(\gamma + \kappa + \lambda)\tau}{2}}}{(\gamma + \kappa + \lambda)(e^{\gamma\tau} - 1) + 2\gamma} \right)^{\frac{2\kappa\theta}{\sigma^2}}; \\ \gamma &= \sqrt{(\kappa + \lambda)^2 + 2\sigma^2} \end{aligned} \quad (3.9)$$

3.2.2. Multi-Faktor-Fall

Erweitert man die klassischen Einfaktoransätze von [Vas77] und [CIR85] um weitere Faktoren, welche derselben Modellierung folgen, so kann man die instantane Short-Rate r als eine Linearkombination von n möglicherweise korrelierter Zustandsvariablen x_1, \dots, x_n modellieren. Konkret wird in [Bol01, S. 18ff] für das Vasicek Modell folgender Mehrfaktorzusammenhang vorgestellt

$$\begin{aligned} P(\tau) &= e^{A(\tau) - \sum_{j=1}^n B_j(\tau)x_j} \\ B_j(\tau) &= \frac{1}{\kappa_j}(1 - e^{-\kappa_j\tau}); \\ A(\tau) &= \sum_{j=1}^n \frac{\gamma_j(B_j(\tau) - \tau)}{\kappa_j^2} - \frac{\sigma_j^2 B_j^2(\tau)}{4\kappa_j} \\ &\quad + \sum_{i,j:i \neq j} \frac{\sigma_{ij}}{2\kappa_i\kappa_j} \left(\tau - B_i(\tau) - B_j(\tau) + \frac{1}{\kappa_i + \kappa_j} (1 - e^{-(\kappa_i + \kappa_j)\tau}) \right); \\ \gamma_j &= \kappa_j^2 \left(\theta_j - \frac{\sigma_j\lambda_j}{\kappa_j} \right) - \frac{\sigma_j^2}{2} \end{aligned} \quad (3.10)$$

und für das CIR Modell entsprechend:

$$\begin{aligned}
 P(\tau) &= e^{\sum_{j=1}^n A_j(\tau) - B_j(\tau)x_j} \\
 B_j(\tau) &= \frac{2(e^{\gamma_j \tau} - 1)}{(\gamma_j + \kappa_j + \lambda_j)(e^{\gamma_j \tau} - 1) + 2\gamma_j}; \\
 A_j(\tau) &= \ln \left(\frac{2\gamma_j e^{\frac{(\gamma_j + \kappa_j + \lambda_j)\tau}{2}}}{(\gamma_j + \kappa_j + \lambda_j)(e^{\gamma_j \tau} - 1) + 2\gamma_j} \right)^{\frac{2\kappa_j \theta_j}{\sigma_j^2}}; \\
 \gamma_j &= \sqrt{(\kappa_j + \lambda_j)^2 + 2\sigma_j^2}
 \end{aligned} \tag{3.11}$$

Ein exemplarisches mehrdimensionales Beobachtungs- und Systemmodell für das Vasicek Modell aus Gleichung (3.10) lässt sich laut [Bol01, S. 29ff] unter Annahme von 3 Zustandsvariablen ($x_t \in \mathbb{R}^3$) und einer Datenbasis von 12 unterschiedlichen Monatslaufzeiten τ_1, \dots, τ_{12} ($z_t \in \mathbb{R}^{12}$) wie folgt ableiten:

$$z_t = \begin{bmatrix} -\frac{A(\tau_1)}{\tau_1} \\ \vdots \\ -\frac{A(\tau_{12})}{\tau_{12}} \end{bmatrix} + \begin{bmatrix} \frac{B_1(\tau_1)}{\tau_1} & \frac{B_2(\tau_1)}{\tau_1} & \frac{B_3(\tau_1)}{\tau_1} \\ \vdots & \vdots & \vdots \\ \frac{B_1(\tau_{12})}{\tau_{12}} & \frac{B_2(\tau_{12})}{\tau_{12}} & \frac{B_3(\tau_{12})}{\tau_{12}} \end{bmatrix} x_t + v_t \tag{3.12}$$

$$x_t = \begin{bmatrix} \theta_1(1 - e^{-\kappa_1 \Delta t}) \\ \theta_2(1 - e^{-\kappa_2 \Delta t}) \\ \theta_3(1 - e^{-\kappa_3 \Delta t}) \end{bmatrix} + \begin{bmatrix} e^{-\kappa_1 \Delta t} & 0 & 0 \\ 0 & e^{-\kappa_2 \Delta t} & 0 \\ 0 & 0 & e^{-\kappa_3 \Delta t} \end{bmatrix} x_{t-1} + w_t \tag{3.13}$$

Dabei kann das Messrauschen des Beobachtungsmodells konventionell in einer Diagonalmatrix R modelliert werden, jedoch ergibt sich für das Systemrauschen folgender etwas komplexerer Zusammenhang:

$$w_t \sim N(0, Q); Q = \begin{bmatrix} \frac{\sigma_1^2}{2\kappa_1}(1 - e^{-2\kappa_1 \Delta t}) & 0 & 0 \\ 0 & \frac{\sigma_2^2}{2\kappa_2}(1 - e^{-2\kappa_2 \Delta t}) & 0 \\ 0 & 0 & \frac{\sigma_3^2}{2\kappa_3}(1 - e^{-2\kappa_3 \Delta t}) \end{bmatrix} \tag{3.14}$$

Analog kann man folgende Modelle für das CIR Modell aufstellen:

$$z_t = \begin{bmatrix} \sum_{j=1}^3 -\frac{A_j(\tau_1)}{\tau_1} \\ \vdots \\ \sum_{j=1}^3 -\frac{A_j(\tau_{12})}{\tau_{12}} \end{bmatrix} + \begin{bmatrix} \frac{B_1(\tau_1)}{\tau_1} & \frac{B_2(\tau_1)}{\tau_1} & \frac{B_3(\tau_1)}{\tau_1} \\ \vdots & \vdots & \vdots \\ \frac{B_1(\tau_{12})}{\tau_{12}} & \frac{B_2(\tau_{12})}{\tau_{12}} & \frac{B_3(\tau_{12})}{\tau_{12}} \end{bmatrix} x_t + v_t \tag{3.15}$$

$$x_t = \begin{bmatrix} \theta_1(1 - e^{-\kappa_1 \Delta t}) \\ \theta_2(1 - e^{-\kappa_2 \Delta t}) \\ \theta_3(1 - e^{-\kappa_3 \Delta t}) \end{bmatrix} + \begin{bmatrix} e^{-\kappa_1 \Delta t} & 0 & 0 \\ 0 & e^{-\kappa_2 \Delta t} & 0 \\ 0 & 0 & e^{-\kappa_3 \Delta t} \end{bmatrix} x_{t-1} + w_t \tag{3.16}$$

Auch hier kann das Messrauschen des Beobachtungsmodells konventionell in einer Diagonalmatrix R modelliert werden, das Systemrauschen ist jedoch bedeutend komplexer als im Vasicek Fall:

$$w_t \sim N(0, Q); Q_t = \begin{bmatrix} \psi_1 & 0 & 0 \\ 0 & \psi_2 & 0 \\ 0 & 0 & \psi_3 \end{bmatrix} x_{t-1} \tag{3.17}$$

mit

$$\psi_j = \frac{\theta_j \sigma_j^2}{2\kappa_j} (1 - e^{-\kappa_j \Delta t})^2 + \frac{\sigma_j^2}{\kappa_j} (e^{-\kappa_j \Delta t} - e^{-2\kappa_j \Delta t}) \tag{3.18}$$

3.3. Parameterschätzung

In der im Anhang A dargestellten praktischen Implementierung der bisher vorgestellten Konzepte, wird dem in Abschnitt 3.1 beschriebenen LIBOR-Datensatz exemplarisch ein eindimensionales CIR-Modell durch Parameterschätzung angepasst. Dabei wird der Kalman-Filter im Geiste des expectation-maximization (EM) Algorithmus [DLR77] zur wiederholten rekursiven Berechnung der Likelihoodwerte für einzelne Parameterhypothesen herangezogen. Die eigentliche Optimierung der Parameterwerte wird in der dieser Arbeit zugrunde liegenden Implementierung (siehe S.21) durch Innere-Punkte-Verfahren (z.B. [WMNO06]) durchgeführt. Das Berechnungsergebnis nach erfolgreichem Durchlaufen des Algorithmus besteht im Idealfalle⁸ aus einer maximum a posteriori (MAP) Schätzung der Parameterwerte für $\theta, \kappa, \sigma, \lambda$ bei der definitionsgemäss der aggregierte Likelihoodwert Λ_T aus Gleichung (2.28) maximal wird

$$\operatorname{argmax}_{\theta, \kappa, \sigma, \lambda} \Lambda_T(\theta, \kappa, \sigma, \lambda) \quad (3.19)$$

Für den vorliegenden Datensatz und die vorliegende Implementierung ergeben sich die folgende aus Tabelle 3.1 ersichtliche maximum a posteriori Schätzung der Parameter.

θ	κ	σ	λ	Λ_T
0.0324	0.0981	0.0189	-0.0903	-1.627*10 ⁵

Tabelle 3.1.: MAP Parameterschätzung für 1-Faktor-CIR-Modell, €-LIBOR 2000-2008, Tageswerte, 1-12 Monate Laufzeit

Die Rauschmatrix des Beobachtungsmodelles wird in der vorliegenden Implementierung ebenfalls als Teil der zu schätzenden bzw. zu optimierenden Parameter modelliert. Als Schätzung ergibt sich untige 12×12 Diagonalmatrix

$$R = \operatorname{diag} [0.24, 0.17, 0.12, 0.09, 0.05, 0.02, 0.00, 0.02, 0.05, 0.06, 0.08, 0.10] * 10^{-2} \quad (3.20)$$

Abschließend sei der Vollständigkeit halber angemerkt, dass ein alternativer Ansatz zur Parameterschätzung im Kontext des Kalmanfilters darin besteht, die zu schätzenden Parameter explizit in einem erweiterten Zustandsraum zu modellieren und durch den Kalman-Filter gemeinsam mit den ursprünglichen Zuständen mitschätzen zu lassen [Maz07, S. 83ff.]. Auf diese Vorgehensweise soll jedoch nicht weiter eingegangen werden.

⁸je nach verwendetem Abbruchkriterium

Literaturverzeichnis

- [AM91] B. D. O. Anderson and J. B. Moore. *Kalman filtering: Whence, what and whither?*, volume 19, pages 42–54. Springer Berlin Heidelberg, 1991. Available from: users.rsise.anu.edu.au/~john/papers/CHAPTERS/02.PDF.
- [AMGC02] Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp. A Tutorial on Particle Filters for On-line Non-linear/Non-Gaussian Bayesian Tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, February 2002. Available from: <http://www.cs.ucsb.edu/~290stat/Papers/ParticleFilterTutorial.pdf>.
- [AP10a] L. Andersen and V. Piterbarg. *Interest Rate Modeling. Volume 1: Foundations and Vanilla Models*. Atlantic Financial Press, 195 Bickenhall Mansions, Bickenhall Street, London W1U 6BX, UK, first edition, 2010.
- [AP10b] L. Andersen and V. Piterbarg. *Interest Rate Modeling. Volume 2: Term Structure Models*. Atlantic Financial Press, 195 Bickenhall Mansions, Bickenhall Street, London W1U 6BX, UK, first edition, 2010.
- [BM07] Damiano Brigo and Fabio Mercurio. *Interest Rate Models - Theory and Practice*. Springer, Berlin, Heidelberg, New York, third edition, 2007.
- [Bol01] David Jamieson Bolder. Affine Term-Structure Models: Theory and Implementation. Working Papers 15, Bank of Canada, 2001. Available from: <http://www.bankofcanada.ca/en/res/wp/2001/wp01-15a.pdf>.
- [Cha04] Don M. Chance. Another look at the forward-futures price differential in libor markets - Working Paper. Technical report, Louisiana State University, Department of Finance, 2163 Patrick F. Taylor Hall, Baton Rouge, LA 70803, USA, June 2004. Available from: <http://www.bus.lsu.edu/academics/finance/research/working%20papers/AnotherLookattheFFDiff.pdf>.
- [CIR85] John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2):385–407, March 1985. Available from: <http://www.javaquant.net/papers/CIR1985.pdf>.
- [DK96] D. Duffie and R. Kan. A YIELD-FACTOR MODEL OF INTEREST RATES. *Mathematical Finance*, 6:379–406, Oct 1996. Available from: www.darrellduffie.com/uploads/pubs/DuffieKan1996.pdf.
- [DLR77] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977. Available from: <http://web.mit.edu/6.435/www/Dempster77.pdf>.

- [DS04] Gregory R. Duffee and Richard H. Stanton. Estimation of Dynamic Term Structure Models - Working Paper. Technical report, Haas School of Business, U.C. Berkeley, 545 Student Services Building 1900, Berkeley, CA 94720-1900, USA, March 2004. Available from: <http://faculty.haas.berkeley.edu/stanton/papers/pdf/affine.pdf>.
- [Gau09] Carl Friedrich Gauss. *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*. Friedrich Perthes and I.H. Besser, Hamburg, 1809. Available from: http://books.google.com.au/books/download/Theoria_motus_corporum_coelestium_in_sec.pdf?id=ORUOAAAAQAAJ&output=pdf&sig=ACfU3U1__v7ZeP-_3-pGGqg1cQBCf00GTw.
- [Hug03] Lane Hughston. The Past, Present and Future of Term Structure Modelling. In Peter Field, editor, *Modern Risk Management: A History*, chapter 7, pages 107–132. Risk Books, Haymarket House, 28-29 Haymarket, London, SW1Y 4RX, UK, 2003. Available from: http://www.mth.kcl.ac.uk/finmath/articles/LPH_risk.pdf.
- [Idl05] Corvin Idler. Visuelle probabilistische Mehrobjektverfolgung im Rahmen der Fahrzeugumfeldbeobachtung. Master's thesis, Fakultät für Informatik der Universität Koblenz-Landau, Jul 2005. Available from: <http://userpages.uni-koblenz.de/~idler/da-00.pdf>.
- [Idl11] Corvin Idler. Zinsstrukturmodelle im HJM-Framework, Bachelor's Thesis, Fakultät für Wirtschaftswissenschaft der Fern-Universität Hagen, Sep 2011. Available from: <http://userpages.uni-koblenz.de/~idler/bachelor.pdf>.
- [Kál60] Rudolph Emil Kálmán. A new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering*, 82(Series D):35–45, 1960. Available from: www.cs.unc.edu/~welch/kalman/media/pdf/Kalman1960.pdf.
- [KB61] R. E. Kálmán and R. S. Bucy. New results in linear filtering and prediction theory. *Transactions of the ASME. Series D, Journal of Basic Engineering*, 83:95–107, 1961.
- [Leg05] Adrien Marie Legendre. *Nouvelles méthodes pour la détermination des orbites des comètes*. F. Didot, Paris, 1805. Available from: http://books.google.com.au/books/download/Nouvelles_m%C3%A9thodes_pour_la_d%C3%A9terminati.pdf?id=spcAAAAAMAAJ&output=pdf&sig=ACfU3U0QLegNSZYLLme09hI7nrHxXbZtGg.
- [May79] Peter S. Maybeck. *Stochastic models, estimation, and control, Vol.1*, volume 141 of *Mathematics in Science and Engineering*. Academic Press Inc., 111 Fifth Avenue, 10003 New York, USA, 1979. Available from: http://www.cs.unc.edu/~welch/kalman/media/pdf/maybeck_ch1.pdf.
- [Maz07] Thomas Mazzoni. *Stetig/diskrete Zustandsraummodelle dynamischer Wirtschaftsprozesse*. Shaker Verlag, Aachen, 1 edition, 2007.

- [O'R96] Randall C. O'Reilly. *The Leabra Model of Neural Interactions and Learning in the Neocortex*. PhD thesis, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA, 1996. Available from: ftp://grey.colorado.edu/pub/oreilly/thesis/oreilly_thesis.all.pdf.
- [Reb04] Riccardo Rebonato. Interest-Rate Term-Structure Pricing Models: A Review. *Proceedings of the Royal Society: Mathematical, Physical and Engineering Sciences*, 460(2043):667–728, 2004. Available from: <http://www.riccardorebonato.co.uk/papers/IRSurvey.pdf>.
- [Rud00] Markus Rudolf. *Zinsstrukturmodelle*. Physica-Verlag, Tiergartenstraße 17, 69121 Heidelberg, D, first edition, 2000.
- [Sch04] Thorsten Schmitt. *Vision-based Probabilistic State Estimation for Cooperating Autonomous Robots*. PhD thesis, Technische Universität München, Fakultät für Informatik, Universitätsbibliothek, Arcisstr. 21, D-80333 München, May 2004. Available from: <http://tumb1.biblio.tu-muenchen.de/publ/diss/in/2004/schmitt.html>.
- [SGA08] Mohinder S. Grewal and Angus P. Andrews. *Kalman Filtering: Theory and Practice Using MATLAB*. John Wiley & Sons, Inc., Hoboken, New Jersey, third edition, 2008.
- [SM85] Stanley F. Schmidt and Leonard A. McGee. Discovery of the Kalman Filter as a Practical Tool for Aerospace and Industry. Technical Memorandum NASA-TM-86847, National Aeronautics and Space Administration, November 1985. Available from: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19860003843_1986003843.pdf.
- [Sor70] Harold Walter Sorenson. Least-squares estimation: from gauss to kalman. *Spectrum, IEEE*, 7:63–68, Jul 1970. Available from: www.cs.unc.edu/~welch/kalman/media/pdf/Sorenson1970.pdf.
- [Swe58] Peter Swerling. A proposed stagewise differential correction procedure for satellite tracking and prediction. Rand Corporation Report P-1292, RAND, Santa Monica, Calif, USA, Jan 1958.
- [Swe59] Peter Swerling. A proposed stagewise differential correction procedure for satellite tracking and prediction. *Journal of Astronautical Sciences*, 6:46–52, 1959.
- [Vas77] Oldrich Vasicek. An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 5(2):177–188, November 1977. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.164.447&rep=rep1&type=pdf>.
- [WB06] Greg Welch and Gary Bishop. An introduction to the kalman filter. Technical Report TR-95-041, Department of Computer Science, University of North Carolina at Chapel Hill, July 2006. Available from: www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf.

- [Wil07] Paul Wilmott. *Paul Wilmott Introduces Quantitative Finance*. John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, UK, second edition, 2007.
- [WMNO06] R. A. Waltz, J. L. Morales, J. Nocedal, and D. Orban. An interior algorithm for nonlinear optimization that combines line search and trust region steps. Technical report, Mathematical Programming 107, 2006. Available from: <http://www.ece.northwestern.edu/~nocedal/PDFfiles/directpaper.pdf>.

A. Anhang

```
% Copyright (c) 2010, Bill, All rights reserved.
% Redistribution and use in source and binary forms, with or without modification,
% are permitted provided that the following conditions are met:
% * Redistributions of source code must retain the above copyright notice, this list
% of conditions and the following disclaimer.
% * Redistributions in binary form must reproduce the above copyright notice, this
% list of conditions and the following disclaimer in the documentation and/or other
% materials provided with the distribution
%
% THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS AS IS AND ANY
% EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES
% OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT
% SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT,
% INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED
% TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS;
% OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN
% CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN
% ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH
% DAMAGE.

function [para, sumll] = TreasuryYieldKF()
% author: biao from www.mathfinance.cn
% http://www.mathfinance.cn/kalman-filter-finance-revisited
% amended by Corvin Idler Idler@uni-koblenz.de

%%CIR parameter estimation using Kalman Filter for given treasury bonds yields
% check paper "estimating and testing exponential-affine term structure
% models by kalman filter " and "affine term structure models: theory and
% implementation" for detail;  $S(t+1) = \mu + F S(t) + \text{noise}(Q)$ 
%  $Y(t) = A + H S(t) + \text{noise}(R)$ 

all=importdata('allmonthlyLibor9908daily.mat');
L = (all/100)'; size(L);
tau = [1/12 2/12 3/12 4/12 5/12 6/12 7/12 8/12 9/12 10/12 11/12 12/12];

for a=1:size(L,1),
L(a,:) = (log(1+(L(a,:).*tau)))./tau;
end

Y= L; [nrow, ncol] = size(Y);
options = optimset('Algorithm','interior-point');
para0 = [0.05, 0.01, 0.01, -0.01, std((Y(1:end-1,:)-Y(2:end,:)))];
[x, fval] = fmincon(@loglik, para0,[],[],[],[],[0.0001,0.0001,0.0001, -1,
0.00001*ones(1,ncol)],[ones(1,length(para0))],[],options,Y, tau, nrow, ncol);

para = x
sumll = fval
```

```

function sumll = loglik(para,Y, tau, nrow, ncol)
% author: biao from www.mathfinance.cn
% modified by Corvin Idler idler@uni-koblenz.de

% initialize the parameter for CIR model
theta = para(1); kappa = para(2); sigma = para(3); lambda = para(4);
sigmai = para(5:end);
R = eye(ncol);
for i = 1:ncol
    R(i,i) = sigmai(i)^2;
end
dt = 1/256; %daily data
initx = theta;
initV = sigma^2*theta/(2*kappa);

% parameter setting for transition equation
mu = theta*(1-exp(-kappa*dt));
F = exp(-kappa*dt);

% parameter setting for measurement equation
A = zeros(1, ncol);
H = A;
for i = 1:ncol
    AffineGamma = sqrt((kappa+lambda)^2+2*sigma^2);
    AffineBeta = 2*(exp(AffineGamma*tau(i))-1)/((AffineGamma+kappa+lambda)*
(exp(AffineGamma*tau(i))-1)+2*AffineGamma);
    AffineAlpha = 2*kappa*theta/(sigma^2)*log(2*AffineGamma*
exp((AffineGamma+kappa+lambda)*tau(i)/2)/((AffineGamma+kappa+lambda)*...
    (exp(AffineGamma*tau(i))-1)+2*AffineGamma));
    A(i) = -AffineAlpha/tau(i);
    H(i) = AffineBeta/tau(i);
end

%now recursive steps
AdjS = initx;
VarS = initV;
ll = zeros(nrow,1); %log-likelihood
for i = 1:nrow
    PredS = mu+F*AdjS; %predict values for S and Y
    Q = theta*sigma*sigma*(1-exp(-kappa*dt))^2/(2*kappa)+sigma*sigma/kappa*
(exp(-kappa*dt)-exp(-2*kappa*dt))*AdjS;
    VarS = F*VarS*F'+Q;
    PredY = A+H*PredS;
    PredError = Y(i,:)-PredY;
    VarY = H'*VarS*H+R;
    InvVarY = inv(VarY);
    DetY = det(VarY);
    %updating
    KalmanGain = VarS*H*InvVarY;
    AdjS = PredS+KalmanGain*PredError';
    VarS = VarS*(1-KalmanGain*H');
    ll(i) = -(ncol/2)*log(2*pi)-0.5*log(DetY)-0.5*PredError*InvVarY*PredError';
end
sumll = -sum(ll);
end

```

Hiermit versichere ich an Eides statt, dass ich die Seminararbeit selbständig und ohne Inanspruchnahme fremder Hilfe angefertigt habe. Ich habe dabei nur die angegebenen Quellen und Hilfsmittel verwendet und die aus diesen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen. Ich erkläre mich damit einverstanden, dass die Arbeit mit Hilfe eines Plagiatserkennungsdienstes auf enthaltene Plagiate überprüft wird.

Datum:

Ort:

Unterschrift: