

Das Chain Ladder Verfahren: Stochastische Modellierung und Überprüfung der Annahmen in ResQ

Matthias Isele (WV/AKS)

19.07.2023



Ihr Fels in der Brandung.

Inhaltsverzeichnis

1	Mathematische Grundlagen	1
1.1	Maßtheorie	1
1.2	Abbildungen zwischen messbaren Räumen	2
1.3	Integrationstheorie	2
1.4	Stochastik	4
1.5	Bedingter Erwartungswert und bedingte Varianz	5
2	Das Chain-Ladder-Verfahren	7
2.1	Die Methode	7
2.2	Zugrundeliegendes stochastisches Modell	8
2.3	Überprüfung der Modellannahmen	10

1 Mathematische Grundlagen

Dieser Abschnitt ist im Wesentlichen ein Excerpt aus dem Skript [1] zur Vorlesung Stochastik1 gehalten im WS 2020 von L. Döring an der Universität Mannheim. Die gesammelten Definitionen und Resultate sind Grundlagen zum mathematischen Verständnis des Chain-Ladder Verfahrens.

1.1 Maßtheorie

Sei $\Omega \neq \emptyset$ eine beliebige Grundmenge. Zu einer Teilmenge $A \subset \Omega$ bezeichnet $A^C := \Omega \setminus A$ das Komplement. $\mathcal{P}(\Omega)$ bezeichnet die Potenzmenge.

Definition 1 (σ -Algebra). Eine Teilmenge $\mathcal{A} \subset \mathcal{P}(\Omega)$ heißt σ -Algebra, falls

- (i) $\Omega \in \mathcal{A}$,
- (ii) $A \in \mathcal{A} \implies \Omega \setminus A \in \mathcal{A}$ (Abgeschlossenheit unter Komplementbildung),
- (iii) $A_1, A_2, \dots \in \mathcal{A} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ (Abgeschlossenheit unter abzählbaren Vereinigungen).

Solch ein Tupel (Ω, \mathcal{A}) nennt man **messbarer Raum**. Ist $\mathcal{B} \subset \mathcal{A}$ eine σ -Algebra so nennt man \mathcal{B} eine Unter- σ -Algebra.

Definition 2 (Maß). Ein Maß zu einer σ -Algebra \mathcal{A} ist eine Abbildung $\mu : \mathcal{A} \rightarrow [0, \infty]$, sodass

- (i) $\mu(\emptyset) = 0$,
- (ii) $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$, falls A_1, A_2, \dots paarweise disjunkt sind (σ -Additivität).

Für ein Maß lässt man explizit den Wert ∞ zu mit $[0, \infty] := [0, \infty) \cup \{\infty\}$. Man definiert zum Beispiel $a + \infty := \infty + a := a \cdot \infty = \infty \cdot a := \infty$, $a \in [0, \infty)$, $0 \cdot \infty = \infty \cdot 0 = 0$ und $\infty + \infty := \infty$. Ein Tripel $(\Omega, \mathcal{A}, \mu)$ nennt man **Maßraum**.

Definition 3 (Erzeugen einer σ -Algebra). Für $\mathcal{E} \subset \mathcal{P}(\Omega)$ heißt

$$\sigma(\mathcal{E}) := \bigcap_{\substack{\mathcal{B} \subset \mathcal{P}(\Omega) \\ \mathcal{B} \text{ } \sigma\text{-Alg.}}} \mathcal{B} \tag{1}$$

die von \mathcal{E} erzeugte σ -Algebra. Ist $\mathcal{A} = \sigma(\mathcal{E})$, so nennt man \mathcal{E} einen Erzeuger von \mathcal{A} .

Man überprüft leicht, dass beliebige Schnitte von σ -Algebren wieder σ -Algebren sind. Erzeuger sind im Allgemeinen nicht eindeutig. Die wichtigste σ -Algebra in der Stochastik ist die **Borel- σ -Algebra** $\mathcal{B}(\mathbb{R}^n)$ auf $\Omega = \mathbb{R}^n$. Sie wird zum Beispiel erzeugt durch

$$\mathcal{E} = \{O \subset \mathbb{R}^n \mid O \text{ offen}\}. \tag{2}$$

1.2 Abbildungen zwischen messbaren Räumen

Definition 4 (Messbare Abbildung). Seien (Ω, \mathcal{A}) , (Ω', \mathcal{A}') messbare Räume und $f : \Omega \rightarrow \Omega'$. Die Abbildung f heißt messbar, falls

$$f^{-1}(A') \in \mathcal{A}, \quad \forall A' \in \mathcal{A}'. \quad (3)$$

Man schreibt auch, dass f messbar bezüglich $(\mathcal{A}, \mathcal{A}')$ ist. Es reicht im Allgemeinen die Messbarkeit von f nur für einen Erzeuger von \mathcal{A}' zu prüfen.

Definition 5 (σ -Algebra erzeugt durch eine Abbildung). Sei $f : \Omega \rightarrow \Omega'$ und (Ω', \mathcal{A}') ein messbarer Raum. Die Menge

$$\mathcal{A} := \{f^{-1}(A') \mid A' \in \mathcal{A}'\} \quad (4)$$

ist die kleinste σ -Algebra, sodass f messbar bezüglich $(\mathcal{A}, \mathcal{A}')$ ist.

Definition 6 (Push-forward). Sei $f : \Omega \rightarrow \Omega'$ messbar bezüglich $(\mathcal{A}, \mathcal{A}')$ und μ ein Maß auf \mathcal{A} . Dann ist

$$f_*\mu(B) := \mu(f^{-1}(B)), \quad B \in \mathcal{A}' \quad (5)$$

ein Maß auf \mathcal{A}' , der sogenannte push-forward.

Nun wird die erweiterte Zahlengerade $\overline{\mathbb{R}} := [-\infty, \infty] := \mathbb{R} \cup \{-\infty, \infty\}$ eingeführt. Hierbei sind Operationen der Form $-\infty + (+\infty)$ nicht definiert. Auf $\overline{\mathbb{R}}$ definiert man die erweiterte Borel- σ -Algebra

$$\mathcal{B}(\overline{\mathbb{R}}) := \{B \subset \overline{\mathbb{R}} \mid B \cap \mathbb{R} \in \mathcal{B}(\mathbb{R})\}. \quad (6)$$

Definition 7 (Messbare numerische Funktion). Sei (Ω, \mathcal{A}) ein messbarer Raum. Eine bezüglich $(\mathcal{A}, \mathcal{B}(\overline{\mathbb{R}}))$ messbare Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ heißt messbare numerische Funktion.

1.3 Integrationstheorie

Definition 8 (Einfache Funktion). Zu einem Messbaren Raum (Ω, \mathcal{A}) heißt eine Abbildung $\phi : \Omega \rightarrow \overline{\mathbb{R}}$ einfache Funktion, falls sie der Form

$$\phi = \sum_{i=1}^N \alpha_i 1_{A_i} \quad (7)$$

ist, wobei $n \in \mathbb{N}$, und $\alpha_i \in [0, \infty]$, $i = 1, \dots, N$. Hierbei sind $A_1, A_2, \dots, A_N \in \mathcal{A}$ paarweise disjunkt und 1_A bezeichnet die Indikatorfunktion von $A \in \mathcal{A}$ auf Ω .

Einfache Funktionen sind numerische Funktionen bezüglich $(\mathcal{A}, \mathcal{B}(\overline{\mathbb{R}}))$ und positiv $\phi \geq 0$. Das Lebesgue-Integral einer Einfachen Funktion ist definiert als

$$\int_{\Omega} \phi d\mu := \sum_{i=1}^N \alpha_i 1_{A_i}. \quad (8)$$

Das Lebesgue-Integral einer $(\mathcal{A}, \mathcal{B}(\overline{\mathbb{R}}))$ -messbaren numerischen positiven Funktion $f \geq 0$ ist definiert als

$$\int_{\Omega} f \, d\mu := \sup \left\{ \int_{\Omega} \phi \, d\mu \mid \phi \text{ ist einfach mit } \phi \leq f \right\}. \quad (9)$$

Folgende Theoreme erklären, warum messbare Funktionen so wichtig sind. Die zugehörigen Beweise sind zu finden in [1, S. 42].

Theorem 1 (Approximation numerischer Funktionen durch einfache Funktionen). *Für jede nicht-negative numerische Funktion f existiert eine wachsende Folge von einfachen Funktionen $(\phi_n)_{n \in \mathbb{N}}$ mit $\phi_n \uparrow f$ punktweise.*

Theorem 2 (Monotone Konvergenz Theorem für einfache Funktionen). *Sei f eine nicht-negative numerische Funktion und $(\phi_n)_{n \in \mathbb{N}}$ eine wachsende Folge von einfachen Funktionen mit $\phi_n \uparrow f$ punktweise. Dann gilt*

$$\lim_{n \rightarrow \infty} \int_{\Omega} \phi_n \, d\mu = \int_{\Omega} f \, d\mu, \quad (10)$$

wobei in der Gleichheit $+\infty = +\infty$ möglich ist.

Man kann Gleichung (10) also auch als alternative äquivalente Definition des Lebesgue Integrals über nicht-negative numerische Funktionen verwenden. Zu einer numerischen Funktion $f : \Omega \rightarrow \overline{\mathbb{R}}$ betrachte man den Positiv- und Negativteil $f^{\pm} : \Omega \rightarrow [0, \infty]$, punktweise definiert durch $f^+ := \max\{f, 0\}$ und $f^- = \max\{-f, 0\}$. Diese sind ebenfalls numerische Funktionen mit

$$f = f^+ - f^- \quad \text{und} \quad |f| = f^+ + f^-. \quad (11)$$

Definition 9 (Lebesgue-Integral). Sei $f : \Omega \rightarrow \overline{\mathbb{R}}$ eine Numerische Funktion, sodass $\int_{\Omega} f^+ \, d\mu$ oder $\int_{\Omega} f^- \, d\mu$ endlich ist. Dann heißt f quasi-integrierbar und das Lebesgue-Integral ist definiert als

$$\int_{\Omega} f \, d\mu := \int_{\Omega} f^+ \, d\mu - \int_{\Omega} f^- \, d\mu. \quad (12)$$

Falls (12) endlich ist, so heißt f integrierbar. Schließlich definiert man für jedes $A \in \mathcal{A}$

$$\int_A f \, d\mu := \int_{\Omega} 1_A f \, d\mu. \quad (13)$$

Es ist wohlbekannt, dass das Lebesgue-Integral linear ist [1, S. 45].

Theorem 3 (Abstrakter Transformationssatz.). *Seien $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, (Ω', \mathcal{A}') ein messbarer Raum, $f : \Omega \rightarrow \Omega'$ und $g : \Omega' \rightarrow \overline{\mathbb{R}}$ messbar. Die numerische Funktion g ist integrierbar bezüglich des push-forward $f_*\mu$ genau dann, wenn $g \circ f$ integrierbar ist bezüglich μ . In diesem Fall gilt der Transformationssatz*

$$\int_{\Omega} g \circ f \, d\mu = \int_{\Omega'} g \, d(f_*\mu). \quad (14)$$

Beweisskizze. Für alle $A \in \Omega'$ gilt

$$\int_{\Omega} 1_A \circ f \, d\mu = \int_{\Omega} 1_{f^{-1}(A)} \, d\mu = \mu(f^{-1}(A)) = \int_{\Omega'} 1_A \, df_*\mu. \quad (15)$$

Unter Ausnutzung der Linearität des Lebesgue-Integrals ist das Theorem also wahr, falls g eine einfache integrierbare Funktion ist. Der allgemeine Fall folgt durch sorgfältige Approximation von g^+ und g^- mit einfachen Funktionen. Für Details siehe [1, S. 50]. \square

1.4 Stochastik

Definition 10 (Wahrscheinlichkeitsraum). Ein Maßraum (Ω, \mathcal{A}, P) heißt Wahrscheinlichkeitsraum, falls $P(\Omega) = 1$. In diesem Fall bezeichnet man Ω als Ergebnismenge und Elemente $A \in \mathcal{A}$ als Ereignisse oder Events. Das Maß P nennt man Wahrscheinlichkeitsmaß.

Definition 11 (Stochastisches Modell). Ein stochastisches Modell ist ein Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) ausgestattet mit einer $(\mathcal{A}, \mathcal{B}(\mathbb{R}^n))$ -messbaren Abbildung $X : \Omega \rightarrow \mathbb{R}^n$. Man nennt X einen **Zufallsvektor** und für $n = 1$ auch eine **Zufallsvariable**.

Definition 12 (Stochastische Unabhängigkeit). Unter- σ -Algebren $\mathcal{B}_1, \dots, \mathcal{B}_n \subset \mathcal{A}$ heißen stochastisch unabhängig, falls $P(B_i \cap B_j) = P(B_i)P(B_j)$ für alle $B_i \in \mathcal{B}_i$, $B_j \in \mathcal{B}_j$ und $i \neq j$ gilt.

Definition 13 (Unabhängigkeit von Zufallsvektoren). Zufallsvektoren $X_i : \Omega \rightarrow \mathbb{R}^{n_i}$, $i = 1, \dots, n$ heißen unabhängig, falls deren erzeugte σ -Algebren $\sigma(X_i)$, $i = 1, \dots, n$ stochastisch unabhängig sind.

Definition 14 (Erwartungswert und Varianz). Der Erwartungswert einer Zufallsvariablen X ist definiert als

$$E(X) := \int_{\Omega} X \, dP, \quad (16)$$

sofern das Integral Sinn macht. Man nimmt oft an, dass der Erwartungswert existiert, das heißt $E(|X|) < \infty$. Die Varianz ist definiert als

$$\text{Var}(X) = E[(X - E(X))^2] = E(X^2) - E(X)^2, \quad (17)$$

sofern der Ausdruck Sinn macht.

Lemma 1 (Faktorisierung des Erwartungswertes). Seien $X_1, \dots, X_n : \Omega \rightarrow \mathbb{R}$ voneinander unabhängige Zufallsvariablen mit existierendem Erwartungswert. Dann gilt

$$E\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n E(X_i). \quad (18)$$

Beweisskizze. Für $n = 2$ betrachte zwei unabhängige Zufallsvariablen X, Y auf Ω mit endlichem Erwartungswert. Für alle $A \in \sigma(X)$, $B \in \sigma(Y)$ gilt

$$\int_{\Omega} 1_A 1_B \, dP = \int_{\Omega} 1_{A \cap B} \, dP = P(A \cap B) = P(A)P(B) = \left(\int_{\Omega} 1_A \, dP\right) \left(\int_{\Omega} 1_B \, dP\right). \quad (19)$$

Unter Ausnutzung der Linearität des Lebesgue-Integrals ist die Behauptung also wahr, falls X und Y einfache Funktionen sind. Der allgemeine Fall folgt durch Approximation der Positiv- und Negativteile X^\pm, Y^\pm mithilfe von einfachen Funktionen $\phi_n^{X^\pm}, \phi_m^{Y^\pm}$, jeweils messbar bezüglich $\sigma(X)$ und $\sigma(Y)$. Hierbei muss der Positivteil $X^+Y^+ + X^-Y^-$ und der Negativteil $X^+Y^- + X^-Y^+$ getrennt betrachtet werden. Der Fall $n \geq 2$ ist analog. \square

Definition 15 (Stochastischer Prozess). Eine Folge von Zufallsvariablen $(X_t)_{t \in T}$ definiert auf demselben Wahrscheinlichkeitsraum heißt stochastischer Prozess.

1.5 Bedingter Erwartungswert und bedingte Varianz

Definition 16 (Bedingter Erwartungswert). Sei $\mathcal{B} \subset \mathcal{A}$ eine Unter- σ -algebra und $X : \Omega \rightarrow \mathbb{R}$ eine \mathcal{A} -messbare Zufallsvariable, dessen Erwartungswert existiert. Eine \mathcal{B} -messbare Zufallsvariable $Z =: E(X|\mathcal{B}) : \Omega \rightarrow \mathbb{R}$ heißt bedingter Erwartungswert von X bezüglich \mathcal{B} , falls

$$E(1_B X) = E(1_B Z) \quad (20)$$

für alle $B \in \mathcal{B}$ gilt. Hier ist 1_B die Indikatorfunktion bezüglich B .

Man kann zeigen dass der bedingte Erwartungswert stets existiert und eindeutig ist fast überall. Für einen \mathcal{A} -messbaren Zufallsvektor $Y = (Y_1, \dots, Y_n) : \Omega \rightarrow \mathbb{R}^n$ schreibt man

$$E(X|Y_1, \dots, Y_n) := E(X|\sigma(Y)), \quad (21)$$

wobei $\mathcal{B} = \sigma(Y)$ die von Y generierte Unter- σ -algebra ist. Für $B = \Omega$ in (20) erhält man für den totalen Erwartungswert

$$E(X) = E(E(X|\mathcal{B})). \quad (22)$$

Ebenso gilt die Turmeigenschaft.

Lemma 2 (Turmeigenschaft). Seien $X : \Omega \rightarrow \mathbb{R}$ eine \mathcal{A} -messbare Zufallsvariable mit endlichem Erwartungswert und $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \mathcal{A}$ zwei Unter- σ -Algebren, so gilt

$$E(X|\mathcal{B}_1) = E(E(X|\mathcal{B}_2)|\mathcal{B}_1). \quad (23)$$

Beweis. Für alle $B_1 \in \mathcal{B}_1$ gilt per definition

$$E(1_{B_1} E(X|\mathcal{B}_2)) = E(1_{B_1} E(E(X|\mathcal{B}_2)|\mathcal{B}_1)). \quad (24)$$

Da $B_1 \in \mathcal{B}_2$ folgt andererseits

$$E(1_{B_1} E(X|\mathcal{B}_2)) = E(1_{B_1} X) = E(1_{B_1} E(X|\mathcal{B}_1)). \quad (25)$$

\square

Lemma 3 (Herausziehen der Zufallsvariablen). Sei $X : \Omega \rightarrow \mathbb{R}$ $\mathcal{B} \subset \mathcal{A}$ -messbar mit existierendem Erwartungswert. Dann gilt

$$X = E(X|\mathcal{B}). \quad (26)$$

Allgemeiner sei $Y : \Omega \rightarrow \mathbb{R}$ \mathcal{A} -messbar mit $E(|XY|) < \infty$, so gilt

$$E(XY|\mathcal{B}) = XE(Y|\mathcal{B}). \quad (27)$$

Beweis. Es wird nur der Spezialfall (26) bewiesen. Für alle $B \in \mathcal{B}$ gilt

$$E(1_B X) = E(1_B E(X|\mathcal{B})), \quad (28)$$

oder äquivalent

$$0 = \int_{\Omega} 1_B \underbrace{(X - E(X|\mathcal{B}))}_{=: \tilde{X}} dP, \quad (29)$$

wobei $E(X|\mathcal{B})$ \mathcal{B} -messbar ist per definition. Insbesondere ist dies auch wahr für $B = \tilde{X}^{-1}((0, \infty))$ und $B = \tilde{X}^{-1}((-\infty, 0))$. Somit gilt die Behauptung fast überall. Der Beweis des allgemeinen Falls ist nachzulesen in [3]. \square

Lemma 4 (Unwichtigkeit unabhängiger Information). Sei $X : \Omega \rightarrow \mathbb{R}$ \mathcal{A} -messbar mit endlichem Erwartungswert und $\mathcal{B}, \mathcal{H} \subset \mathcal{A}$ σ -Unteralgebren, sodass \mathcal{H} unabhängig ist von $\sigma(\mathcal{B}, \sigma(X))$. Dann gilt

$$E(X|\sigma(\mathcal{B}, \mathcal{H})) = E(X|\mathcal{B}). \quad (30)$$

Beweis. Zu zeigen ist

$$E(X1_A) = E(E(X|\mathcal{B})1_A) \quad (31)$$

für alle $A \in \sigma(\mathcal{B}, \mathcal{H})$. Hierbei reicht es, die Aussage für ein Erzeugendensystem S von $\sigma(\mathcal{B}, \mathcal{H})$ zu zeigen, gewählt sei $S = \{B \cap H | B \in \mathcal{B}, H \in \mathcal{H}\}$. Nach Voraussetzung sind die Zufallsvariablen 1_H und $E(X|\mathcal{B})1_B$ unabhängig, da sie \mathcal{H} - beziehungsweise \mathcal{B} -messbar sind. Ebenso sind 1_H und $X1_B$ unabhängig, da $\sigma(X1_B) \subset \sigma(\mathcal{B}, \sigma(X))$. Man erhält

$$\begin{aligned} E(E(X|\mathcal{B})1_{B \cap H}) &= E(E(X|\mathcal{B})1_B 1_H) = E(E(X|\mathcal{B})1_B)E(1_H) \\ &= E(X1_B)E(1_H) = E(X1_{B \cap H}). \end{aligned} \quad (32)$$

\square

Definition 17 (Bedingte Varianz). Sei $\mathcal{B} \subset \mathcal{A}$ eine Unter- σ -algebra und $X : \Omega \rightarrow \mathbb{R}$ eine \mathcal{A} -messbare Zufallsvariable, dessen Erwartungswert existiert. Die bedingte Varianz ist definiert als

$$\text{Var}(X|\mathcal{B}) := E(X^2|\mathcal{B}) - E(X|\mathcal{B})^2. \quad (33)$$

Wie beim bedingten Erwartungswert schreibt man

$$\text{Var}(X|Y_1, \dots, Y_n) := \text{Var}(X|\sigma(Y)) \quad (34)$$

für einen beliebigen \mathcal{A} -messbaren Zufallsvektor $Y = (Y_1, \dots, Y_n) : \Omega \rightarrow \mathbb{R}^n$.

2 Das Chain-Ladder-Verfahren

2.1 Die Methode

Das Chain-Ladder Verfahren ist eine Methode zur Vervollständigung von Abwicklungsdreiecken zu Abwicklungsquadraten, also zur Prognose zukünftiger Schadenständen. Ein Abwicklungsdreieck ist gegeben als eine Familie von Schadenständen $(S_{i,k})$ in \mathbb{R} mit $0 \leq i, k$ und $i + k \leq n$. Hierbei bezeichnet i das **Anfalljahr**, k das **Abwicklungsjahr** und n das **aktuelle Kalenderjahr**, veranschaulicht in Abbildung 1.

Anfall- jahr	Abwicklungsjahr								
	0	1	...	k	...	$n-i$...	$n-1$	n
0	$S_{0,0}$	$S_{0,1}$...	$S_{0,k}$...	$S_{0,n-i}$...	$S_{0,n-1}$	$S_{0,n}$
1	$S_{1,0}$	$S_{1,1}$...	$S_{1,k}$...	$S_{1,n-i}$...	$S_{1,n-1}$	
\vdots	\vdots	\vdots		\vdots		\vdots			
i	$S_{i,0}$	$S_{i,1}$...	$S_{i,k}$...	$S_{i,n-i}$			
\vdots	\vdots	\vdots		\vdots					
$n-k$	$S_{n-k,0}$	$S_{n-k,1}$...	$S_{n-k,k}$					
\vdots	\vdots	\vdots							
$n-1$	$S_{n-1,0}$	$S_{n-1,1}$							
n	$S_{n,0}$								

Abbildung 1: Abwicklungsdreieck mit Schadenständen $S_{i,k}$.

Eine Erweiterung des Abwicklungsdreieckes zu einem Abwicklungsquadrat $(S_{i,k})_{0 \leq i,k \leq n}$ in \mathbb{R} ist veranschaulicht in Abbildung 2. Das Chain-Ladder Verfahren definiert, wie Abwicklungsdreiecke zu abwicklungsquadraten zu erweitern sind.

Definition 18 (Chain-Ladder Verfahren). Sei $(S_{i,k})$ ein Abwicklungsdreieck zum Kalenderjahr n sodass die individuellen Abwicklungsfaktoren

$$\hat{\varphi}_{i,k} := \frac{S_{i,k}}{S_{i,k-1}} \quad (35)$$

existieren für alle $i + k \leq n$, $k \geq 1$, $i \geq 0$. Bilde die Chain-Ladder Faktoren

$$\hat{\varphi}_k^{\text{CL}} := \frac{\sum_{i=0}^{n-k} S_{i,k}}{\sum_{i=0}^{n-k} S_{i,k-1}} = \sum_{i=0}^{n-k} W_{i,k} \hat{\varphi}_{i,k}, \quad k = 1, \dots, n, \quad (36)$$

mit Gewichtungen

$$W_{i,k} := \frac{S_{i,k-1}}{\sum_{j=0}^{n-k} S_{j,k-1}}. \quad (37)$$

Das Chain-Ladder Verfahren ist die Erweiterung des Abwicklungsdreieckes zu einem Abwicklungsquadrat durch

$$S_{i,k} = \hat{\varphi}_k^{\text{CL}} \hat{\varphi}_{k-1}^{\text{CL}} \cdots \hat{\varphi}_{n-i+1}^{\text{CL}} S_{i,n-i}, \quad i+k > n. \quad (38)$$

Anfall- jahr	Abwicklungsjahr								
	0	1	...	k	...	$n-i$...	$n-1$	n
0	$S_{0,0}$	$S_{0,1}$...	$S_{0,k}$...	$S_{0,n-i}$...	$S_{0,n-1}$	$S_{0,n}$
1	$S_{1,0}$	$S_{1,1}$...	$S_{1,k}$...	$S_{1,n-i}$...	$S_{1,n-1}$	$S_{1,n}$
\vdots	\vdots	\vdots		\vdots		\vdots		\vdots	\vdots
i	$S_{i,0}$	$S_{i,1}$...	$S_{i,k}$...	$S_{i,n-i}$...	$S_{i,n-1}$	$S_{i,n}$
\vdots	\vdots	\vdots		\vdots		\vdots		\vdots	\vdots
$n-k$	$S_{n-k,0}$	$S_{n-k,1}$...	$S_{n-k,k}$...	$S_{n-k,n-i}$...	$S_{n-k,n-1}$	$S_{n-k,n}$
\vdots	\vdots	\vdots		\vdots		\vdots		\vdots	\vdots
$n-1$	$S_{n-1,0}$	$S_{n-1,1}$...	$S_{n-1,k}$...	$S_{n-1,n-i}$...	$S_{n-1,n-1}$	$S_{n-1,n}$
n	$S_{n,0}$	$S_{n,1}$...	$S_{n,k}$...	$S_{n,n-i}$...	$S_{n,n-1}$	$S_{n,n}$

Abbildung 2: Erweiterung des Abwicklungsdreieckes zu einem Abwicklungsquadrat. Die Schadenstände $S_{i,k}$ mit $i+k > n$ sind Prognosen für die Zukunft.

2.2 Zugrundeliegendes stochastisches Modell

Unter welchen Annahmen das Chain-Ladder Verfahren sinnvolle Prognosen liefert, kann durch das folgende stochastische Modell erklärt werden. Wir betrachten einen Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) sowie eine Familie $(S_{i,k})_{0 \leq i,k \leq n}$ von $(\mathcal{A}$ -messbaren) Zufallsvariablen $\Omega \rightarrow \mathbb{R}$. Die Schadenstände werden also als Realisierungen von Zufallsvariablen aufgefasst. Sie sollen folgendes Muster (Chain-Ladder Voraussetzungen) erfüllen.

(CL1) Es gibt Abwicklungsfaktoren $\varphi_1, \dots, \varphi_n \in \mathbb{R}$ mit $E(S_{i,k} | S_{i,0}, \dots, S_{i,k-1}) = \varphi_k S_{i,k-1}$ für alle $k = 1, \dots, n$ und $i = 0, \dots, n$.

(CL2) Die Zufallsvektoren $(S_{i,0}, \dots, S_{i,n})$, $i = 0, \dots, n$, sind unabhängig.

Theorem 4. *Es gilt*

$$E(S_{i,k} | S_{i,0}, \dots, S_{i,n-i}) = \varphi_k \varphi_{k-1} \cdots \varphi_{n-i+1} S_{i,n-i} \quad (39)$$

für alle $i+k > n$.

Beweis. Aus der Turmeigenschaft, CL1 und Herausziehen der Zufallsvariablen folgt

$$\begin{aligned}
E(S_{i,k}|S_{i,0}, \dots, S_{i,n-i}) \\
&= E(E(S_{i,k}|S_{i,0}, \dots, S_{i,k-1})|S_{i,0}, \dots, S_{i,n-i}) \\
&= \varphi_k E(S_{i,k-1}|S_{i,0}, \dots, S_{i,n-i}) \\
&= \dots = \varphi_k \varphi_{k-1} \dots \varphi_{n-i+1} E(S_{i,n-i}|S_{i,0}, \dots, S_{i,n-i}) \\
&= \varphi_k \varphi_{k-1} \dots \varphi_{n-i+1} S_{i,n-i}
\end{aligned} \tag{40}$$

□

Theorem 5. Die Schätzer $(\hat{\varphi}_k^{CL})$ aus (36), aufgefasst als Zufallsvariablen, sind unabhängig und erwartungstreu zu den Abwicklungsfaktoren (φ_k) aus CL1. Das heißt

$$E(\hat{\varphi}_{k_1}^{CL} \dots \hat{\varphi}_{k_m}^{CL}) = E(\hat{\varphi}_{k_1}^{CL}) \dots E(\hat{\varphi}_{k_m}^{CL}) \quad \text{und} \quad E(\hat{\varphi}_k^{CL}) = \varphi_k \tag{41}$$

für alle $k = 1, \dots, n$ und $k_1, \dots, k_m = 1, \dots, n$ paarweise unterschiedlich.

Beweis. Setze $B_k := (S_{i,l})_{i=0, \dots, n-k, l=0, \dots, k-1}$. Bezüglich der Erwartungstreue gilt

$$\begin{aligned}
E(\hat{\varphi}_k^{CL}) &= \sum_{i=0}^{n-k} E\left(\frac{S_{i,k}}{\sum_{j=0}^{n-k} S_{j,k-1}}\right) = \sum_{i=0}^{n-k} E\left(\frac{1}{\sum_{j=0}^{n-k} S_{j,k-1}} E(S_{i,k}|B_k)\right) \\
&= \sum_{i=0}^{n-k} E\left(\frac{1}{\sum_{j=0}^{n-k} S_{j,k-1}} E(S_{i,k}|S_{i,0}, \dots, S_{i,k-1})\right) \\
&= \sum_{i=0}^{n-k} E\left(\frac{\varphi_k S_{i,k-1}}{\sum_{j=0}^{n-k} S_{j,k-1}}\right) = \varphi_k.
\end{aligned} \tag{42}$$

Für $l < k$ gilt

$$E(\hat{\varphi}_k^{CL} \hat{\varphi}_l^{CL}) = E(E(\hat{\varphi}_k^{CL} \hat{\varphi}_l^{CL}|B_k)) = E(\hat{\varphi}_l^{CL} \underbrace{E(\hat{\varphi}_k^{CL}|B_k)}_{=\varphi_k}) = \varphi_k \varphi_l. \tag{43}$$

Der allgemeine Fall ist analog. □

Die Schätzer in (38) des Chain-Ladder Verfahrens für zukünftige Schadenstände $S_{i,k}$, $i + k > n$, stimmen also mit dem postulierten stochastischen Modell im bedingten Erwartungswert (bekanntes Abwicklungsdreieck) überein. Das heißt

$$\begin{aligned}
E(S_{i,k}|S_{i,0}, \dots, S_{i,n-i}) &= E(\hat{\varphi}_k^{CL} \dots \hat{\varphi}_{n-i+1}^{CL} S_{i,n-i}|S_{i,0}, \dots, S_{i,n-i}) \\
&= E(\varphi_k \dots \varphi_{n-i+1} S_{i,n-i}|S_{i,0}, \dots, S_{i,n-i}).
\end{aligned} \tag{44}$$

Die wesentlichen Annahmen des Modells sind das Abwicklungsmuster CL1 sowie die Unabhängigkeit in den Anfalljahren CL2. Durch die Wahl der Schätzer $\hat{\varphi}_k^{CL}$ wurde implizit eine Annahme CL3 über die Varianz getroffen im folgenden Sinne. Man kann zeigen [2, S. 41-34 und 248-249], dass die Schätzer $\hat{\varphi}_k^{CL}$ genau dann optimal sind bezüglich Minimierung der Varianz, wenn gilt:

(CL3) Es gibt Proportionalitätskonstanten $\sigma_1, \dots, \sigma_n \geq 0$ mit

$$\text{Var}\left(\frac{S_{i,k}}{S_{i,k-1}}|S_{i,0}, \dots, S_{i,k-1}\right) = \sigma_k^2 / S_{i,k-1}, \quad k = 1, \dots, n, \quad i = 0, \dots, n.$$

2.3 Überprüfung der Modellannahmen

Aus der Modellannahme (CL1) ergibt sich für festes $k \in \{1, \dots, n\}$ und für alle $i = 0, \dots, n$

$$\frac{S_{i,k}}{S_{i,k-1}} = \frac{\varphi_k S_{i,k}}{E_i} = \varphi_k \frac{E_i + S_{i,k} - E_i}{E_i} = \varphi_k + \varepsilon_i, \quad (45)$$

wobei $E_i := E(S_{i,k} | S_{i,0}, \dots, S_{i,k-1})$ abgekürzt und $E_i \neq 0$ angenommen wurde. Der Störterm $\varepsilon_i := \varphi_k(S_{i,k} - E_i)/E_i$ verschwindet im bedingten Erwartungswert

$$E(\varepsilon_i | S_{i,0}, \dots, S_{i,k-1}) = 0. \quad (46)$$

Gleichung (45) kann unter CL1 in die form

$$S_{i,k} = \varphi_k S_{i,k-1} + \tilde{\varepsilon}_i \quad (47)$$

gebracht werden mit Störterm $\tilde{\varepsilon}_i = S_{i,k} - E_i$, der im bedingten Erwartungswert verschwindet. Aus der Modellannahme CL3 ergibt sich für die bedingte Varianz

$$\text{Var}(\tilde{\varepsilon}_i | S_{i,0}, \dots, S_{i,k-1}) = \text{Var}(S_{i,k} | S_{i,0}, \dots, S_{i,k-1}) = S_{i,k-1} \sigma_k^2. \quad (48)$$

Die Varianz des Störterms ist also nicht konstant; Sie hängt vom Anfalljahr i ab. Dies macht Gleichung (47) unbrauchbar als Ansatz für eine lineare Regression. Jedoch kann reskaliert werden

$$S_{i,k}/\sqrt{S_{i,k-1}} = \varphi_k S_{i,k-1}/\sqrt{S_{i,k-1}} + \tilde{\varepsilon}_i/\sqrt{S_{i,k-1}}, \quad (49)$$

sodass die Varianz des reskalierten Störterms

$$\text{Var}(\tilde{\varepsilon}_i/\sqrt{S_{i,k-1}} | S_{i,0}, \dots, S_{i,k-1}) = \sigma_k^2 \quad (50)$$

konstant ist. (k ist fest.) Der reskalierte Störterm hat ebenfalls einen verschwindenden bedingten Erwartungswert. Nun kann die übliche Methode der kleinsten Quadrate zur Schätzung von φ_k angewandt werden. Dazu minimiert man

$$\sum_{i=0}^n (S_{i,k} \sqrt{S_{i,k-1}} - S_{i,k-1} \varphi_k / \sqrt{S_{i,k-1}})^2 = \sum_{i=0}^n (S_{i,k} - S_{i,k-1} \varphi_k)^2 / |S_{i,k-1}| \quad (51)$$

nach φ_k . Die Lösung ist genau der Chain-Ladder Schätzer [2, S. 256]

$$\varphi_k = \hat{\varphi}_k^{\text{CL}}. \quad (52)$$

Aus (CL1) und (CL3) folgt also, dass die Faktoren $S_{i,k}/S_{i,k-1}$ für festes k und für alle i konstant gleich dem Chain-Ladder Schätzer $\hat{\varphi}_k^{\text{CL}}$ sein sollten bis auf zufällige schwankungen mit verschwindendem Erwartungswert.

In ResQ kann dies mithilfe der Plots zu finden unter 'Ratios' -> 'Single Development' (DFM-Methode) überprüft werden.

Die normierten Residuen

$$r_{i,k} := \frac{S_{i,k} - S_{i,k-1}\varphi_k}{\sigma_k \sqrt{S_{i,k-1}}} \stackrel{(CL1)}{=} \frac{S_{i,k} - E(S_{i,k}|S_{i,0}, \dots, S_{i,k-1})}{\sigma_k \sqrt{S_{i,k-1}}} \quad (53)$$

haben verschwindenden Erwartungswert

$$E(r_i|S_{i,0}, \dots, S_{i,k-1}) = 0 \quad (54)$$

und Einheitsvarianz

$$\begin{aligned} \text{Var}(r_{i,k}|S_{i,0}, \dots, S_{i,k-1}) &= \left[E(S_{i,k}^2|S_{i,0}, \dots, S_{i,k-1}) - E(S_{i,k}|S_{i,0}, \dots, S_{i,k-1})^2 \right] \frac{1}{\sigma_k^2 S_{i,k-1}} \\ &= \left[E(S_{i,k}^2/S_{i,k-1}^2|S_{i,0}, \dots, S_{i,k-1}) - E(S_{i,k}/S_{i,k-1}|S_{i,0}, \dots, S_{i,k-1})^2 \right] \frac{S_{i,k-1}}{\sigma_k^2} \\ &= \text{Var}(S_{i,k}/S_{i,k-1}|S_{i,0}, \dots, S_{i,k-1}) \frac{S_{i,k-1}}{\sigma_k^2} \stackrel{(CL3)}{=} 1. \end{aligned} \quad (55)$$

Für die erzeugte σ -Algebra gilt

$$\sigma(r_{i,k}) \subset \sigma(S_{i,0}, \dots, S_{i,n}). \quad (56)$$

Laut (CL2) sind also zwei Residuen $r_{i,k}$ und $r_{j,l}$ unabhängig, falls $i \neq j$. Zusammenfassend betrachte man stochastische Prozesse der form

$$\tilde{r} := (\tilde{r}_i)_{i=1, \dots, n} := (r_{i,k_i})_{i=1, \dots, n}. \quad (57)$$

Es wird also zu jedem Anfalljahr ein beliebiges Residuum ausgewählt. Aus den drei Chain-Ladder Annahmen (CL1) - (CL3) folgt dann, dass \tilde{r} Weißes Rauschen ist, das heißt,

$$\text{i) } E(\tilde{r}_i|S_{i,0}, \dots, S_{i,k_i-1}) = 0,$$

$$\text{ii) } \text{Var}(\tilde{r}_i|S_{i,0}, \dots, S_{i,k_i-1}) = 1,$$

iii) Die Elemente in \tilde{r} sind paarweise unabhängig.

In ResQ kann dies mithilfe des Residuenplots zu finden unter 'Checks' -> 'Residuals Graph' (DFM-Methode) überprüft werden. Hier werden die normierten Residuen für alle verfügbaren Abwicklungsjahre je Anfalljahr abgebildet. Im Schnitt sollten die Residuen keinen Trend in den Anfalljahren zeigen und gleichmäßig (konstante Varianz = 1) um 0 herum schwanken.

Literatur

- [1] L. DÖRING, *Stochastik 1*. https://www.wim.uni-mannheim.de/media/Lehrstuehle/wim/doering/WIM_LS_Doering/Stochastik_HWS20/Stochastik1.pdf, Vorlesungsskript WS 2020, Universität Mannheim.
- [2] T. MACK, *Schriftenreihe Angewandte Versicherungsmathematik, Heft 28, Schaden-versicherungsmathematik*, Verlag Versicherungswirtschaft E.V., Karlsruhe, 1997.
- [3] G. ZITKOVIC, *Theory of probability 1*. https://web.ma.utexas.edu/users/gordanz/notes/conditional_expectation.pdf, Vorlesungsskript WS 2013, The University of Texas at Austin.