

Zeitdiskrete Markowketten

Marcel Wienöbst

25. April 2022

Einleitung

Ich beginne zunächst damit, an die Ausarbeitung von Eva Martenstein [3] anknüpfend, das unbeschränkte Gambler's Ruin Problem zu diskutieren. Im Anschluss führe ich wichtige Grundlagen diskreter Markowketten ein. Darauf aufbauend betrachte ich die Modellierung von allgemeinen Geburts- und Sterbeprozessen durch Markowketten sowie logistische Wachstumsprozesse. Dabei beschränke ich mich auf endliche Modelle.

Für eine bessere Intuition der verschiedenen Populationsmodelle, habe ich eine Vielzahl von Plots und Simulationen angefertigt. Diese sind als Ergänzung zu dieser Ausarbeitung zu verstehen und unter <https://nbviewer.jupyter.org/github/mwien/ZeitdiskreteMarkovketten/blob/master/Seminar.ipynb> abrufbar.

Das unbeschränkte Gambler's Ruin Problem

Das unbeschränkte Gambler's Ruin Problem beschreibt das folgende Szenario: Die Position x repräsentiert das Kapital eines Spielers und in jedem Schritt erhöht sich das Kapital auf $x + 1$ oder es sinkt auf $x - 1$. Diese Übergänge treten mit Wahrscheinlichkeit p bzw. q auf. Der Wertebereich ist $\{0, 1, 2, \dots\}$. Dies kann als Spiel gegen die *Bank* interpretiert werden. Das bedeutet wiederum, dass der Spieler entweder unendlich lange spielt oder die Bank gewinnt (der Spieler verliert sein gesamtes Kapital und landet im Zustand $x = 0$). Es ergeben sich damit die Fragen nach der Wahrscheinlichkeit, dass der Spieler vom Kapital k ausgehend verliert, und die erwartete Dauer bis dies geschieht.

Betrachten wir zunächst erstere Fragestellung. In der Ausarbeitung von Eva Martenstein [3] wird diese Wahrscheinlichkeit als $A_k(1)$ beschrieben, wobei die allgemeine Lösung von $A_k(t)$ die Form

$$A_k(t) = c_1 \lambda_1^k + c_2 \lambda_2^k$$

hatte. Nun ändern sich die Randbedingungen zu:

$$A_0(t) = 1 \text{ und } \lim_{k \rightarrow \infty} |A_k(t)| < \infty$$

Für die genaue Herleitung der Lösung sei auf Kapitel 3.4 von Allen in [1] verwiesen. Es ergibt sich:

$$a_k = A_k(1) = \begin{cases} 1 & \text{wenn } p \leq q, \\ (q/p)^k & \text{wenn } p > q \end{cases}$$

Das bedeutet, dass der Spieler immer verliert, solange das Spiel nicht zu seinen Gunsten ist. Insbesondere also auch im Falle des *fairen* Spiels für $p = q$. Dieses paradoxe Resultat verleiht dem Gambler's Ruin Problem seinen Namen.

Nun betrachten wir die Frage nach der erwarteten Dauer bis zum Verlust, wenn der Spieler mit Kapital k beginnt. Wir geben diese als τ_k an:

$$\tau_k = \begin{cases} k/(q-p), & \text{wenn } p < q, \\ \infty, & \text{wenn } p \geq q. \end{cases}$$

Interessant ist erneut den Fall $p = q$ zu betrachten. Hier verliert der Spieler mit Wahrscheinlichkeit 1, die Dauer ist jedoch ∞ .

Abschließend sei angemerkt, dass das Gambler's Ruin Problem auch als Modellierung einer Population aufgefasst werden kann, wobei der Zustand x in diesem Fall die Größe der Population angibt. Dieses Modell hat jedoch einige Einschränkungen. So bleibt die Populationsgröße nie konstant, sondern vergrößert oder verkleinert sich in jedem Schritt. Außerdem sind die Übergangswahrscheinlichkeiten p und q unabhängig von der Populationsgröße. Wir werden dieses Modell daher im weiteren Verlauf dieser Ausarbeitung verallgemeinern.

Stationäre Wahrscheinlichkeitsverteilungen

Wir betrachten nun einige wichtige theoretische Ergebnisse, die uns bei der späteren Analyse der Populationsmodelle behilflich sind.

Definition 1. Eine stationäre Wahrscheinlichkeitsverteilung einer Markowkette mit den Zuständen $\{1, 2, \dots\}$ ist der nichtnegative Vektor $\pi = (\pi_1, \pi_2, \dots)^T$, der die Bedingungen $P\pi = \pi$ und $\sum_{i=1}^{\infty} \pi_i = 1$ erfüllt.

Eine solche stationäre Verteilung stellt also einen Gleichgewichtszustand dar. Wird diese Verteilung zu einem bestimmten Zeitpunkt erreicht, so bleibt diese von dort an konstant. Wir betrachten den Spezialfall, dass die Markowkette endlich ist. In diesem Fall existiert immer eine stationäre Verteilung. Dies ist leicht zu sehen, wenn die obige Bedingung einer stationären Wahrscheinlichkeitsverteilung als Eigenwertproblem betrachtet wird:

$$P\pi = \lambda\pi$$

Folglich ist die stationäre Verteilung π der Eigenvektor zum Eigenwert $\lambda = 1$ und dieser existiert auch, da stochastische Matrizen den Eigenwert 1 besitzen. Der Eigenvektor kann nun skaliert werden, sodass die sich die Elemente zu 1 aufaddieren. Die gefundene stationäre Verteilung ist jedoch nicht notwendigerweise eindeutig, da bei dem Falle von mehreren Eigenvektoren zum Eigenwert 1 einen Vektorraum aufgespannt wird.

Wir spezifizieren die Bedingungen an eine Markowkette, die eine eindeutige stationäre Verteilung hat, nun genauer. Ist die Markowkette irreduzibel, positiv rekurrent und aperiodisch (und damit dann stark ergodisch), so gilt das folgende Theorem, welches außerdem eine Aussage über den Grenzwert der Verteilung angibt:

Theorem 1. Eine Markowkette sei stark ergodisch mit den Zuständen $\{1, 2, \dots\}$ und der Übergangsmatrix P . Dann existiert eine eindeutige stationäre Verteilung $\pi = (\pi_1, \pi_2, \dots)^T$, $P\pi = \pi$, sodass

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_i, \text{ für } i, j = 1, 2, \dots$$

Der Beweis dieser Aussage kann bei Karlin und Taylor (1975) [2] nachgelesen werden.

Weitere Eigenschaften endlicher Markowketten

Das folgende Theorem benennt zentrale Eigenschaften endlicher Markowketten.

Theorem 2. *In einer endlichen Markowkette kann es keine null rekurrenten Zustände geben und nicht alle Zustände können transient sein. Daher ist jede endliche Markowkette positiv rekurrent.*

Wir beweisen diese Ergebnisse kurz unter Zuhilfenahme des folgenden Lemmas (ohne Beweis).

Lemma 1. *Ist j ein transienter Zustand einer Markowkette und i ein beliebiger Zustand der Markowkette, dann gilt:*

$$\lim_{n \rightarrow \infty} p_{ji}^{(n)} = 0$$

Beweis. Wir beweisen nun Theorem 2: Aus Lemma 1 folgt direkt, dass nicht alle Zustände transient sein können, da dann

$$\lim_{n \rightarrow \infty} p_{ji}^{(n)} = 0, \text{ für } i = 1, 2, \dots, N,$$

wobei N die Anzahl der Zustände ist. Dann würde aber gelten:

$$\lim_{n \rightarrow \infty} P^n = 0_{N \times N}$$

Die Matrix P^n soll jedoch eine stochastische Matrix sein, also $\sum_{j=1}^N p_{ji}^{(n)} = 1$. Dann widerspricht

$$\sum_{j=1}^N \lim_{n \rightarrow \infty} p_{ji}^{(n)} = 1$$

jedoch dem obigen Grenzwert.

Wir zeigen nun durch einen Widerspruchsbeweis, dass es keine null rekurrenten Zustände geben kann. Nehmen wir also an, es gibt einen null rekurrenten Zustand $i \in C$, wobei C eine Klasse von Zuständen ist. Es kann gezeigt werden, dass diese Klasse geschlossen ist und alle Zustände null rekurrent sind (s. Korollar 2.2 in [1]). Dann würde aber nach dem bekannten Grenzwertsatz $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 1/\mu_{ii}$ und damit $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$ für alle $i, j \in C$ gelten. Dies wäre jedoch ein Widerspruch zu der Voraussetzung, dass die Submatrix P_C von P eine stochastische Matrix ist (s. obige Argumentation).

Ist die Markowkette nun irreduzibel, so gibt es nur eine Klasse, und da weder alle Zustände transient sein können noch ein null rekurrenter Zustand existiert, müssen folglich alle Zustände positiv rekurrent sein. \square

Eine direkte Folgerung aus obigem Theorem ist es, dass Theorem 1 für endliche Markowketten gilt, wenn diese irreduzibel und aperiodisch sind.

Wir betrachten nun eine Berechnungsmethode für die erwartete Wiederkehrzeit und die erwartete Ersteintrittszeit. Dazu definieren wir die Matrix dieser als:

$$M = (\mu_{ij}) = \begin{pmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{1N} \\ \mu_{21} & \mu_{22} & \dots & \mu_{2N} \\ \vdots & \vdots & \dots & \vdots \\ \mu_{N1} & \mu_{N2} & \dots & \mu_{NN} \end{pmatrix}$$

Anstatt die Wahrscheinlichkeiten $\{f_{ii}^{(n)}\}$ und $\{f_{ji}^{(n)}\}$ zu berechnen, kann ein lineares Gleichungssystem aufgestellt werden. Dabei wird die folgende Beziehung ausgenutzt:

$$\mu_{ji} = p_{ji} + \sum_{k=1, k \neq j}^N p_{ki}(1 + \mu_{jk}) = 1 + \sum_{k=1, k \neq j}^N p_{ki}\mu_{jk}.$$

Die Idee dieses Ansatzes ist, dass wir entweder in einem Schritt von i nach j gelangen oder zunächst einen weiteren Zustand k erreichen. Die erwartete Zeit in diesem Fall ist nun die Zeit von k nach j zu gelangen, wobei der eine Schritt der bereits absolviert wurde, auf das Resultat addiert werden muss. Diese erwarteten Zeiten sind gewichtet mit den Übergangswahrscheinlichkeiten. Diese Gleichungen können in Matrixform aufgeschrieben werden:

$$M = 1_{N,N} + (M - \text{diag}(M))P$$

Es kann gezeigt werden, dass dieses lineare Gleichungssystem eine eindeutige Lösung besitzt.

Geburts- und Todesprozesse

Wir modellieren nun eine Population durch eine diskrete Markowkette. Dabei gibt der Zustand X_n die Größe der Population nach n Schritten an. Diese kann entweder endlich mit $X_n \in \{0, 1, 2, \dots, N\}$ oder unendlich mit $X_n \in \{0, 1, 2, \dots\}$ sein. In jedem Schritt $n \rightarrow n+1$ wird angenommen, dass entweder eine Geburt oder ein Tod eintritt oder die Populationsgröße unverändert bleibt. Ein solcher Schritt muss also ein angemessen kleines Zeitintervall modellieren. Die Wahrscheinlichkeit einer Geburt bzw. eines Todes werden mit b_i bzw. d_i angegeben, wobei das i die Populationsgröße angibt. Wir setzen $b_0 = d_0 = 0$ fest, ebenso wie im Falle einer endlichen Markowkette (bzw. einer beschränkten Population mit maximaler Größe N) $b_N = 0$. Dann sind also die Übergangswahrscheinlichkeiten folgendermaßen definiert:

$$p_{ji} = \text{Prob}\{X_{n+1} = j | X_n = i\} \\ = \begin{cases} b_i & \text{für } j = i + 1 \\ d_i & \text{für } j = i - 1 \\ 1 - (b_i + d_i) & \text{für } j = i \\ 0 & \text{für } j \neq i - 1, i, i + 1 \end{cases}$$

Damit die zugehörige Übergangsmatrix P eine stochastische Matrix ist, muss gelten:

$$\sup_{i \in \{1, 2, \dots\}} \{b_i + d_i\} \leq 1$$

Wir betrachten von nun an den endlichen Fall einer durch N beschränkten Population. Es kann leicht nachvollzogen werden, dass der Zustand 0 positiv rekurrent ist, während alle anderen transient sind. Es existiert eine eindeutige stationäre Verteilung π , wobei $\pi_0 = 1$ und $\pi_i = 0$ für $i = 1, 2, \dots, N$. Außerdem gilt

$$\lim_{n \rightarrow \infty} P(X_n = 0) = \lim_{n \rightarrow \infty} p_0(n) = 1.$$

Die Population stirbt folglich sicher aus. Es bleibt die Frage, wie lange es im Erwartungswert bis zum Aussterben dauert. Dazu sei τ_k die erwartete Dauer für eine Population der Größe k . Dann ist $\tau_0 = 0$, τ_k für $0 < k < N$

$$\tau_k = b_k(1 + \tau_{k+1}) + d_k(1 + \tau_{k-1}) + (1 - (b_k + d_k))(1 + \tau_k).$$

und $\tau_N = d_N(1 + \tau_{N-1}) + (1 - d_N)(1 + \tau_N)$. Diese Differenzengleichungen können als Matrixgleichung $D\tau = c$ interpretiert werden, wobei $\tau = (\tau_0, \tau_1, \dots, \tau_N)^T$, $c = (0, -1, \dots, -1)^T$ und

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ d_1 & -b_1 - d_1 & b_1 & 0 & \cdots & 0 & 0 \\ 0 & d_2 & -b_2 - d_2 & b_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & d_N & -d_N \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ D_1 & D_N \end{pmatrix}$$

Die Submatrix D_N von D ist irreduzibel diagonal dominant. Daraus folgt $\det(D_N) \neq 0$ (ohne Beweis). Weil darüber hinaus $\det(D) = \det(D_N)$ gilt, ist D invertierbar und damit gilt:

$$\tau = D^{-1}c$$

Eine analytische Lösung wurde 1982 von Nisbet und Gurney präsentiert [4]:

$$\tau_m = \begin{cases} 1/d_1 + \sum_{i=2}^N \frac{b_1 \cdots b_{i-1}}{d_1 \cdots d_i}, & m = 1 \\ \tau_1 + \sum_{s=1}^{m-1} \left(\frac{d_1 \cdots d_s}{b_1 \cdots b_s} \sum_{i=s+1}^N \frac{b_1 \cdots b_{i-1}}{d_1 \cdots d_i} \right), & m = 2, \dots, N. \end{cases}$$

Logistische Wachstumsprozesse

Wir betrachten weiterhin allgemeine Geburts- und Sterbeprozesse. Wir wählen dabei die Wahrscheinlichkeiten b_i und d_i so, dass der Prozess logistische Form hat. Man erinnere sich dazu an das deterministische logistische Modell, dem folgende Differentialgleichung zugrunde liegt:

$$\frac{dy}{dt} = ry \left(1 - \frac{y}{K} \right), \quad y(0) = y_0 > 0.$$

Dabei ist der Parameter r die Wachstumsrate und K eine Obergrenze für die Population. Es ist bekannt, dass die eindeutige Lösung $y(t)$ den Grenzwert $\lim_{t \rightarrow \infty} y(t) = K$ besitzt.

Wir charakterisieren nun einen logistischen Wachstumsprozess durch die folgende Gleichung:

$$b_i - d_i = ri(1 - i/K),$$

wobei $i = 0, 1, 2, \dots, N$ mit $N > K$. Es sei angemerkt, dass im Falle von $i = 0$ sowie $i = N$ die Geburtswahrscheinlichkeit gleich der Sterbewahrscheinlichkeit ist. Wir betrachten zwei verschiedene Möglichkeiten b_i und d_i zu wählen:

$$b_i = r \left(i - \frac{i^2}{2K} \right) \quad \text{und} \quad d_i = r \frac{i^2}{2K}, \quad i = 0, 1, 2, \dots, 2K \quad (1)$$

$$b_i = \begin{cases} ri, & i = 0, 1, 2, \dots, N-1 \\ 0, & i \geq N \end{cases} \quad \text{und} \quad d_i = r \frac{i^2}{K}, \quad i = 0, 1, \dots, N \quad (2)$$

Es fallen einige Unterschiede zum deterministischen Modell auf: Die Population wird mit Sicherheit Aussterben (dies war uns bereits durch Erkenntnisse des vorigen Kapitels bekannt) und damit nicht den Grenzwert K erreichen. Darüber hinaus sind auch Populationsgrößen möglich, die K überschreiten.

Quasistationäre Verteilungen

Um die beiden Modelle (das deterministische logistische Wachstum und den logistischen Wachstumsprozess) miteinander zu vereinbaren, schließen wir die Möglichkeit des Aussterbens in letzterem Modell aus. Dazu betrachten wir die folgende bedingte Wahrscheinlichkeit:

$$\begin{aligned} q_i(n) &= P(X_n = i | X_j \neq 0, j = 0, 1, 2, \dots, n-1) \\ &= \frac{p_i(n)}{1 - p_0(n)} \end{aligned}$$

für $i = 1, 2, \dots, N$. Dabei sei $\{X_n\}$ für $n = 0, 1, 2, \dots$ ein allgemeiner Geburts- und Sterbeprozess mit $p_i(n) = P(X_n = i), i = 0, 1, 2, \dots, N$. Die Verteilung $q(n) = (q_1(n), q_2(n), \dots, q_N(n))^T$ ist eine Wahrscheinlichkeitsverteilung, weil

$$\sum_{i=1}^N q_i(n) = \frac{\sum_{i=1}^N p_i(n)}{1 - p_0(n)} = \frac{1 - p_0(n)}{1 - p_0(n)} = 1$$

gilt. Wir betrachten also eine Verteilung, welche bedingt durch das Nichtaussterben in n Schritten ist. Die zugehörige Markowkette kann auf eine stationäre Verteilung untersucht werden und diese wird nicht (wie zuvor) das Aussterben sein. Dazu stellen wir zunächst die Differenzengleichung für $q_i(n+1)$ auf.

$$\begin{aligned} q_i(n+1) &= \frac{p_i(n+1)}{1 - p_0(n+1)} \\ &= \left(\frac{p_i(n+1)}{1 - p_0(n)} \right) \left(\frac{1 - p_0(n)}{1 - p_0(n+1)} \right) \\ &= \left(\frac{p_i(n+1)}{1 - p_0(n)} \right) \left(\frac{1 - p_0(n)}{1 - p_0(n) - d_1 p_1(n)} \right) \end{aligned}$$

Dies kann wiederum umgeformt werden zu:

$$\begin{aligned} q_i(n+1)(1 - d_1 q_1(n)) &= \left(\frac{p_i(n+1)}{1 - p_0(n)} \right) \\ &= b_{i-1} q_{i-1}(n) + (1 - b_i - d_i) q_i(n) + d_{i+1} q_{i+1}(n) \end{aligned}$$

Eine analytische Lösung der stationären Verteilung aus dieser Gleichung kann nicht direkt gefunden werden. Nåsell hat in [5] eine numerische Methode vorgeschlagen. In dieser Ausarbeitung hingegen beschränken wir uns auf eine Approximation der Lösung, indem wir $d_1 = 0$ annehmen. Insbesondere für den logistischen Wachstumsprozess ist dies eine vernünftige Annahme, da hier im Allgemeinen $d_1 \approx 0$ gilt. Damit folgt:

$$\bar{q}_i(n+1) = b_{i-1} \bar{q}_{i-1}(n) + (1 - b_i - d_i) \bar{q}_i(n) + d_{i+1} \bar{q}_{i+1}(n)$$

für $i = 2, \dots, N-1$. Für $i = 1$ bzw. $i = N$ gilt:

$$\bar{q}_1(n+1) = (1 - b_1) \bar{q}_1(n) + d_2 \bar{q}_2(n)$$

und

$$\bar{q}_N(n+1) = b_{N-1} \bar{q}_{N-1}(n) + (1 - d_N) \bar{q}_N(n).$$

Die zu dieser Approximation gehörige Übergangsmatrix \bar{P} ist eine Submatrix von P , wobei $d_1 = 0$.

$$\bar{P} = \begin{pmatrix} 1-b_1 & d_2 & \dots & 0 & 0 \\ b_1 & 1-(b_2+d_2) & \dots & 0 & 0 \\ 0 & b_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1-(b_{N-1}+d_{N-1}) & d_N \\ 0 & 0 & \dots & b_{N-1} & 1-d_N \end{pmatrix}$$

Die Markowkette ist stark ergodisch und hat damit eine eindeutige Wahrscheinlichkeitsverteilung \bar{q}^* , $\bar{P}\bar{q}^* = \bar{q}^*$. Es kann darüber hinaus gezeigt werden, dass diese die folgende Bedingung erfüllt:

$$\bar{q}_{i+1}^* = \frac{b_i \dots b_1}{d_{i+1} \dots d_2} \bar{q}_i^*$$

Literatur

- [1] ALLEN, L. *An Introduction to Stochastic Processes with Applications to Biology*. CRC Press, 2003.
- [2] KARLIN, S., AND TAYLOR, H. *A First Course in Stochastic Processes*. Academic Press, 1975.
- [3] MARTENSTEIN, E. Zeitdiskrete Markowketten. Seminar Mathematische Modelle der Lebenswissenschaften, 2017.
- [4] NISBET, R., AND GURNEY, W. *Modelling Fluctuating Populations*. John Wiley And Sons, 1982.
- [5] NÅSELL, I. On the quasi-stationary distribution of the stochastic logistic epidemic. *Math. Biosci.* 156 (1999), 21–40.