Stochastische Matrizen

Eine Matrix heißt stochastische Matrix, wenn

- sie quadratisch ist und
- alle Einträge zwischen 0 und 1 sind, wobei 0 und 1 auch möglich sind und
- die Spaltensummen in allen Spalten gleich 1 sind.

Beispiele:

$S=\begin{pmatrix}0,2 & 0,5 & 0,1\\ 0,25 & 0,05 & 0,7\\ 0,55 & 0,45 & 0,2\end{pmatrix}$ ist eine stochastische Matrix.

- quadratisch ✓
- alle Einträge zwischen 0 und 1 ✓
- Spaltensummen gleich 1 ✓

$A=\begin{pmatrix}0,2 & 0,8 & 0,1\\ 0,25 & 0,05 & 0,7\\ 0,55 & 0,45 & 0,2\end{pmatrix}$ ist **keine** stochastische Matri

- quadratisch ✓
- alle Einträge zwischen 0 und 1 ✓
- Spaltensummen gleich 1 ✗ (2. Spalte 1,3)

Verwendung

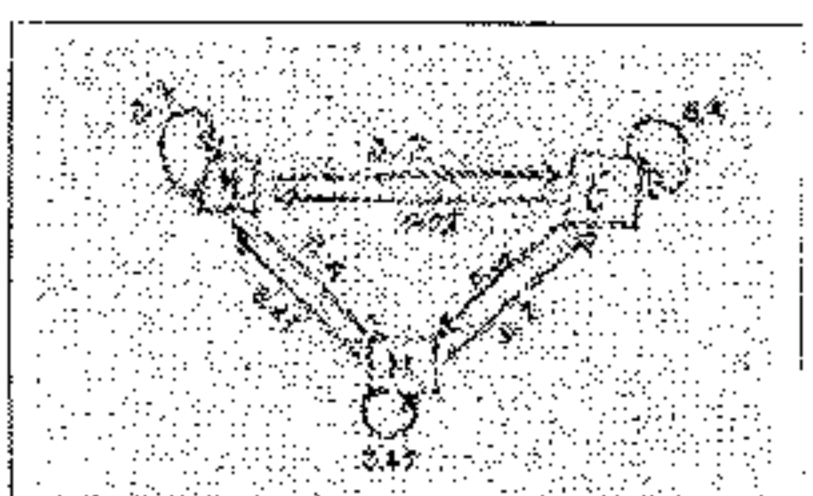
Stochastische Matrizen werden benötigt, um Übergänge zwischen zwei sogenannten Zuständen zu beschreiben. Beispiele:

- Wählerwanderungen zwischen Parteien
- Abstellorte von Leihfahrrädern, Mietwagen oder E-Scootern
- Marktanteile von konkurrierenden Unternehmen

Die Bewegungen werden in Form von Tabellen oder Diagrammen dargestellt:

Beispiele: Standorte eines Mietwagenunternehmens (Zahlen aus „Alles klar? Buch Seite 567)

Von → Nach ↓	Hamburg	Essen	München
Hamburg	0,7	0,05	0,15
Essen	0,2	0,9	0,2
München	0,1	0,05	0,65



E-Mail: Carsten.Vooren@bkcr.info

Rechnung:

Ausgehend von einem Startzustand Startvektor v_0 kann mit Hilfe einer stochastischen Matrix S der nächste Zustand v_1 berechnet werden, indem die Matrix S mit dem Zustandsvektor v_0 multipliziert wird:

Zustand 1: $S \cdot v_0 = v_1$ Zustand 2: $S \cdot v_1 = v_2$ Zustand 3: $S \cdot v_2 = v_3$ Zustand 4: $S \cdot v_3 = v_4$

Die Aneinanderreihung der Zustandsvektoren $v_0, v_1, v_2, v_3, v_4, \ldots$ nennt man eine Markov-Kette.

Beispiel: („Alles klar? Buch Seite 567)

Startzustand $v_0 = \begin{pmatrix} 90 \\ 70 \\ 110 \end{pmatrix}$

Das ist die Anzahl der Autos am 5. Juni in Hamburg (90), Essen (70) und München (110).

Stochastische Übergangsmatrix $S = \begin{pmatrix} 0,7 & 0,05 & 0,15 \\ 0,2 & 0,9 & 0,2 \\ 0,1 & 0,05 & 0,65 \end{pmatrix}$

Zustand 1 (Anzahl der Autos nach einer Woche): $v_1 = S \cdot v_0$

$$S \cdot v_0 = \begin{pmatrix} 0,7 & 0,05 & 0,15 \\ 0,2 & 0,9 & 0,2 \\ 0,1 & 0,05 & 0,65 \end{pmatrix} \cdot \begin{pmatrix} 90 \\ 70 \\ 110 \end{pmatrix} = \begin{pmatrix} 90 \cdot 0,7 + 70 \cdot 0,05 + 110 \cdot 0,15 \\ 90 \cdot 0,2 + 70 \cdot 0,9 + 110 \cdot 0,2 \\ 90 \cdot 0,1 + 70 \cdot 0,05 + 110 \cdot 0,65 \end{pmatrix} = \begin{pmatrix} 83 \\ 103 \\ 84 \end{pmatrix} = v_1$$

Erläuterung der Rechnung am Beispiel der 1. Zeile:

90 Autos stehen in Hamburg, davon werden 70% also $90 \cdot 0,7 = 63$ Autos auch wieder am Ende der Woche in Hamburg zurückgegeben.

70 Autos stehen in Essen, davon werden 5%, also $70 \cdot 0,05 = 3,5$ Autos am Ende der Woche in Hamburg zurückgegeben.

110 Autos stehen in München, davon werden 15%, also $110 \cdot 0{,}15 = 16{,}5$ Autos am Ende der Woche in Hamburg zurückgegeben.

Die Summe ist $63 + 3{,}5 + 16{,}5 = 83$.

Anmerkung: Es kann vorkommen, dass am Ende gerundet werden muss (siehe Zustand 2)!

Zustand 2 (Anzahl der Autos nach zwei Wochen): $v_2 = S \cdot v_1$

$$S \cdot v_1 = \begin{pmatrix} 0{,}7 & 0{,}05 & 0{,}15 \\ 0{,}2 & 0{,}9 & 0{,}2 \\ 0{,}1 & 0{,}05 & 0{,}65 \end{pmatrix} \cdot \begin{pmatrix} 83 \\ 103 \\ 84 \end{pmatrix} = \begin{pmatrix} 83 \cdot 0{,}7 + 103 \cdot 0{,}05 + 84 \cdot 0{,}15 \\ 83 \cdot 0{,}2 + 103 \cdot 0{,}9 + 84 \cdot 0{,}2 \\ 83 \cdot 0{,}1 + 103 \cdot 0{,}05 + 84 \cdot 0{,}65 \end{pmatrix} = \begin{pmatrix} 75{,}85 \\ 126{,}1 \\ 68{,}05 \end{pmatrix} \approx \begin{pmatrix} 76 \\ 126 \\ 68 \end{pmatrix} = v_2$$

Kontrolle der gerundeten Ergebnisse: 76+126+68 = 270 ✓

Zustand 3 (Anzahl der Autos nach drei Wochen): $v_3 = S \cdot v_2$

$$S \cdot v_2 = \begin{pmatrix} 0{,}7 & 0{,}05 & 0{,}15 \\ 0{,}2 & 0{,}9 & 0{,}2 \\ 0{,}1 & 0{,}05 & 0{,}65 \end{pmatrix} \cdot \begin{pmatrix} 76 \\ 126 \\ 68 \end{pmatrix} = \begin{pmatrix} 76 \cdot 0{,}7 + 126 \cdot 0{,}05 + 68 \cdot 0{,}15 \\ 76 \cdot 0{,}2 + 126 \cdot 0{,}9 + 68 \cdot 0{,}2 \\ 76 \cdot 0{,}1 + 126 \cdot 0{,}05 + 68 \cdot 0{,}65 \end{pmatrix} = \begin{pmatrix} 69{,}7 \\ 142{,}2 \\ 58{,}1 \end{pmatrix} \approx \begin{pmatrix} 70 \\ 142 \\ 58 \end{pmatrix} = v_3$$

Kontrolle der gerundeten Ergebnisse: 70+142+58 = 270 ✓

Die Auflistung der Zustände nennt man **Markov-Kette**:

$$v_0, v_1, v_2, v_3, \ldots = \begin{pmatrix} 90 \\ 70 \\ 110 \end{pmatrix}, \begin{pmatrix} 83 \\ 103 \\ 84 \end{pmatrix}, \begin{pmatrix} 76 \\ 126 \\ 68 \end{pmatrix}, \begin{pmatrix} 70 \\ 142 \\ 58 \end{pmatrix}, \ldots$$

Beobachtung: Ohne Eingreifen der Autovermietung werden irgendwann alle Autos in Essen stehen, wenn sich das Kundenverhalten und somit die Wahrscheinlichkeiten aus der stochastischen Matrix nicht ändern.

Übung

$$V_{10} = \begin{pmatrix} 52 \\ 177 \\ 41 \end{pmatrix}$$

$$V_{15} = \begin{pmatrix} 50 \\ 179 \\ 40 \end{pmatrix}$$

$$V_{20} = \begin{pmatrix} 50 \\ 180 \\ 40 \end{pmatrix}$$

CAS: S definieren

V0 definieren

Rechnung $V_{10} = S^{10} \cdot V0$

Fixvektor

Fragestellung: Gibt es einen Zustandsvektor v*, so dass sich am Ende der Woche überall so viele Autos an den drei Standorten befinden wie am Anfang der Woche? Wenn so ein Vektor existiert, nennt man ihn Fixvektor.

Ansatz: $S \cdot v^* = v^*$ muss nach v^* aufgelöst werden.

$S \cdot v^* = v^* \Leftrightarrow S \cdot v^* = E \cdot v^* \mid -E \cdot v^* \Leftrightarrow S \cdot v^* - E \cdot v^* = 0 \mid v^*$ nach rechts ausklammern

$\Leftrightarrow (S - E) \cdot v^* = 0$

Das entstehende lineare Gleichungssystem kann mit dem CAS mit den Befehlen linsolve oder solve gelöst werden.

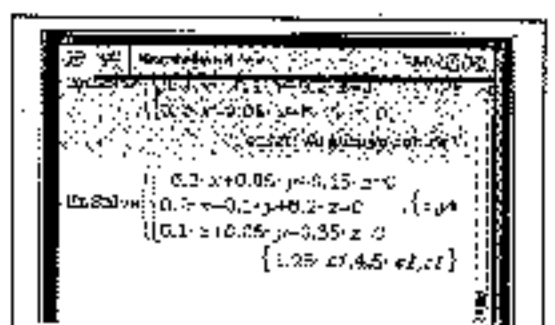
$v^* = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ mit x_1: Anz. Autos in Hamburg, x_2: Anz. Autos in Essen und x_3 = Anz. Autos in München

$$S - E = \begin{pmatrix} 0,7 & 0,05 & 0,15 \\ 0,2 & 0,9 & 0,2 \\ 0,1 & 0,05 & 0,65 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -0,3 & 0,05 & 0,15 \\ 0,2 & -0,1 & 0,2 \\ 0,1 & 0,05 & -0,35 \end{pmatrix}$$

$$(S - E) \cdot v^* = 0 \Leftrightarrow \begin{pmatrix} -0,3 & 0,05 & 0,15 \\ 0,2 & -0,1 & 0,2 \\ 0,1 & 0,05 & -0,35 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -0,30x_1 + 0,05x_2 + 0,15x_3 = 0 \\ 0,20x_1 - 0,10x_2 + 0,20x_3 = 0 \\ 0,10x_1 + 0,05x_2 - 0,35x_3 = 0 \end{pmatrix}$$

Mit CAS:

$$v^* = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1,25 \cdot t \\ 4,5 \cdot t \\ t \end{pmatrix}, t \geq 0$$



In der Praxis nimmt man üblicherweise den Buchstaben t statt c1, wie im CAS c1.

$S \cdot v^* = v^* \quad | -v^*$

$\Leftrightarrow S \cdot v^* - v^* = 0$

$\Leftrightarrow S \cdot v^* - E \cdot v^* = 0$

$\Leftrightarrow (S - E) \cdot v^* = 0$

⋮

Siehe links

$v^* = \begin{pmatrix} 1,25 \cdot t \\ 4,5 \cdot t \\ t \end{pmatrix} \geq 0$

Es muss gelten

$1,25 \cdot t + 4,5 \cdot t + t = 270$ (Anzahl aller Autos)

$\Leftrightarrow t = 40 \Rightarrow v^* = \begin{pmatrix} 50 \\ 180 \\ 40 \end{pmatrix}$

E-Mail: Carsten.Vooren@bkcr.info

Je nach Aufgabe und Bedeutung der Variablen muss darauf geachtet werden, dass die Lösungen positiv oder ganzzahlig sind oder eine bestimmte Vorgabe erfüllen. In dieser Aufgabe hat die Firma 270 Autos, die vermietet werden, das heißt t muss so bestimmt werden, dass die Summe der Autos gleich 270 ist und alle Lösungen ganzzahlig sind:

$1{,}25 \cdot t + 4{,}5 \cdot t + t = 270 \Leftrightarrow 6{,}75 \cdot t = 270 \,|\, :6{,}75 \Leftrightarrow t = 40$

<u>Lösung:</u> In Hamburg müssen $1{,}25 \cdot 40 = 50$, in Essen $4{,}5 \cdot 40 = 180$ und in München 40 Autos stehen. Dann stehen am Ende der Woche in allen Standorten so viele Autos wie zu Beginn der Woche.

Fixvektor: $v^* = \begin{pmatrix} 50 \\ 180 \\ 40 \end{pmatrix}$ *und* $S \cdot v^* = \begin{pmatrix} 0,7 & 0,05 & 0,15 \\ 0,2 & 0,9 & 0,2 \\ 0,1 & 0,05 & 0,65 \end{pmatrix} \cdot \begin{pmatrix} 50 \\ 180 \\ 40 \end{pmatrix} = \begin{pmatrix} 50 \\ 180 \\ 40 \end{pmatrix} = v^*$

<u>Effiziente Berechnung der einzelnen Zustände</u>

Für die Berechnung eines Zustands v_n ist es bei der bisherigen Betrachtung immer notwendig, den vorherigen Zustand v_{n-1} zu kennen: $v_n = S \cdot v_{n-1}$

Bei genauerer Betrachtung stellt man fest, dass man v_n auch nur mit der stochastischen Matrix S und der Startverteilung v_0 bestimmen kann.

$v_1 = S \cdot v_0$

$v_2 = S \cdot v_1 = S \cdot S \cdot v_0 = S^2 \cdot v_0$

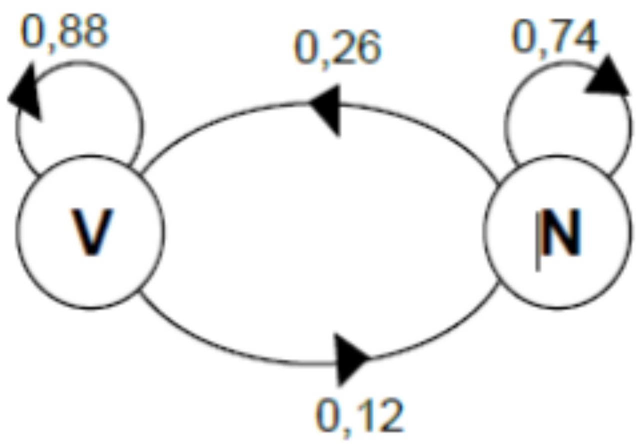
$v_3 = S \cdot v_2 = S \cdot S \cdot v_1 = S \cdot S \cdot S \cdot v_0 = S^3 \cdot v_0$

$v_4 = S \cdot v_3 = S \cdot S \cdot v_2 = S \cdot S \cdot S \cdot v_1 = S \cdot S \cdot S \cdot S \cdot v_0 = S^4 \cdot v_0$

$\Rightarrow$ $\mathbf{v_n = S^n \cdot v_0}$

Aufgabe 1

In der Schulmensa gibt es zwei Essensangebote. Es gibt ein vegetarisches Gericht (V) und ein nichtvegetarisches Gericht (N). Es kommt vor, dass die Schülerinnen und Schüler von Tag zu Tag zwischen den Angeboten wechseln. Das Übergangsdiagramm stellt diesen Wechsel dar.



nach \ von	V	N
V	0,88	0,26
N	0,12	0,74

Am Montag wählen 100 Schüler das vegetarische und 200 Schüler das nichtvegetarische Gericht.

a) Berechnen Sie jeweils die Anzahl der Essen für Dienstag, Mittwoch, Donnerstag und Freitag. Berechnen Sie dabei die Anzahl der Essen für Dienstag „von Hand“.

b) Prüfen Sie, ob es einen Fixvektor v* gibt. Ein anderer Begriff, der auch häufig verwendet wird, ist „Grenzverteilung“. Für die Grenzverteilung oder den Fixvektor gilt $S \cdot v^* = v^*$.

c) Manchmal existiert eine Grenzmatrix G für die stochastische Matrix S und es gilt:

$\lim_{n \to \infty} S^n = G$. Überprüfen Sie, ob eine solche Grenzmatrix existiert, indem Sie verschiedene „hohe“ Potenzen für S berechnen und die Ergebnisse vergleichen.

Stochastische Matrix S

$S = \begin{pmatrix} 0{,}88 & 0{,}26 \\ 0{,}12 & 0{,}74 \end{pmatrix}$ in CAS definieren

Startzustand $v_0 = \begin{pmatrix} 100 \\ 200 \end{pmatrix}$ Mo.

Di.: $v_1 = S \cdot v_0 = \begin{pmatrix} 140 \\ 160 \end{pmatrix}$

Mi.: $v_2 = S^2 \cdot v_0 \approx \begin{pmatrix} 165 \\ 135 \end{pmatrix}$

Do.: $v_3 = S^3 \cdot v_0 = \begin{pmatrix} 180 \\ 120 \end{pmatrix}$

Fr.: $v_4 = S^4 \cdot v_0 = \begin{pmatrix} 190 \\ 110 \end{pmatrix}$

Durch Ausprobieren: Fixvektor: $v^x = \begin{pmatrix} 205 \\ 95 \end{pmatrix}$

~~ao~~ ab v_{11}

Aufgabe 2

Ein E-Scooter-Unternehmen hat die Verteilung der Roller zu Beginn und am Ende eines Tages an drei Hotspots untersucht und folgenden Werte erhalten:

	Von A (Hbf)	Von B (Stadtpark)	Von C (Marktplatz)
nach A (Hbf)	20%	20%	30%
nach B (Stadtpark)	40%	40%	10%
nach C (Marktplatz)	40%	40%	60%

Eine Startverteilung am Morgen des 1. März sieht so aus: A: 200 E-Scooter, B: 200 E-Scooter, C: 200 E-Scooter.

a) Berechnen Sie die Verteilung der E-Scooter nach 1 Tag, nach 10 Tagen, nach 20 Tagen, nach 40 Tagen und nach 50 Tagen.
b) Prüfen Sie, ob eine Grenzverteilung und eine Grenzmatrix existieren und geben Sie diese, falls ja, an.

Aufgabe 3

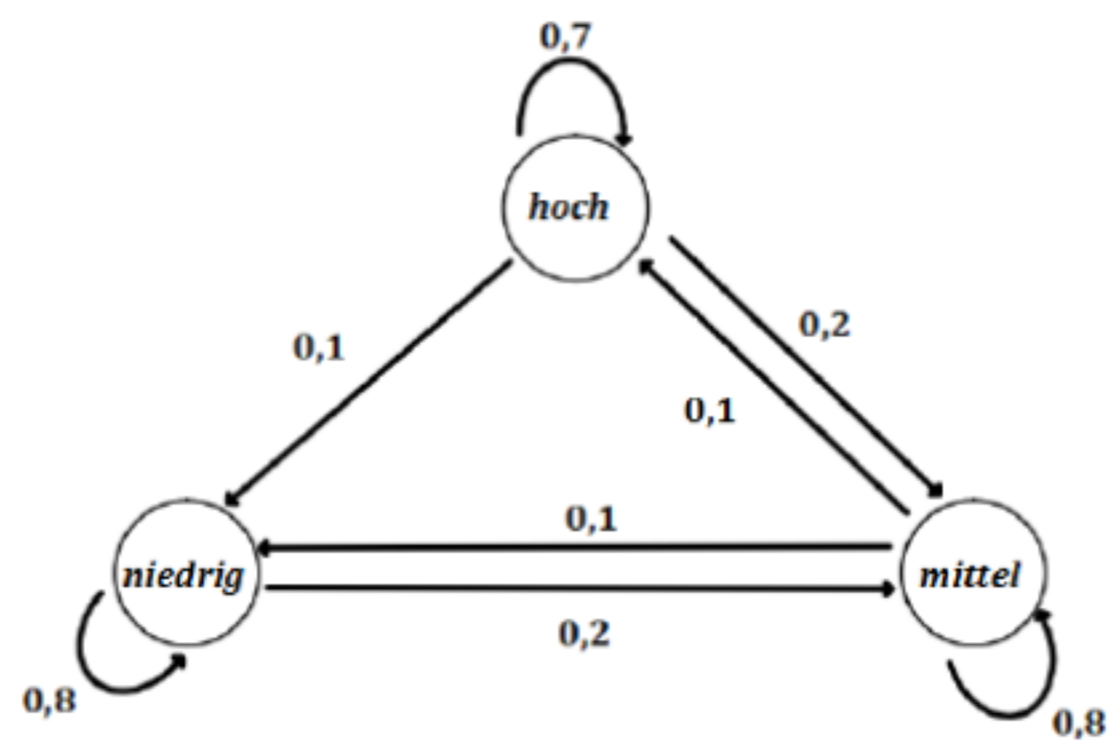
Laut einer Untersuchung des Deutschen Instituts für Touristik blieben im vergangenem Jahr 25 % der Bevölkerung im Urlaub zu Hause (H), 42 % verreisten innerhalb von Deutschland (D) und 33 % verbrachten Ihren Urlaub im Ausland (A). Die untenstehende Tabelle beschreibt das Wechselverhalten der deutschen Urlauber von Jahr zu Jahr.

	Von H	Von D	Von A
nach H	0,3	0,3	0,1
nach D	0,2	0,6	0,5
nach A	0,5	0,1	0,4

a) Erläutern Sie die Bedeutung der Einträge 0,6 und 0,4 im Sachkontext.
b) Berechnen Sie jeweils die Anzahl der Urlauber, die zu Hause bleiben, den Urlaub in Deutschland und im Ausland verbringen für das aktuelle Jahr.
c) Bestimmen Sie die Anzahlen für weitere zwei Jahre.
d) Berechnen Sie die Übergangsmatrix S^2. Vergleichen Sie diese mit S.
e) Bestimmen Sie S^5, S^{10}, S^{15}, S^{20} und beschreiben Sie, was Ihnen auffällt.
f) Bestimmen Sie $S^{-1} \cdot \begin{pmatrix} 40 \\ 35 \\ 25 \end{pmatrix}$ und erläutern Sie die Bedeutung im Sachkontext.

Aufgabe 4

In einem kleinen Land gibt es ca. 1.000.000 Erwerbstätige, die vom Finanzamt jährlich einer der drei Einkommensgruppen 'niedrig', 'mittel' und 'hoch' zugeordnet werden.
Es hat sich im Laufe der Jahre gezeigt, dass es zwischen diesen Gruppen von Jahr zu Jahr Verschiebungen gibt. Das untenstehende Übergangsdiagramm beschreibt diese Verschiebungen von Jahr zu Jahr.

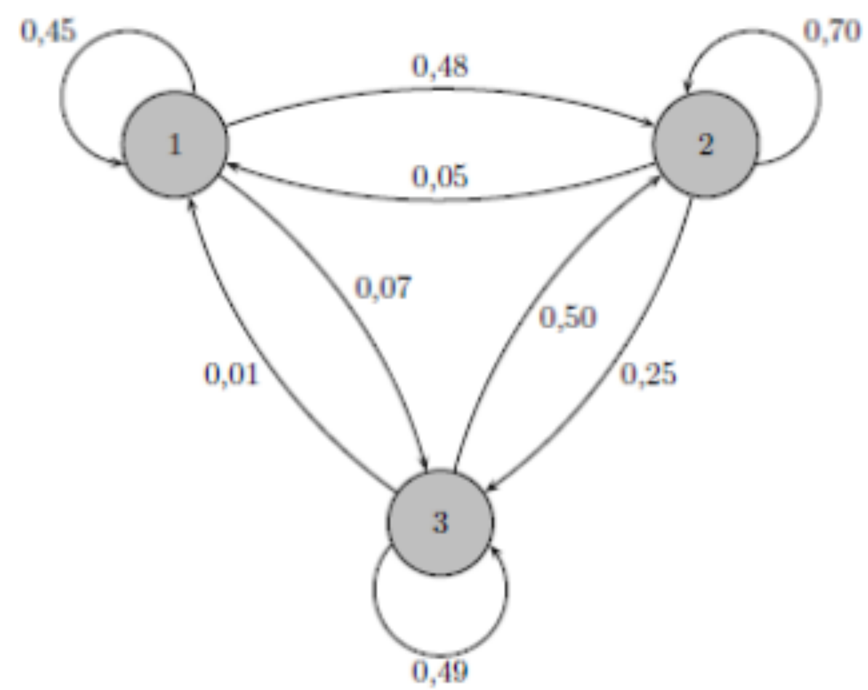


Am Anfang befinden sich 200.000 Erwerbstätige in der Gruppe *niedrig,* 700.000 Erwerbstätige in der Gruppe *mittel* und 100.000 Erwerbstätige in Gruppe *hoch*.

a) Bestimmen Sie die Übergangsmatrix S und begründen Sie, warum S eine stochastische Matrix ist.
b) Berechnen Sie die Anzahl der Erwerbstätigen, die sich im nächsten Jahr in den Einkommensgruppen befinden.
c) Bestimmen Sie die Gruppenstärken für weitere zwei Jahre.
d) Berechnen Sie die Übergangsmatrix S^2. Vergleichen Sie diese mit S.
e) Bestimmen Sie S^5, S^{10}, S^{15}, S^{20} und beschreiben Sie, was Ihnen auffällt.
f) Ermitteln Sie, falls sie existieren, eine Grenzverteilung und eine Grenzmatrix.
g) Bestimmen Sie $S^{-1}\cdot\begin{pmatrix}200000\\700000\\100000\end{pmatrix}$ und erläutern Sie die Bedeutung im Sachkontext.

Aufgabe 5

In einer soziologischen Studie wurde untersucht, in welchem Ausmaß der soziale Status des Vaters den des Sohnes in England und Wales beeinflusst. Im Diagramm sind die Übergänge enthalten.



a) Stellen Sie die Übergangsmatrix S auf und beachten Sie, dass die Spaltenüberschriften der Schicht des Vaters und die Zeilenüberschriften der Schicht des Sohnes entsprechen.
b) Bestimmen Sie die Wahrscheinlichkeit in zwei Generationen von 3 nach 1 aufzusteigen, bzw. von 1 nach 3 abzusteigen.
c) Wie lautet die stationäre Verteilung (oder Grenzverteilung oder Fixvektor)?
d) Untersuchen Sie, ob sich jede Anfangsverteilung der Grenzverteilung nähert.
e) Eine Anfangsverteilung ist durch $v_0 = \begin{pmatrix} 50 \\ 300 \\ 2000 \end{pmatrix}$ gegeben. Ermitteln Sie die zu dieser Anfangsverteilung gehörende Grenzverteilung.