Übungsaufgaben Lösungen

Stochastische Matrizen, Markov-Prozesse

MV5.1

- **Eine N×N-Matrix P heißt stochastisch, wenn ihre Matrixelemente nicht-negativ sind und alle Zeilensummen 1 ergeben.**
  In Formeln: $P_{ij} \geq 0$, $\sum_{j=1}^{N} P_{ij} = 1$ für alle i=1,...,N.
  Wir entwickeln im Folgenden die Grundzüge der Theorie dieser Matrizen. Dazu ist es manchmal ratsam, sich die Lösung der Aufgaben anzusehen, bevor man weitermacht.
- **Eine wichtige Anwendung ist die Theorie der endlichen Markov-Prozesse. Dafür werden die Matrixelemente P_{ij} als Übergangswahrscheinlichkeiten gedeutet. Genauer: Ein System hat endlich viele Zustände, die mit den Zahlen i=1,...N, bezeichnet werden. Wenn das System im Zustand i ist, sei P_{ij} die (bedingte) Wahrscheinlichkeit dafür, dass es im nächsten Zeitschritt in den Zustand j übergeht.**
  Interpretieren Sie die Forderungen $P_{ij} \geq 0$, $\sum_{j=1}^{N} P_{ij} = 1$ für alle i=1,...,N, für diese Wahrscheinlichkeits-Interpretation!
- **Wahrscheinlichkeiten sind reelle Zahlen zwischen 0 und 1, also muss $0 \leq P_{ij} \leq 1$ gelten. Im nächsten Zeitschritt ist das System mit Sicherheit in irgendeinem Zustand $1 \leq j \leq N$, also addiert sich die Gesamtwahrscheinlichkeit $\sum_{j=1}^{N} P_{ij}$ zu 1.**
- **Zeigen Sie: Jede stochastische Matrix hat den Eigenvektor $u = (1, 1, \ldots, 1)^T$ zum Eigenwert 1!**
- **$(P\,u)_i = \sum_{j=1}^{N} P_{ij}\, u_j = \sum_{j=1}^{N} P_{ij} \overset{P \text{ stoch.}}{=} 1$, also P u = u.**
- **Ist E_N eine stochastische Matrix? Wie sind die entsprechenden Übergangswahrscheinlichkeiten zu interpretieren?**
- **Ja. Das System bleibt mit Wahrscheinlichkeit 1 im gleichen Zustand.**
- **Wie lautet die stochastische Matrix für ein System, das mit gleicher Wahrscheinlichkeit in irgendeinen anderen Zustand übergeht?**
- **$P_{ij} = \frac{1}{N-1}$ für $i \neq j$, und $P_{ii} = 0$.**

MV5.2

- **Wir setzen voraus, dass der Übergang in einen neuen Zustand nur vom augenblicklichen Zustand abhängt und nicht von der Vorgeschichte. Das System hat also kein "Gedächtnis". Man nennt dann den Übergang in neue Zustände nach endlich vielen Zeitschritten einen endlichen Markov-Prozess.**
  Wie berechnet man die Wahrscheinlichkeit $P_{ij}^{(2)}$ für einen Übergang von i nach j in zwei Zeitschritten? Bilden diese Wahrscheinlichkeiten wieder eine stochastische Matrix?

- **Wir nennen die Wahrscheinlichkeit für den Übergang i→k→j : W(i→k→j). Dann ist die Wahrscheinlichkeit**
  $P_{ij}^{(2)} = \sum_{k=1}^{N} W(i \to k \to j)$, da das System nach einem Zeitschritt in irgendeinen Zustand k gehen kann. Andererseits ist
  auf Grund der Markov-Annahme W(i→k→j)= $P_{ik}\, P_{kj}$, da die Wahrscheinlichkeiten für die Übergänge i→k und k→j unabhängig voneinander sind und sich multiplizieren.
  Insgesamt gilt also $P_{ij}^{(2)} = \sum_{k=1}^{N} P_{ik}\, P_{kj} = (P^2)_{ij}$. Der Übergang in zwei Zeitschritten ist durch das Quadrat der Matrix P gegeben. P^2 ist wieder stochastisch wegen $P_{ik}\, P_{kj} = (P^2)_{ij} \geq 0$ und $\sum_{j=1}^{N} (P^2)_{ij} = \sum_{j,k=1}^{N} P_{ik}\, P_{kj} = \sum_{k=1}^{N} P_{ik} = 1$.

- **Wie berechnet man die Wahrscheinlichkeit $P_{ij}^{(n)}$ für einen Übergang von i nach j in *n* Zeitschritten? Bilden diese Wahrscheinlichkeiten wieder eine stochastische Matrix?**

- **Analog zur obigen Rechnung ergibt sich $P_{ij}^{(n)} = (P^n)_{ij}$ mit der stochastischen Matrix P^n.**

MV5.3

- **Wir nehmen jetzt an, dass das System zu einem bestimmten Zeitpunkt durch eine diskrete Wahrscheinlichkeitsverteilung charakterisiert ist. Das heißt, p_i, i=1,...,N, sei die Wahrscheinlichkeit dafür, das System im Zustand i zu finden. Natürlich muss dann $0 \leq p_i \leq 1$ und $\sum_{i=1}^{N} p_i = 1$ gelten.**
  Wie lautet unter den gemachten Voraussetzungen die Wahrscheinlichkeitsverteilung $p_i^{(1)}$ nach einem Zeitschritt? Nach *n* Zeitschritten ?

- **Die Wahrscheinlichkeit $p_j^{(1)}$ ergibt sich als Summe der Produkte p_i mit P_{ij}, also $p_j^{(1)} = \sum_{i=1}^{N} p_i\, P_{ij}$.**
  Analog gilt $p_j^{(n)} = \sum_{i=1}^{N} p_i\, (P^n)_{ij}$.

- **Drücken Sie $p_i^{(1)}$ und $p_j^{(n)}$ durch ein Produkt Matrix mal Vektor aus!**

In $p_j^{(1)} = \sum_{i=1}^{N} p_i\, P_{ij}$ wird über den ersten Index von P_{ij} summiert, beim Produkt Matrix mal Vektor aber über den zweiten Index der Matrix. Also müssen wir schreiben: $p_j^{(1)} = \sum_{i=1}^{N} p_i\, P_{ij} = \sum_{i=1}^{N} \left(P^T\right)_{ji}\, p_i$ oder, kompakt, $p^{(1)} = P^T$p.
Analog: $p^{(n)} = P^{T\,n}$p $= P^{nT}$ p.
Die Wahrscheinlichkeitsverteilungen werden also mit der transponierten stochastischen Matrix P^T transformiert.

MV5.4

- **Betrachten Sie folgende Simulation und versuchen Sie zu verstehen, was dabei passiert. Dabei werden die Rechenschritte erläutert, weil wir nicht voraussetzen können, dass alle mit MATHEMATICA vertraut sind.**

```
n = 5; P = Table[RandomReal[], {n}, {n}]; a = Table[Sum[P[[i, j]], {j, n}], {i, n}];
P = Table[P[[i, j]] / a[[i]], {i, n}, {j, n}]; P // MatrixForm
```

$$\begin{pmatrix} 0.17933 & 0.120941 & 0.34277 & 0.112524 & 0.244434 \\ 0.252026 & 0.256939 & 0.181891 & 0.184791 & 0.124354 \\ 0.240461 & 0.349775 & 0.0109782 & 0.379561 & 0.0192243 \\ 0.0819523 & 0.36007 & 0.194197 & 0.0520092 & 0.311771 \\ 0.401193 & 0.121456 & 0.0764949 & 0.29199 & 0.108866 \end{pmatrix}$$

- **Wir haben eine zufällige stochastische Matrix P erzeugt. Man beachte, dass “stochastische Matrix” und “Zufallsmatrix” nicht synonym sind!**

```
Q = Transpose[P];
```

- **Q ist die transponierte stochastische Matrix.**

```
Eigensystem[P]
```

```
{{1., -0.44395, -0.0267853 + 0.129422 i, -0.0267853 - 0.129422 i, 0.105643},
 {{-0.447214, -0.447214, -0.447214, -0.447214, -0.447214},
  {0.428613, -0.0481544, -0.581586, 0.494272, -0.481067},
  {-0.523165 - 0.0852916 i, -0.104608 + 0.165809 i,
   0.209395 - 0.450392 i, 0.561283 + 0. i, -0.0139501 + 0.344465 i},
  {-0.523165 + 0.0852916 i, -0.104608 - 0.165809 i,
   0.209395 + 0.450392 i, 0.561283 + 0. i, -0.0139501 - 0.344465 i},
  {0.0921482, -0.402896, -0.472689, 0.156394, 0.762424}}}
```

- **Die ersten 5 Zahlen sind die Eigenwerte von P. Die folgenden 5 Listen sind die entsprechenden Eigenvektoren. Entspricht dies Ergebnis unserer Theorie?**

- **Ja, wir haben tatsächlich die 1 als Eigenwert von P. Der entsprechende Eigenvektor {-0.447214,-0.447214,-0.447214,-0.447214,-0.447214} ist proportional zu {1,1,1,1,1}, wie wir oben gezeigt haben.**
  Die anderen Eigenwerte sind zum Teil komplex, was aber nicht verwunderlich ist, weil P nicht hermitesch ist.

```
es = Eigensystem[Q]
```

```
{{1., -0.44395, -0.0267853 + 0.129422 i, -0.0267853 - 0.129422 i, 0.105643},
 {{-0.499174, -0.530793, -0.3837, -0.430245, -0.369794},
  {0.445358, -0.046898, -0.478399, 0.572602, -0.492664},
  {-0.324949 - 0.355355 i, 0.66869 + 0. i, -0.435608 + 0.0287673 i,
   -0.0386246 + 0.334398 i, 0.130491 - 0.0078097 i},
  {-0.324949 + 0.355355 i, 0.66869 + 0. i, -0.435608 - 0.0287673 i,
   -0.0386246 - 0.334398 i, 0.130491 + 0.0078097 i},
  {0.274434, -0.837538, -0.00986801, 0.114772, 0.458199}}}
```

- **Die ersten 5 Zahlen sind die Eigenwerte von Q=P^T. Die folgenden 5 Listen sind die entsprechenden Eigenvektoren. Was fällt Ihnen auf? Was vermuten Sie? Können Sie Ihre Vermutung allgemein beweisen?**

- **Die Eigenwerte von Q sind identisch mit den Eigenwerten von P, während die Eigenvektoren verschieden sind.**
  Vermutung: P und P^T haben immer die gleichen Eigenwerte.
  Beweis: $\det(P-\lambda E_N)=\det(P-\lambda E_N)^T = \det(P^T-\lambda E_N)$: Die charakteristischen Polynome von P und P^T sind identisch, also auch die Eigenwerte. Dabei wird nicht benutzt, dass P stochastisch ist.

```
u = es[[2, 1]]; u = u / Sum[u[[i]], {i, n}]
```

```
{0.225492, 0.239776, 0.173329, 0.194355, 0.167047}
```

- **u ist der Eigenvektor von Q zum Eigenwert 1. Er wurde so normiert, dass die Summe der Komponenten gleich 1 ist.**

```
p = Table[RandomReal[], {n}]; p = p / Sum[p[[i]], {i, n}]
```

```
{0.134338, 0.359443, 0.109869, 0.298244, 0.0981054}
```

- **p ist eine zufällige Wahrscheinlichkeitsverteilung.**

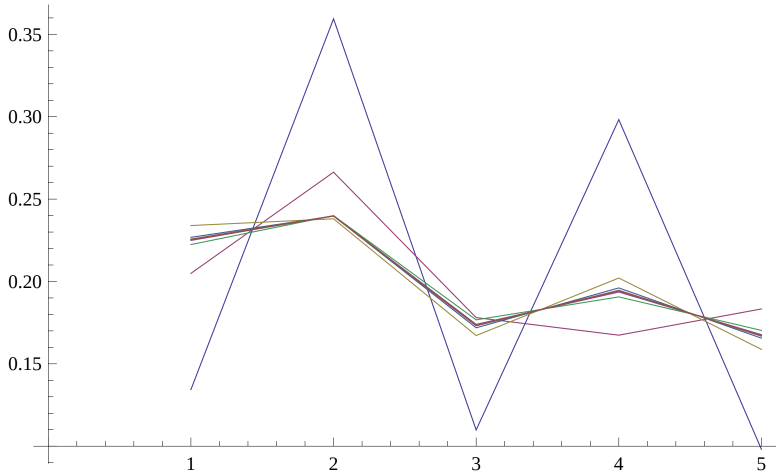
```
pp = Table[Nest[Q.# &, p, m], {m, 0, 10}]
```

```
{{0.134338, 0.359443, 0.109869, 0.298244, 0.0981054},
 {0.2049, 0.266336, 0.178055, 0.167397, 0.183311},
 {0.233946, 0.238031, 0.167163, 0.202087, 0.158773},
 {0.2224, 0.239972, 0.17671, 0.19063, 0.170288},
 {0.226795, 0.239687, 0.171867, 0.196079, 0.165572},
 {0.224901, 0.23984, 0.173966, 0.193589, 0.167703},
 {0.225756, 0.239747, 0.173047, 0.194694, 0.166756},
 {0.225375, 0.239788, 0.173455, 0.194205, 0.167177},
 {0.225544, 0.23977, 0.173274, 0.194422, 0.16699},
 {0.225469, 0.239778, 0.173354, 0.194326, 0.167073},
 {0.225503, 0.239775, 0.173318, 0.194368, 0.167036}}
```

- **pp ist die Liste der Vektoren (Wahrscheinlichkeitsverteilungen) Q^n p, n=0,...10. Der erste Vektor ist gleich p.**
  Was fällt Ihnen auf? Was vermuten Sie?

- **Die Folge der Vektoren Q^n p scheint gegen u, dem Eigenvektor von Q zum Eigenwert 1, zu konvergieren.**
  u wäre dann eine stationäre Wahrscheinlichkeitsverteilung, die von jeder Anfangsverteilung nach unendlich vielen Zeitschritten angenommen wird.
  Das zeigt auch ein Plot der Verteilungen:

```
ListPlot[pp, Joined → True]
```



- **Ein Beweis dieser Vermutung überschreitet den Stoff dieser Veranstaltung. Nur einige Hinweise: Wenn u der Eigenvektor einer Matrix Q mit vollständiger Eigenbasis ist, der zum betragsmäßig größten, nicht entarteten Eigenwert gehört, dann konvergiert $\frac{Q^n p}{\|Q^n p\|}$ für n → ∞ gegen α u (mit einem Faktor $\alpha \neq 0$), wobei p≠0 beliebig ist.**
  In unserem Fall kann man sich die Division durch die Norm sparen, weil alle $Q^n p$ wieder Wahrscheinlichkeitsverteilungen sind. Dann ist auch α=1.
  Eine stochastische Matrix P hat immer 1 als betragsmäßig größten Eigenwert. Allerdings kann der auch entartet sein,
  siehe das obige Beispiel mit E_N als stochastischer Matrix. Wir müssten die entsprechende Voraussetzung, dass 1 nicht entartet ist, also noch hinzufügen, um

unsere Vermutung Q^n p → u zu beweisen.