# §11 Der kontinuierliche Entscheidungsprozess mit Diskontierung

## 11.1 Einführung

Im Kapitel 7 studierten wir den diskreten sequentiellen Prozess mit Diskontierung oder mit einer unbestimmten Dauer. Wir können kontinuierliche Entscheidungsprozess untersuchen mit ähnlichen Elementen unter Verwendung einer analogen Methode.

**11.1.1 Definition:** Definieren wir einen Diskontsatz

$$0 < \alpha < \infty$$

so, dass eine Geldeinheit, die nach einem sehr kurzen Zeitraum $\mathrm{d}\, t$ eingenommen wird, jetzt den Wert

$$1 - \alpha\, \mathrm{d}\, t$$

hat. Diese Definition entspricht einem stetigen Wachstum mit der Wachstumsrate $\alpha$.

Eine andere Interpretation von $\alpha$, die für den Prozess eine unbestimmte Dauer zulässt, ist, dass wir eine Wahrscheinlichkeit von $\alpha\, \mathrm{d}\, t$ dafür haben, dass der Prozess im Intervall $\mathrm{d}\, t$ zuende geht.

Wenn $\mathrm{v}_i(t)$, $i \in E$, der total zu erwartende Erlös des Systems in der Zeit $t$ ist, dann haben wir in Analogie zu Gleichung (9.1) die Behauptung:

**11.1.2 Satz:** Für die total zu erwartende Erlöse gilt das System von Differentialgleichungen

$$\frac{\mathrm{d}}{\mathrm{d}\, t}\, \mathrm{v}_i(t) + \alpha\, \mathrm{v}_i(t) = q_i + \sum_{j=1}^{N} a_{i\,j}\, \mathrm{v}_j(t),\;\; i \in E \tag{11.1}$$

oder in Matrixschreibweise

$$\frac{\mathrm{d}}{\mathrm{d}\, t}\, \mathbf{v}(t) + \alpha\, \mathbf{v}(t) = \mathbf{q} + \mathbf{A}\, \mathbf{v}(t), \tag{11.2}$$

wobei

$$q_i = c_{i\,i} + \textstyle\sum_{j \neq i} a_{i\,j}\, c_{i\,j}.$$

▼

**Beweis:** Wir können den total zu erwartenden Erlös in der Zeit $t + \mathrm{d}\, t$, $\mathrm{v}_i(t + \mathrm{d}\, t)$, $i \in E$, mit $\mathrm{v}_i(t)$ mittels der folgenden Gleichung in Verbindung bringen. Mit dem Satz von der totalen Wahrscheinlichkeit erhält man also

$$\mathrm{v}_i(t + \mathrm{d}\, t) = (1 - \alpha\, \mathrm{d}\, t)[(1 - \textstyle\sum_{j \neq i} a_{i\,j}\, \mathrm{d}\, t)\, (c_{i\,i}\, \mathrm{d}\, t + \mathrm{v}_i(t)) + \sum_{j \neq i} a_{i\,j}\, \mathrm{d}\, t (c_{i\,j} + \mathrm{v}_j(t))],\; i \in E. \tag{11.3}$$

In dieser Gleichung nehmen wir an, dass die Erlöse am ende des Intervalls $\mathrm{d}\, t$ ausgezahlt werden und dass der Prozess bei vorzeitigem Abbruch keinen Erlös hat. Wenn wir die Gleichung

$$a_{i\,i} = -\textstyle\sum_{j \neq i} a_{i\,j}$$

verwenden, können wir die Gleichung (11.3) neu schreiben in der Form

$$\mathrm{v}_i(t + \mathrm{d}\, t) = (1 - \alpha\, \mathrm{d}\, t)[(1 + a_{i\,i}\, \mathrm{d}\, t)\, (c_{i\,i}\, \mathrm{d}\, t + \mathrm{v}_i(t)) + \textstyle\sum_{j \neq i} a_{i\,j}\, \mathrm{d}\, t (c_{i\,j} + \mathrm{v}_j(t))],\; i \in E$$

oder

$$\mathrm{v}_i(t+\mathbb{d}\,t)=(1-\alpha\,\mathbb{d}\,t)[(c_{i\,i}+\sum_{j\neq i}a_{i\,j}\,c_{i\,j})\,\mathbb{d}\,t+\mathrm{v}_i(t)+\sum_{j=1}^{N}a_{i\,j}\,\mathbb{d}\,t\,\mathrm{v}_j(t)],\ i\in E$$

und

$$\mathrm{v}_i(t+\mathbb{d}\,t)=(c_{i\,i}+\sum_{j\neq i}a_{i\,j}\,c_{i\,j})\,\mathbb{d}\,t+\mathrm{v}_i(t)+\sum_{j=1}^{N}a_{i\,j}\,\mathbb{d}\,t\,\mathrm{v}_j(t)-\alpha\,\mathbb{d}\,t\,\mathrm{v}_i(t),\ i\in E,$$

wobei Glieder von höherer Ordnung als $\mathbb{d}\,t$ vernachlässigt wurden.

Die Einführung der Erlösrate aus der Gleichung

$$q_i=c_{i\,i}+\sum_{j\neq i}a_{i\,j}\,c_{i\,j}$$

und eine neue Anordnung ergibt

$$\mathrm{v}_i(t+\mathbb{d}\,t)-\mathrm{v}_i(t)+\alpha\,\mathbb{d}\,t\,\mathrm{v}_i(t)=q_i\,\mathbb{d}\,t+\sum_{j=1}^{N}a_{i\,j}\,\mathbb{d}\,t\,\mathrm{v}_j(t),\ i\in E.$$

Wenn diese Gleichung durch $\mathbb{d}\,t$ dividiert wird und $\mathbb{d}\,t$ gegen Null geht erhalten wir

$$\frac{\mathbb{d}}{\mathbb{d}\,t}\,\mathrm{v}_i(t)+\alpha\,\mathrm{v}_i(t)=q_i+\sum_{j=1}^{N}a_{i\,j}\,\mathrm{v}_j(t),\ \ i\in E.$$

Die Gleichungen (11.1) sind ähnlich den Gleichungen (9.1), und sie reduzieren sich auf diesen, wenn $\alpha=0$ ist. In Vektorform wird aus den Gleichungen (11.1)

$$\frac{\mathbb{d}}{\mathbb{d}\,t}\,\mathbf{v}(t)+\alpha\,\mathbf{v}(t)=\mathbf{q}+\mathbf{A}\,\mathbf{v}(t).$$

---

Da die Gleichung (11.2) eine lineare Differentialgleichung mit konstanten Koeffizienten ist wird sich die Laplace-Transformation als nützlich erweisen.

**11.1.3 Satz:** Für die Laplace-Transformierte $\boldsymbol{v}(s)$ des Vektors $\mathbf{v}(t)$ gilt

(11.4) $$\boldsymbol{v}(s)=\frac{1}{s}\,((s+\alpha)\,\mathbf{I}-\mathbf{A})^{-1}\,\mathbf{q}+((s+\alpha)\,\mathbf{I}-\mathbf{A})^{-1}\,\mathbf{v}(0).$$

Dabei ist **I** die Einheitsmatrix.

▼

**Beweis:**

Transformieren wir die Gleichung (11.2),

$$\frac{\mathbb{d}}{\mathbb{d}\,t}\,\mathbf{v}(t)+\alpha\,\mathbf{v}(t)=\mathbf{q}+\mathbf{A}\,\mathbf{v}(t)$$

dann erhalten wir

$$s\,\boldsymbol{v}(s)-\mathbf{v}(0)+\alpha\,\boldsymbol{v}(s)=\frac{1}{s}\,\mathbf{q}+\mathbf{A}\,\boldsymbol{v}\,(\mathrm{s})$$

oder

$$((s+\alpha)\,\mathbf{I}-\mathbf{A})\,\boldsymbol{v}(s)=\frac{1}{s}\,\mathbf{q}+\mathbf{v}(0)$$

und schliesslich

$$\boldsymbol{v}(s)=\frac{1}{s}\,((s+\alpha)\,\mathbf{I}-\mathbf{A})^{-1}\,\mathbf{q}+((s+\alpha)\,\mathbf{I}-\mathbf{A})^{-1}\,\mathbf{v}(0).$$

---

Man könnte die Gleichung (11.4) und die inverse Transformation verwenden, um $\mathbf{v}(t)$ für einen gegebenen Prozess zu bestimmen. Wie üblich sind wir jedoch an Prozessen von langer Dauer interessiert, so dass uns nur die asymptotische Form von $\mathbf{v}(t)$ für grosse $t$ interessiert.

**11.1.4 Satz:** Für die Laplace-Transformierte $\boldsymbol{v}(s)$ des Vektors $\mathbf{v}(t)$ gilt

(11.5) $$\boldsymbol{v}(s) = \frac{1}{s}\left(\frac{1}{s+\alpha}\,\mathbf{S} + \mathcal{T}(s+\alpha)\right)\mathbf{q} + \left(\frac{1}{s+\alpha}\,\mathbf{S} + \mathcal{T}(s+\alpha)\right)\mathbf{v}(0).$$

Dabei ist $\mathbf{I}$ die Einheitsmatrix.

▼

**Beweis:**

Aus Gleichung (8.11) wissen wir noch, dass

$$(s\,\mathbf{I} - \mathbf{A})^{-1} = \frac{1}{s}\,\mathbf{S} + \mathcal{T}(s)$$

ist, wobei $\mathbf{S}$ die Matrix der Grenzwahrscheinlichkeiten,

$$\mathbf{S} = \begin{pmatrix} \pi_1 & \pi_2 & \cdots & \pi_N \\ \pi_1 & \pi_2 & \cdots & \pi_N \\ \vdots & \vdots & \cdots & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_N \end{pmatrix}$$

und $\mathcal{T}(s)$ eine Matrix aus nur transienten komponenten ist. Daraus folgt, dass

(11.6) $$((s+\alpha)\,\mathbf{I} - \mathbf{A})^{-1} = \frac{1}{s+\alpha}\,\mathbf{S} + \mathcal{T}(s+\alpha),$$

so dass

$$((s+\alpha)\,\mathbf{I} - \mathbf{A})^{-1}$$

nur transiente Komponenten hat. Wir die Gleichung (11.6) in Gleichung (11.4) eingesetzt, dann erhalten wir

$$\boldsymbol{v}(s) = \frac{1}{s}\left(\frac{1}{s+\alpha}\,\mathbf{S} + \mathcal{T}(s+\alpha)\right)\mathbf{q} + \left(\frac{1}{s+\alpha}\,\mathbf{S} + \mathcal{T}(s+\alpha)\right)\mathbf{v}(0).$$

## 11.2 Die Werbestimmung

Nun wollen wir wissen, welche Komponenten von $\mathbf{v}(t)$ nicht verschwinden für grosse $t$. Bezeichnen wir mit

$$\mathbf{v} = \lim_{t\to\infty}\mathbf{v}(t)$$

den Vektor der Gegenwartswerten $\mathrm{v}_i$, $i \in E$.

**11.2.1 Satz:** Der Vektor $\mathbf{v}(t)$ für $t \to \infty$ strebt asymptotisch gegen $\mathbf{v}$,

(11.7) $$\mathbf{v} = (\alpha\,\mathbf{I} - \mathbf{A})^{-1}\,\mathbf{q}.$$

▼

**Beweis:**

Betrachten wir die Gleichung (11.5)

$$\boldsymbol{v}(s) = \frac{1}{s}\left(\frac{1}{s+\alpha}\,\mathbf{S} + \mathcal{T}(s+\alpha)\right)\mathbf{q} + \left(\frac{1}{s+\alpha}\,\mathbf{S} + \mathcal{T}(s+\alpha)\right)\mathbf{v}(0).$$

- Für den ersten Summand von

$$\frac{1}{s}\left(\frac{1}{s+\alpha}\,\mathbf{S} + \mathcal{T}(s+\alpha)\right)\mathbf{q}$$

mit der Parzialbruchzerlegung gilt

$$\frac{1}{s(s+\alpha)}\,\mathbf{S} = \frac{1}{\alpha\, s}\,\mathbf{S} - \frac{1}{\alpha\,(s+\alpha)}\,\mathbf{S}.$$

Für das Urbild erhält man also

$$\frac{1}{\alpha}\,\mathbf{S} - \frac{1}{\alpha}\,e^{-\alpha t}\,\mathbf{S}$$

was für $t \to \infty$ gegen der Funktion

$$\frac{1}{\alpha}\,\mathbf{S}$$

konvergiert.

■ Für den zweiten Summand setzen wir

$$\mathcal{T}\,(s+\alpha) = \sum_i \frac{D_i}{s+\alpha-\alpha_i},$$

wobei $D_i$ die Matrizen, die von $s$ nicht abhängig sind, und $\alpha_i < 0$ die charakteristische Werte sind.

Daraus folgt

$$\frac{1}{s}\,\mathcal{T}\,(s+\alpha) = \sum_i \left(\frac{D_i}{s(\alpha-\alpha_i)} - \frac{D_i}{(s+\alpha-\alpha_i)\,(\alpha-\alpha_i)}\right),$$

diese Laplace-Transformierte entschpricht also der Funktion

$$\sum_i \left(\frac{D_i}{\alpha-\alpha_i}\right) + \sum_i e^{(\alpha_i-\alpha)\,t}\,\frac{D_i}{\alpha-\alpha_i}.$$

Zweite Term ist Null, wenn $t \to \infty$, und

$$\sum_i \left(\frac{D_i}{\alpha-\alpha_i}\right) = \mathcal{T}(\alpha)$$

Also die Transformierte

$$\frac{1}{s}\,\mathcal{T}\,(s+\alpha)\,\mathbf{q}$$

entspricht der Funktion

$$\mathcal{T}\,(\alpha)\,\boldsymbol{q}.$$

■ Alle andere Komponente des Vektors $\mathbf{v}(t)$ gehen gegen 0, wenn $t \to \infty$

Also

$$\mathbf{v} = \lim_{t\to\infty} \mathbf{v}(t) = \left(\frac{1}{\alpha}\,\mathbf{S} + \mathcal{T}\,(\alpha)\right)\mathbf{q}$$

oder mit der Gleichung $(s\,\mathbf{I} - \mathbf{A})^{-1} = \frac{1}{s}\,\mathbf{S} + \mathcal{T}(s)$

$$\mathbf{v} = (\alpha\,\mathbf{I} - \mathbf{A})^{-1}\,\mathbf{q}.$$

Der Vektor **v** stellt die diskontierten zukünftigen Erlöse einer sehr langen Zeitspanne für alle Anfangszustände des Systems dar. Die Gleichung (11.7) zeigt, wie diese Gegenwartswerte abhängen von dem Diskontsatz, der Übergangsraten-Matrix **A** und dem Erlösraten-Vektor **q**.

**11.2.2 Bemerkung:** Gleichung (11.7) kann auch folgendermassen geschrieben werden:

(11.8) $$\alpha\,\mathrm{v}_i = q_i + \sum_{j=1}^{N} a_{i\,j}\,\mathrm{v}_j, \quad i \in E.$$

Wir können die Gleichungen (11.8) lösen, um die Gegenwartswerte irgendeines kontinuierlichen Entscheidungsprozess mit Diskontierung zu bestimmen.

## 11.3 Die Politik-Verbesserung

Wir sind nicht nur an der Auswertung einer gegebenen Politik interessiert, sondern wir möchten auch die Politik mit den höchsten Gegewartswerten in allen Zuständen feststellen. Wir möchten ein Problem lösen können, wie es in Tabelle gestellt wurde, wenn die Diskontierung ein wichtiges Element bedeutet.

| Zustand $i$ | Strategie $k$ | $a_{i1}^k$ | $a_{i2}^k$ | Erlösrate $q_i^k$ |
|---|---|---|---|---|
| 1 Laufende Anlage | 1 (gewöhnliche Unterhalt) | −5 | 5 | 6 |
| 1 | 2 (teurer Unterhalt) | −2 | 2 | 4 |
| 2 Anlage ausser Betrieb | 1 (Reparatur in eigener Werkstatt) | 4 | −4 | −3 |
| 2 | 2 (Reparatur ausserhalb) | 7 | −7 | −5 |

Die Gleichungen (11.8) ergeben eine Wertbestimmung. Wir benötigen nun noch eine Politik-Verbesserung.

Wenn wir in Gleichung (11.1) die Zuwachsrate von $v_i(t)$, $i \in E$, zur Zeit $t$ maximieren wollen, müssen wir

$$q_i^k + \sum_{j=1}^N a_{ij}^k \, v_j(t) - \alpha \, v_i(t)$$

maximieren bezüglich aller Strategien $k$ im $i$-ten Zustand. Wenn wir nur an grossen $t$ interessiert sind, können wir den asymptotischen Gegenwartswert $v_i$, $i \in E$, anstelle von $v_i(t)$ benutzen und so zu der Testgrösse

$$q_i^k + \sum_{j=1}^N a_{ij}^k \, v_j - \alpha \, v_i$$

gelangen.

Da jedoch $v_i$ nicht von $k$ abhängig ist, ist der Ausdruck

$$q_i^k + \sum_{j=1}^N a_{ij}^k \, v_j$$

maximiert, wobei die Gegenwartswerte der vorangehenden Politik einzusetzen sind.

Diese Strategie wird die neue Entscheidung im $i$-ten Zustand. Wenn das Verfahren in sämtlichen Zuständen wiederholt worden ist, ist eine neue Politik bestimmt worden. Diese neue Politik muss grössere Gegenwartswerte aufweisen als die vorhergehende Politik, es sei denn, die beiden Politiken seien identisch. Im letzteren Fall hätte man die optimale Politik gefunden.

Die Wertbestimmung und Politik-Verbesserung sind im Iterationszyklus ersichtlich. Die Regeln für den Beginn und das Ende des Zyklus sind die gleichen wie in den früheren Fällen. Wir werden nun die Eigenschaften des Zyklus beweisen, wobei wir dem Beweisweg für den diskreten Fall (Kapitel 7) folgen.

**11.3.1 Satz:** Der Iterationszyklus für die Ermittlung der optimalen Politik,

**Schritt 0 (Initialisierung)**: Fixieren wir eine Anfangspolitik **d**

**Schritt 1 (Wertbestimmung)**: Man verwende $a_{ij}$ und $q_i$ für die gegebene Politik und löse

$$\alpha \, v_i = q_i + \sum_{j=1}^N a_{ij} \, v_j = 0, \quad i \in E$$

für alle Gegenwartswerte $v_i$, $i \in E$.

**Schritt 2** (**Verbesserung der Politik**): Für jeden Zustand $i \in E$ stelle man die Strategie $k'$ der neuen Politik **d**' fest, welche

$$q_i^k + \sum_{j=1}^N a_{ij}^k \, v_j$$

maximiert unter Verwendung der Gegenwartswerte $v_i$, $i \in E$, der vorhergehenden Politik **d**.

$k'$ ergibt dann die neue Entscheidung im $i$-ten Zustand, $q_i^{k'}$ wird $q_i$, und $a_{ij}^{k'}$ wird $a_{ij}$.

**Schritt 3** (**Prüfung von Konvergenz**): Der Iterationszyklus endet, wenn die Politiken **d** und **d**' bei zwei aufeinanderfolgenden Iterationen stimmen überein. Sonst muss den Schritt 1 mit den neuen Werten wiederholt werden.

▼

**Beweis:** Nehmen wir an, der Iterationszyklus führe zu einer Politik $B$ als Nachfolger der Politik $A$. Da $B$ auf $A$ folgt, wissen wir, dass

$$q_i^B + \sum_{j=1}^{N} a_{i\,j}^B \mathrm{v}_j^A \geq q_i^A + \sum_{j=1}^{N} a_{i\,j}^A \mathrm{v}_j^A$$

ist in jedem Zustand $i \in E$ oder, was gleichbedeutend ist, dass

$$\gamma_i = q_i^B + \sum_{j=1}^{N} a_{i\,j}^B \mathrm{v}_j^A - q_i^A - \sum_{j=1}^{N} a_{i\,j}^A \mathrm{v}_j^A \geq 0$$

für alle $i$, wobei $\gamma_i$ den Zuwachs der Testgrösse darstellt, den die Politik-Verbesserung im $i$-ten Zustand erlangen konnte. Für die verschiedenen Politiken ergibt die Wertbestimmung

$$\alpha\,\mathrm{v}_i^A = q_i^A + \sum_{j=1}^{N} a_{i\,j}^A \mathrm{v}_j^A, \quad i \in E,$$

$$\alpha\,\mathrm{v}_i^B = q_i^B + \sum_{j=1}^{N} a_{i\,j}^B \mathrm{v}_j^B, \quad i \in E.$$

Wenn die ertste Gleichung von der zweiten subtrahiert wird und die Ralation für $\gamma_i$ benutzt wird, um $q_i^B - q_i^A$ zu eliminieren, erhalten wir

$$\alpha(\mathrm{v}_i^B - \mathrm{v}_i^A) = \gamma_i + \sum_{j=1}^{N} a_{i\,j}^B(\mathrm{v}_j^B - \mathrm{v}_j^A)$$

oder

$$\alpha\,\mathrm{v}_i^\Delta = \gamma_i + \sum_{j=1}^{N} a_{i\,j}^B \mathrm{v}_j^\Delta,$$

wobei

$$\mathrm{v}_i^\Delta = \mathrm{v}_j^B - \mathrm{v}_j^A$$

ist. Diese Gleichungen sind dieselben wie unsere Wertbestimmungsgleichungen mit der Ausnahme, dass sie in Differenzen von Gegenwartswerten geschrieben werden. Die Lösung in Vektorform lautet

$$\mathbf{v}^\Delta = (\alpha\,\mathbf{I} - \mathbf{A})^{-1}\,\boldsymbol{\gamma},$$

wobei $\boldsymbol{\gamma}$ den Vektor mit den Komponenten $\gamma_i$ darstellt. Alle Elemente von $(\alpha\,\mathbf{I} - \mathbf{A})^{-1}$ sind nichtnegative wie diejenigen von $(\mathbf{I} - \beta\,\mathbf{P})^{-1}$ im diskreten Fall, wiederum sowohl aus physicalischen als auch aus mathematischen Gründen.

Wenn irgendein

$$\gamma_i > 0$$

ist, muss mindestens ein $\mathrm{v}_i^\Delta$ grösser sein als Null, und kein $\mathrm{v}_i^\Delta$ kann kleiner als Null sein. Die Politik-Verbesserung muss die Gegenwartswerte von mindestens einem Zustand erhöhen und kann für keinen Zustand den Gegenwartswert vermindern.

Ähnlich kann eine Politik $B$ mit höheren Gegenwartswerten als Politik $A$ nicht unentdeckt bleiben wegen der Konvergenz gegen $A$. Das gilt deshalb, weil in einem solchen Fall alle $\gamma_i \leq 0$ sein würden, während mindestens ein $\mathrm{v}_i^\Delta > 0$ sein würde. Diese Situation würde der oben hergeleiteten Relation widersprechen. Wenn der Iterationszyklus gegen eine Politik konvergiert, so hat diese Politik höhere Gegenwartswerte als irgendeine nicht äquivalente Politik.

## 11.4 Ein Beispiel

Wir wollen nun unsere Resultate für die Lösung des sequentiellen Entscheidungsproblems

| Zustand $i$ | Strategie $k$ | $a_{i\,1}^k$ | $a_{i\,2}^k$ | Erlösrate $q_i^k$ |
|---|---|---|---|---|
| 1 Laufende Anlage | 1 (gewöhnliche Unterhalt) | −5 | 5 | 6 |
| 1 | 2 (teurer Unterhalt) | −2 | 2 | 4 |
| 2 Anlage ausser Betrieb | 1 (Reparatur in eigener Werkstatt) | 4 | −4 | −3 |
| 2 | 2 (Reparatur ausserhalb) | 7 | −7 | −5 |

mit $\alpha = \frac{1}{9}$ anwenden. Das können wir so interpretieren, dass die Dauer des Einsatzes des Werkmeisters einer Exponentialverteilung mit einem Mittelwert von 9 Stunden unterliegt, oder wir können an eine Investition denken, bei welcher der Zinsfuss wichtig ist.

Schritt 0. Wie gewohnt werden wir  als Anfangspolitik diejenige wählen, die die Erlösrate maximiert, d.h.

$$\mathbf{d} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; \mathbf{A} = \begin{pmatrix} -5 & 5 \\ 4 & -4 \end{pmatrix}, \; \mathbf{q} = \begin{pmatrix} 6 \\ -3 \end{pmatrix}.$$

**Schritt 1**. Die Wertbestimmungsgleichungen

$$\alpha\, \mathrm{v}_i = q_i + \sum_{j=1}^N a_{i\,j}\, \mathrm{v}_j(t), \;\; i \in E$$

lauten

$$\frac{1}{9}\, \mathrm{v}_1 = 6 - 5\, \mathrm{v}_1 + 5\, \mathrm{v}_2,$$
$$\frac{1}{9}\, \mathrm{v}_2 = -3 + 4\, \mathrm{v}_1 - 4\, \mathrm{v}_2.$$

Ihre Lösung ist

$$\mathrm{v}_1 = \frac{783}{82}, \; \mathrm{v}_2 = \frac{702}{82}.$$

```
Solve[{1 / 9 v1 == 6 - 5 v1 + 5 v2, 1 / 9 v2 == -3 + 4 v1 - 4 v2}, {v1, v2}]
```

$$\left\{\left\{\mathrm{v1} \to \frac{783}{82}, \; \mathrm{v2} \to \frac{351}{41}\right\}\right\}$$

**Schritt 2**. Um eine bessere Politik zu finden, wenden wir die Politik-Verbesserung aus dem Satz 11.3.1 an.

| Zustand | Strategie $d(i) = k$ | $q_i^k + \sum_{j=1}^M a_{i\,j}^k\, \mathrm{v}_j$ |
|---|---|---|
| 1 | 1 | $6 - 5 \cdot \left(\frac{783}{82}\right) + 5 \cdot \left(\frac{702}{82}\right) = \frac{87}{82}$ |
| 1 | 2 | $4 - 2 \cdot \left(\frac{783}{82}\right) + 2 \cdot \left(\frac{702}{82}\right) = \frac{166}{82} \leftarrow$ |
| 2 | 1 | $-3 + 4 \cdot \left(\frac{783}{82}\right) - 4 \cdot \left(\frac{702}{82}\right) = \frac{78}{82}$ |
| 2 | 2 | $-5 + 7 \cdot \left(\frac{783}{82}\right) - 7 \cdot \left(\frac{702}{82}\right) = \frac{157}{82} \leftarrow$ |

$$6-5\left(\frac{783}{82}\right)+5\left(\frac{702}{82}\right)$$
$$4-2\left(\frac{783}{82}\right)+2\left(\frac{702}{82}\right)$$
$$-3+4\left(\frac{783}{82}\right)-4\left(\frac{702}{82}\right)$$
$$-5+7\left(\frac{783}{82}\right)-7\left(\frac{702}{82}\right)$$

$$\frac{87}{82}$$

$$\frac{83}{41}$$

$$\frac{39}{41}$$

$$\frac{157}{82}$$

**Schritt 3**.

$$\begin{pmatrix} 2 \\ 2 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Schritt 0. Die zweite Strategie in jedem Zustand bildet eine bessere Politik, so dass nun

$$\mathbf{d} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \mathbf{A} = \begin{pmatrix} -2 & 2 \\ 7 & -7 \end{pmatrix}, \mathbf{q} = \begin{pmatrix} 4 \\ -5 \end{pmatrix}.$$

**Schritt 1**. Die Wertbestimmungsgleichungen

$$\alpha \, \mathrm{v}_i = q_i + \sum_{j=1}^{N} a_{i\,j} \, \mathrm{v}_j(t), \;\; i \in E$$

lauten

$$\frac{1}{9} \, \mathrm{v}_1 = 4 - 2 \, \mathrm{v}_1 + 2 \, \mathrm{v}_2,$$
$$\frac{1}{9} \, \mathrm{v}_2 = -5 + 7 \, \mathrm{v}_1 - 7 \, \mathrm{v}_2.$$

Ihre Lösung ist

$$\mathrm{v}_1 = \frac{1494}{82}, \, \mathrm{v}_2 = \frac{1413}{82}.$$

```
Solve[{1/9 v1 == 4 - 2 v1 + 2 v2, 1/9 v2 == -5 + 7 v1 - 7 v2}, {v1, v2}]
```

$$\left\{\left\{\mathrm{v1} \to \frac{747}{41}, \, \mathrm{v2} \to \frac{1413}{82}\right\}\right\}$$

Man beachte, dass die Gegenwartswerte sich wiederum erhöht haben. Die Politik-Verbesserung wird wiederum durchgeführt mit den Resultaten

| Zustand | Strategie $d(i) = k$ | $q_i^k + \sum_{j=1}^M a_{ij}^k \mathrm{v}_j$ |
|---|---|---|
| 1 | 1 | $\frac{87}{82}$ |
| 1 | 2 | $\frac{166}{82}$ ← |
| 2 | 1 | $\frac{78}{82}$ |
| 2 | 2 | $\frac{157}{82}$ ← |

Da

$$\mathrm{v}_1 - \mathrm{v}_2$$

unverändert geblieben ist, sind die Werte für die Testgrösse due gleichen wie in diese Tabelle.

**Schritt 3**. Die Politik

$$\mathbf{d} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

wurde zweimal nacheinander gefunden. Somit ist sie die optimale Politik. Sie hat höhere Gegenwartswerte in allen Zuständen als irgendeine andere Politik. Auch wenn die zu erwartende Dauer des Prozesses nur 9 Stunden beträgt, sollte der Werkmeister die kostspielige Unterhaltmethode und auswärtige Reparaturarbeiten wählen.

## 11.5 Vergleich mit dem diskreten Fall

Im diskreten sequentiellen Entscheidungsprozess mit Diskontierung haben wir folgende Wertbestimmungsgleichungen:

(7.9) $$\mathrm{v}_i = q_i + \beta \sum_{j=1}^N p_{ij} \mathrm{v}_j, \quad i \in E = \{1, 2, \ldots, N\}, \quad n \in \mathbb{N}.$$

Wenn wir für dieses Problem ein Computerprogramm geschrieben haben, möchten wir wissen, ob ein solches Programm auch im kontinuierlichen Fall brauchbar ist. Im kontinuierlichen Fall lauten die analogen Gleichungen

(11.8) $$\alpha \, \mathrm{v}_i = q_i' + \sum_{j=1}^N a_{ij} \mathrm{v}_j, \quad i \in E,$$

wobei $q_i'$ benutzt wurde, um den kontinuierlichen vom diskreten Fall zu unterscheiden. Wir können

$$a_{ij} = p_{ij} - \delta_{ij}$$

setzen und dann die Gleichungen (11.8) in der Form

$$\alpha \, \mathrm{v}_i = q_i' + \sum_{j=1}^N \left(p_{ij} - \delta_{ij}\right) \mathrm{v}_j$$

oder

$$(1 + \alpha) \, \mathrm{v}_i = q_i' + \sum_{j=1}^N p_{ij} \mathrm{v}_j$$

schreiben. Daraus folgt

$$\mathrm{v}_i = \frac{1}{1+\alpha} q_i' + \frac{1}{1+\alpha} \sum_{j=1}^N p_{ij} \mathrm{v}_j.$$

Wenn wir

$$\beta = \frac{1}{1+\alpha}$$

und

$$q_i = \frac{1}{1+\alpha} q_i'$$

setzen, dann erhalten wir

$$v_i = q_i + \beta \sum_{j=1}^{N} p_{i\,j}\, v_j, \quad i \in E,$$

also ein Gleichungssystem von der gleichen Form wie im diskreten Fall. Somit können, wenn wir ein kontinuierliches Problem haben, das bestimmt ist durch $\alpha$, **q**' und **A**, das Programm für das diskrete Problem, das gegeben ist durch $\beta$, **q** und **P**, benützen, indem wir die Transformationen durchführen:

$$\beta = \frac{1}{1+\alpha}, \; \mathbf{q} = \beta\, \mathbf{q}', \; \mathbf{P} = \mathbf{A} + \mathbf{I}.$$

In der Politik-Verbesserung für den diskreten Fall ist die Testgrösse

$$q_i^k + \beta \sum_{j=1}^{N} p_{i\,j}^k\, v_j.$$

Für den ko ntinuierlichen Fall lauten es

$$q'^k_i + \sum_{j=1}^{N} \left(p_{i\,j}^k - \delta_{i\,j}\right) v_j,$$

wobei

$$a_{i\,j}^k = p_{i\,j}^k - \delta_{i\,j}$$

ist. Wir haben nun einen Ausdruck, der äquivalent ist zu

$$q'^k_i + \beta \sum_{j=1}^{N} p_{i\,j}^k\, v_j,$$

da $v_j$ von $k$ unabhängig ist. Wenn

$$q'^k_i = \frac{1}{\beta}\, q_i^k$$

gesetzt wird, wobei

$$\beta = \frac{1}{1+\alpha}$$

ist, haben wir

$$\frac{1}{\beta}\, q_i^k + \sum_{j=1}^{N} p_{i\,j}^k\, v_j,$$

und dies ist natürlich proportional zu

$$q_i^k + \beta \sum_{j=1}^{N} p_{i\,j}^k\, v_j,$$

der Testgrösse für den diskreten Fall. Somit erlaubt uns die gleiche Transformation, ein für den diskreten Fall geschriebenes Programm für die Lösung der Wertbestimmung zu benützen und ein Programm für die Politik-Verbesserung, das auf dem diskreten Prozess basiert, anzuwenden.

Wir sehen, dass bei geeigneten Transformationen ein einziges Programm für den diskreten wie kontinuierlichen Fall mit Diskontierung genügt. Da wir die gleiche Tatsache früher schon für Fälle ohne Diskontierung gezeigt haben, ist es klar, dass der kontinuierliche Entscheidungsprozess, mit oder ohne Diskontierung, rechnerisch äquivalent zu seinem diskreten Gegenstück ist.