stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE BESLISKUNDE BC 1/70 DECEMBER

A. HORDIJK en H.C. TIJMS
COLLOQUIUM MARKOV-PROGRAMMERING

januari 1970 - december 1970

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

Inhoud

Hoofdstuk 1. Inleiding.

1.1 HOWARD's iteratiemethodes uit de Markov-programmering.

In de Markov-programmeringsmodellen is sprake van een physisch systeem dat gecontroleerd wordt door een beheerder. Deze beheerder kan de ontwikkeling van de toestand van het systeem beïnvloeden door beslissingen (akties) te nemen. In de te beschouwen modellen is het beslissingsmechanisme bekend en het proces dat het verloop van de toestand beschrijft. Voorts is de opbrengst- en kostenstructuur van de beslissingsproblemen gegeven. Deze vage beschrijving zullen wij later voor voor de specifieke modellen nader uitwerken.

In de Markov-programmeringsmodellen kunnen wij het volgende onderscheid maken.

a) De mogelijke toestanden van het systeem zijn uitsluitend van belang op discrete tijdstippen $t = 0, 1, 2, \ldots$ en alleen op deze tijdstippen kan de beslisser ingrijpen.
b) De toestand van het systeem wordt op ieder moment geobserveerd en de beslisser kan op elk moment ingrijpen.

Wij zullen ons in de rest van deze paragraaf tot het volgende model beperken: Het systeem kan slechts een eindig aantal toestanden aannemen. De toestanden zijn genummerd als $i = 1, \ldots, N$. De toestand van het systeem wordt uitsluitend geobserveerd op de equidistant gelegen tijdstippen $t = 0, 1, \ldots$. Voor iedere toestand i is een eindige verzameling $A(i)$ van toegelaten beslissingen (akties) gegeven. Als in toestand i de beslissing a wordt genomen, dan wordt een direkte opbrengst $r(i,a)$ verkregen en is met kans $p(j|i,a)$ de toestand op het volgende beslissingstijdstip gelijk aan j. De opbrengsten $r(i,a)$ zijn eindig. De som van de $p(j|i,a)$ over j is 1 voor alle i en a.

Wij zullen hoofdzakelijk beslissingsproblemen beschouwen waarin het systeem gedurende een onbegrensd aantal perioden wordt beheerd.

Laat F de speciale klasse van strategieën zijn, welke aan iedere toestand één en slechts één beslissing toewijzen, die niet afhangt van het beschouwde beslissingstijdstip t. In hoofdstuk 2 zullen wij algemenere klassen van strategieën beschouwen en wij zullen daar aantonen dat in vele gevallen men zich uiteindelijk kan beperken tot de klasse F. Een strategie uit F wordt meestal genoteerd of als f^{∞} of als f. In deze paragraaf zullen wij een strategie uit F aanduiden door $f(f^{*},f_n)$.

Als wij de opbrengsten verdisconteren met een factor β, $0 \leq \beta < 1$, dan is op tijdstip 0 de waarde van een op tijdstip n verkregen opbrengst r gelijk aan $\beta^n r$ ($\beta = 1/(1+\rho)$, waarbij ρ een vaste rentefactor is).

In paragraaf 1.1.1 worden de opbrengsten verdisconteerd met een vaste factor β, $0 \leq \beta < 1$ en formuleren wij een iteratiemethode die na een eindig aantal stappen leidt tot een strategie uit F waarvoor de <u>totale verwachte opbrengst</u> over de tijdstippen t = 0, 1, ... maximaal is, ongeacht de begintoestand van het systeem.

In paragraaf 1.2.1 verdisconteren wij de opbrengsten niet. Wij formuleren dan een iteratiemethode die convergeert naar een strategie uit F waarvoor de <u>verwachte gemiddelde opbrengst per tijdseenheid</u>, genomen over de tijdstippen t = 0, 1, ... , maximaal is voor elke begintoestand.

1.1.1 <u>De totale verwachte verdisconteerde opbrengst als kriterium.</u>

Stel ter afkorting voor elke $f \in F$

$$p_{ij}(f) = p(j|i,f(i)), \tag{1.1}$$

waarbij f(i) de beslissing is die strategie f in toestand i voorschrijft. Laat P(f) de matrix zijn met als elementen de overgangskansen $p_{ij}(f)$. Definieer voor alle $f \in F$, i, j = 1, ..., N

$$p_{ij}^{(0)}(f) = \begin{cases} 1 & \text{als } j = i \\ 0 & \text{als } j \neq i \end{cases} \tag{1.2}$$

en

(1.3) $$p_{ij}^{(n)}(f) = \sum_{k=1}^{N} p_{ik}(f)\, p_{kj}^{(n-1)}(f) \qquad \text{voor } n \geqq 1.$$

Voor $n \geqq 1$ geldt dat $p_{ij}^{(n)}(f)$ de kans is dat het systeem op tijdstip n in toestand j is, gegeven dat strategie f wordt toegepast en dat het systeem op t = 0 in toestand i is.

Laten wij $v_i(n;f;\beta)$ definieren als de totale verwachte verdisconteerde opbrengst over de eerste n beslissingstijdstippen gegeven dat strategie $f \in F$ wordt toegepast en dat het systeem op t = 0 in toestand i is. Dan geldt

(1.4) $$v_i(n;f;\beta) = \sum_{k=0}^{n-1} \sum_{j=1}^{N} \beta^k\, p_{ij}^{(k)}(f)\, r(j,f(j))$$

$$\text{voor } i = 1, \ldots, N;\ n \geqq 1.$$

Aangezien

(1.5) $$\lim_{n\to\infty} \beta^n\, p_{ij}^{(n)}(f) = 0 \qquad \text{voor alle } i, j = 1, \ldots, N$$

en de opbrengstfunktie r(i,a) begrensd is, geldt dat $\lim_{n\to\infty} v_i(n;f;\beta)$ bestaat en eindig is. Deze limiet is gelijk aan

(1.6) $$v_i(f;\beta) \overset{\text{def}}{=} \lim_{n\to\infty} v_i(n;f;\beta) = \sum_{k=0}^{\infty} \sum_{j=1}^{N} \beta^k\, p_{ij}^{(k)}(f) r(j,f(j))$$

$$\text{voor } i = 1, \ldots, N.$$

Met behulp van (1.2) en (1.3) volgt op eenvoudige wijze uit (1.6) dat

(1.7) $$v_i(f;\beta) = r(i,f(i)) + \beta \sum_{j=1}^{N} p_{ij}(f)\, v_j(f;\beta)$$

$$\text{voor } i = 1, \ldots, N.$$

Aangezien $0 \leqq \beta\, p_{ij}(f) < 1$ voor alle i en j, heeft de matrix $I - \beta P(f)$ een inverse (paragraaf 1.3, stelling 1.3.2). Dus het stelsel lineaire vergelijkingen (1.7) heeft precies één oplossing. Wij zien dus dat voor elke strategie $f \in F$ de totale verwachte verdisconteerde opbrengst $v_i(f;\beta)$ gevonden kan worden door N lineaire vergelijkingen in N onbekenden op te lossen.

In hoofdstuk 2 tonen wij aan dat een strategie $f^* \in F$ bestaat, waarvan geldt

$$(1.8) \qquad v_i(f^*;\beta) \geqq v_i(f;\beta) \qquad \text{voor } i = 1, \ldots, N; \text{ alle } f \in F.$$

Wij zullen nu een iteratiemethode geven die naar een dergelijke strategie convergeert in een eindig aantal stappen. Het eerste deel van elke iteratiestap bestaat uit het oplossen van een stelsel lineaire vergelijkingen. Het tweede deel van elke deel van elke iteratiestap verkrijgen wij door het volgende probleem te beschouwen. Stel eens dat de beslisser op t = 0 vrij is om een beslissing te kiezen, maar dat hij op de volgende tijdstippen t = 1, 2, ... strategie $f \in F$ moet toepassen. Als de beslisser op tijdstip 0 in toestand i de beslissing a neemt en daarna strategie f toepast, dan is de totale verwachte verdisconteerde opbrengst gelijk aan

$$(1.9) \qquad r(i,a) + \beta \sum_{j=1}^{N} p(j|i,a)\, v_j(f;\beta).$$

De beslisser kiest uiteraard die beslissing a waarvoor (1.9) maximaal is. Als wij nu voor elke toestand i een beslissing a bepalen waarvoor (1.9) maximaal is, dan vinden wij een strategie $f_1 \in F$ met de eigenschap (zie hoofdstuk 2)

$$(1.10) \qquad v_i(f_1;\beta) \geqq v_i(f;\beta) \qquad \text{voor } i = 1, \ldots, N.$$

Wij zien dus dat uitgaande van een willekeurige strategie $f \in F$ wij deze strategie kunnen verbeteren tot een strategie $f_1 \in F$ door eerst het stelsel lineaire vergelijkingen (1.7) op te lossen en vervolgens de grootheid (1.9) naar $a \in A(i)$ te maximaliseren. Wij kunnen nu HOWARD's iteratiemethode voor het verdisconteringsmodel formuleren. De n^{de} stap uit deze iteratiemethode luidt als volgt (de eerste stap beginnen wij met een willekeurige strategie $f_0 \in F$):

n^{de} stap

a) Zij $f_{n-1} \in F$ de strategie verkregen aan het eind van de $(n-1)^{ste}$ stap. Bepaal de unieke oplossing van het stelsel lineaire vergelijkingen.

$$v_i(f_{n-1};\beta) = r(i,f_{n-1}(i)) + \beta \sum_{j=1}^{N} p_{ij}(f_{n-1})\, v_j(f_{n-1};\beta) \tag{1.11}$$

voor $i = 1, \ldots, N$.

b) Bepaal voor elke $i = 1, \ldots, N$ een beslissing $a \in A(i)$ welke

$$r(i,a) + \beta \sum_{j=1}^{N} p(j|i,a)\, v_j(f_{n-1};\beta) \tag{1.12}$$

maximaal maakt (terwille van de convergentie spreken wij hierbij af dat wij $a = f_{n-1}(i)$ kiezen als deze beslissing (1.12) maximaliseert). Op deze wijze vinden wij strategie $f_n \in F$.

Einde van de n^{de} stap.

In hoofdstuk 2 zullen wij aantonen dat voor elke $n \geq 1$ geldt dat $v_i(f_n;\beta) \geq v_i(f_{n-1};\beta)$ is voor alle i. Voorts bewijzen wij in hoofdstuk 2 dat bovenstaande iteratiemethode na een eindig stappen, zeg m, leidt tot $f_m = f_{m-1}$. De strategie $f^* = f_m$ voldoet dan aan $v_i(f^*;\beta) \geq v_i(f;\beta)$ voor alle i en alle $f \in F$.

1.2 De verwachte gemiddelde opbrengst per tijdseenheid als kriterium.

De stochastische variable $\underline{r}_i(n;f)$ definieren wij als de opbrengst op het n^{de} beslissingstijdstip, gegeven dat strategie $f \in F$ wordt toegepast en dat het systeem op $t = 0$ in toestand i is. De opbrengsten worden niet verdisconteerd. De verwachte opbrengst over de eerste n beslissingstijdstippen, gegeven dat strategie $f \in F$ wordt toegepast en dat i de begintoestand is, wordt dus gegeven door

$$(1.13) \qquad v_i(n;f) = \sum_{k=0}^{n-1} E\underline{r}_i(k;f) = \sum_{k=0}^{n-1} \sum_{j=1}^{N} p_{ij}^{(k)}(f)\; r(j,f(j)).$$

Met behulp van de relaties (1.2) en (1.3) volgt op eenvoudige wijze uit (1.13) dat

$$(1.14) \qquad v_i(n+1;f) = r(i,f(i)) + \sum_{j=1}^{N} p_{ij}(f)\; v_j(n;f)$$

$$\text{voor } 1 \leq i \leq N;\; n \geq 0,$$

waarbij

$$(1.15) \qquad v_i(0;f) \overset{\text{def}}{=} 0 \qquad \text{voor } i = 1, \ldots, N.$$

De gemiddelde verwachte opbrengst per tijdseenheid genomen over de eerste n beslissingstijdstippen, gegeven strategie $f \in F$ en begintoestand i, is gelijk aan

$$(1.16) \qquad \frac{1}{n}\, v_i(n;f) = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{j=1}^{N} p_{ij}^{(k)}(f)\; r(j,f(j)).$$

In paragraaf 1.2, stelling 1.2.12 tonen wij aan dat

$$(1.17) \qquad p^{*}_{ij}(f) \overset{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} p^{(k)}_{ij}(f)$$

bestaat uit alle i en j. Uit (1.16) en (1.17) volgt nu dat, gegeven strategie $f \in F$ en begintoestand i, de gemiddelde verwachte opbrengst per tijdseenheid genomen over $t = 0, 1, \dots,$ gegeven wordt door

$$(1.18) \qquad \lim_{n\to\infty} \frac{1}{n} v_i(n;f) = x_i(f) \overset{\text{def}}{=} \sum_{j=1}^{N} p^{*}_{ij}(f)\, r(j,f(j)).$$

Uit stelling 1.2.23 (paragraaf 1.2) volgt dat de stochast

$$(1.19) \qquad \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \underline{r}_i(k;f)$$

met kans 1 bestaat. Passen wij de stelling van Lebesque over gemajoreerde convergentie uit de maattheorie toe, dan vinden wij

$$(1.20) \qquad E\left(\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \underline{r}_i(k;f)\right) = \lim_{n\to\infty} E\left(\frac{1}{n} \sum_{k=0}^{n-1} \underline{r}_i(k;f)\right) =$$

$$= \lim_{n\to\infty} \frac{1}{n} v_i(n;f) = x_i(f)$$

$$\text{voor } i = 1, \dots, N.$$

Dus ook de verwachte gemiddelde opbrengst per tijdseenheid genomen over $t = 0, 1, \dots$ is gelijk aan $x_i(f)$, ingeval strategie $f \in F$ wordt toegepast en i de begintoestand is. Als de begintoestand i tot een priemfuik (zie def. 1.2.5 in paragraaf 1.2) behoort van de Markovketen met $P(f)$ als matrix van overgangskansen, dan kunnen wij zelfs nog een sterkere interpretatie aan $x_i(f)$ geven. Dan geldt namelijk dat de stochastische variabele (1.19) (= de werkelijke gemiddelde opbrengst per tijdseenheid) met kans 1 de waarde $x_i(f)$ aanneemt (zie stelling 1.2.23). Wij merken nog op dat uit stelling 1.2.20 en definitie(1.18)volgt dat voor toestanden k en h uit een zelfde priemfuik van P(f) geldt

$$(1.21) \qquad x_h(f) = x_k(f).$$

Dus als de verzameling van alle toestanden van de Markov-keten met P(f) als matrix van overgangskansen een priemfuik vormt, zoals in praktische problemen veelal het geval is, dan geldt dat de verwachte gemiddelde opbrengst per tijdseenheid niet afhangt van de begintoestand en de de gemiddelde opbrengst per tijdseenheid met kans 1 daaraan gelijk is.

In hoofdstuk 2 tonen wij aan dat een strategie $f^* \in F$ bestaat met de eigenschap dat

$$(1.22) \qquad x_i(f^*) \geqq x_i(f) \qquad \text{voor } i = 1, \ldots, N; \text{ alle } f \in F.$$

Wij zullen nu een iteratiemethode geven die na een eindig stappen convergeert naar een strategie $f^* \in F$ die aan (1.22) voldoet. Alvorens wij de iteratiemethode formuleren, zullen wij laten zien hoe op een heuristische wijze de structuur van de iteratiemethode verkregen kan worden. Daartoe merken wij op dat $\lim_{n\to\infty} v_i(n;f)/n = x_i(f)$ voor alle i (zie (1.18)). Laten wij nu eens aannemen dat een natuurlijk getal n_0 bestaat en getallen $v_i(f)$, $i = 1, \ldots, N$, zodat

$$(1.23) \qquad v_i(n;f) = nx_i(f) + v_i(f) \quad \text{voor } i = 1, \ldots, N;\ n \geqq n_0.$$

Substitueren wij (1.23) in (1.14) dan vinden wij voor $n \geqq n_0$

$$(1.24) \qquad nx_i(f) + x_i(f) + v_i(f) = r(i,f(i)) + n \sum_{j=1}^{N} p_{ij}(f)\, x_j(f) + \sum_{j=1}^{N} p_{ij}(f)\, v_j(f) \quad \text{voor } i = 1, \ldots, N.$$

Aangezien (1.24) geldt voor iedere $n \geq n_0$, vinden wij dat de getallen $x_i(f)$ en $v_i(f)$ voldoen aan

$$(1.25)\qquad \begin{cases} x_i(f) = \sum_{j=1}^{N} p_{ij}(f)\, x_j(f) & \text{voor } i = 1, \ldots, N \\ x_i(f) + v_i(f) = r(i,f(i)) + \sum_{j=1}^{N} p_{ij}(f)\, v_j(f) & \\ & \text{voor } i = 1, \ldots, N. \end{cases}$$

In hoofdstuk 2 tonen wij aan dat het stelsel van 2N lineair vergelijkingen in de onbekenden (u_i, t_i)

$$(1.26)\qquad \begin{cases} u_i = \sum_{j=1}^{N} p_{ij}(f)\, u_j & \text{voor } i = 1, \ldots, N \\ u_i + t_i = r(i,f(i)) + \sum_{j=1}^{N} p_{ij}(f)\, t_j & \\ & \text{voor } i = 1, \ldots, N. \end{cases}$$

oplosbaar is. Voor elke oplossing (U,T) van (1.26) geldt

$$(1.27)\qquad u_i = x_i(f) \qquad \text{voor } i = 1, \ldots, N.$$

Om dit laatste aan te tonen merken wij eerst op dat uit een n-voudige toepassing van de eerste gelijkheid uit (1.26) volgt dat $u_i = \sum_j p_{ij}^{(n)}(f)\, u_j$ voor $i = 1, \ldots, N$, $n \geq 1$; nemen wij van deze relatie de Cesarolimiet dan vinden wij dat (zie (1.17))

$$(1.28)\qquad u_i = \sum_{j=1}^{N} p_{ij}^{*}(f)\, u_j \qquad \text{voor } i = 1, \ldots, N.$$

Vermenigvuldigen wij beide leden van de tweede gelijkheid uit (1.26) met de invariante kans $p_{ki}^{*}(f)$ en sommeren wij daarna over i dan vinden wij met behulp van stelling 1.2.13 (paragraaf 1.2)

$$(1.29) \qquad \sum_{i=1}^{N} p_{ki}^{*}(f)\; u_i = \sum_{i=1}^{N} p_{ki}^{*}(f)\; r(i,f(i)) \text{ voor } k = 1, \ldots, N.$$

Uit (1.28), (1.29) en (1.19) volgt nu (1.27). Wij merken op dat (1.26) geen unieke oplossing heeft, immers als (U,T) voldoet, dan voldoet ook (U,T'), waarbij $t_i' = t_i + c$ voor $i = 1, \ldots, N$ en c een willekeurige constante is. Door aan (1.26) op de volgende wijze bijvoorwaarden toe te voegen verkrijgt men een unieke oplossing: Bepaal een splitsing van de toestandsruimte $\{1,\ldots,N\}$ van de Markov-keten met P(f) als matrix van overgangskansen in disjuncte priemfuiken $E_1, \ldots, E_m$ zodat $\{1,\ldots,N\}\setminus\bigcup_{i=1}^{m} E_i$ geen fuik is (zie def. 1.2.5). Laat e_i een willekeurige toestand uit priemfuik E_i zijn. Dan heeft (1.26) onder de voorwaarden

$$(1.30) \qquad t_{e_i} = 0 \qquad\qquad \text{voor } i = 1, \ldots, m$$

een ondubbelzinnig bepaalde oplossing. *)

Het stelsel lineaire vergelijkingen (1.26) vormt tezamen met de voorwaarde (1.30) de ene bouwsteen van HOWARD's iteratiemethode. Om tot de andere bouwsteen te komen, beschouwen wij het volgende probleem. Stel eens dat het systeem alleen gedurende de eerste n + 1 beslissingstijdstippen beheerd wordt. Op het eerste beslissingstijdstip $t = 1,\ldots,n$ moet hij strategie $f \in F$ toepassen. Als de beslisser op $t = 0$ in toestand i beslissing a neemt en daarna f toepast, dan wordt de totale verwachte opbrengst over de eerste n + 1 beslissingstijdstippen gegeven door (zie (1.13))

---

*) Een bewijs kan gevonden worden in deel 7b van de Leergang besliskunde, G. de Leve en H.C. Tijms, Dynamische programmering II.

$$(1.31) \qquad r(i,a) + \sum_{j=1}^{N} p(j|i,a)\, v_j(n;f).$$

De beslisser kiest uiteraard een beslissing a waarvoor (1.31) maximaal is. Nemen wij aan dat (1.23) juist is en substitueren wij deze relatie in (1.31), dan gaat (1.31) voor $n \geqq n_0$ over in

$$(1.32) \qquad r(i,a) + \sum_{j=1}^{N} p(j|i,a)\, v_j(f) + n \sum_{j=1}^{N} p(j|i,a)\, x_j(f).$$

Als n voldoende groot is dan is (1.32) maximaal voor een beslissing a uit de verzameling A'(i) waarbij A'(i) bestaat uit die beslissing(en) $a \in A(i)$ waarvoor

$$(1.33) \qquad \sum_{j=1}^{N} p(j|i,a)\, x_j(f)$$

maximaal is. Voor n voldoende groot is (1.32) dus maximaal voor een beslissing $a \in A'(i)$, welke

$$(1.34) \qquad r(i,a) + \sum_{j=1}^{N} p(j|i,a)\, v_j(f)$$

onder de voorwaarde $a \in A'(i)$ maximaliseert. Voegen wij nu aan elke i een beslissing $a \in A'(i)$ toe, welke (1.34) onder de voorwaarde $a \in A'(i)$ maximaliseert, dan vinden wij een strategie $f_1 \in F$ waarvoor geldt (hoofdstuk 2)

$$(1.35) \qquad x_i(f_1) \geqq x_i(f) \qquad \text{voor } i = 1, \ldots, N.$$

Na de bouwstenen van HOWARD's iteratiemethode op een nogal heuristische wijze gevonden te hebben, formuleren wij nu deze iteratiemethode. Wij beschrijven de n^{de} stap (in de eerste stap beginnen wij met een willekeurige $f_0 \in F$).

n^{de} stap van HOWARD's iteratiemethode

a) Zij strategie f_{n-1} verkregen aan het eind van de $(n-1)^{ste}$ stap. Bepaal een splitsing van de toestandsruimte $\{1,\dots,N\}$ van de Markovketen met $P(f_{n-1})$ als matrix van overgangskansen in disjuncte kernfuiken $E_1, \dots, E_m$, zodat $\{1,\dots,N\} \setminus \bigcup_{1}^{m} E_i$ geen fuik is. Kies in elke kernfuik E_j de toestand met de grootste index, zeg toestand e_j. Bepaal de unieke oplossing $(x_i(f_{n-1}), t_i(f_{n-1}))$ van het stelsel lineaire vergelijkingen (vgl. (1.27))

$$
(1.36) \quad \begin{cases} u_i = \sum_{j=1}^{N} p_{ij}(f_{n-1})\, u_j & \text{voor } i=1,\dots,N \\ u_i + t_i = r(i,f_{n-1}(i)) + \sum_{j=1}^{N} p_{ij}(f_{n-1})\, t_j & \text{voor } i=1,\dots,N \\ t_{e_j} = 0 & \text{voor } j=1,\dots,m. \end{cases}
$$

b) Bepaal voor elke $i = 1, \dots, N$ de verzameling $A'(i)$ van die beslissingen $a \in A(i)$ waarvoor

$$
(1.37) \qquad \sum_{j=1}^{N} p(j|i,a)\, x_j(f_{n-1})
$$

maximaal is. Voeg aan elke i een beslissing $a \in A'(i)$ toe, welke

$$
(1.38) \qquad r(i,a) + \sum_{j=1}^{N} p(j|i,a)\, t_j(f_{n-1})
$$

onder de voorwaarde $a \in A'(i)$ maximaliseert (terwille van de convergentie spreken wij af dat wij $a = f_{n-1}(i)$ kiezen als deze beslissing tot $A'(i)$ behoort en (1.38) onder $a \in A'(i)$ maximaliseert). Op deze wijze verkrijgen wij een strategie $f_n \in F$.

Einde van de n^{de} stap.

In hoofdstuk 2 tonen wij aan dat $x_i(f_n) \geqq x_i(f_{n-1})$ voor $i = 1,\ldots,N$ en alle $n \geqq 1$. In het in de voetnoot op blz. 10 genoemde boek kan een bewijs worden gevonden dat deze iteratiemethode van HOWARD na een eindig aantal stappen, zeg m, leidt tot $f_m = f_{m-1}$. Voor de strategie $f^* = f_m$ geldt dan $x_i(f^*) \geqq x_i(f)$ voor alle $i = 1, \ldots, N$ en alle $f \in F$. Het bewijs herhalen wij hier niet omdat wij in de eerste plaats geïnteresseert zijn in de existentie van een dergelijke strategie $f^* \in F$. In hoofdstuk 2 zullen wij de existentie van zo'n strategie f^* aantonen met behulp van een iets gewijzigde versie van HOWARD's iteratiemethode.

Opmerking.

De componenten (u_i) van een oplossing (u_i,t_i) van (1.26) kunnen een physische interpretatie gegeven worden, omdat $u_i = x_i(f)$ voor alle i en $x_i(f)$ de verwachte gemiddelde opbrengst per tijdseenheid is genomen over $t = 0, 1, \ldots$, gegeven strategie $f \in F$ en begintoestand i. Onder zekere voorwaarden kunnen wij ook de getallen t_i een interpretatie geven. Als toestanden k en h tot een zelfde priemfuik behoren en als deze priemfuik aperiodiek is (d.w.z. een terugkeertoestand met periode 1 bevat, zie paragraaf 1.2) dan kunnen wij aantonen dat (vgl. (1.13))

$$t_k - t_h = \lim_{n\to\infty} \{v_k(n;f) - v_h(n;f)\}. \tag{1.39}$$

Dus $t_k - t_h$ is het verschil in de totale verwachte opbrengst die ontstaat als het systeem toestand k i.p.v. toestand h als begintoestand heeft. Om relatie (1.39) aan te tonen, passen wij op (1.26) een n-voudige herhaalde substitutie toe. Wij vinden dan dat voor $i = 1, \ldots, N$ en elke $n \geqq 1$ geldt

$$u_i = \sum_{j=1}^{N} p_{ij}^{(n)}(f)\, u_j \tag{1.40}$$

$$u_i + t_i = \sum_{k=0}^{n-1} \sum_{j=1}^{N} p_{ij}^{(k)}(f)\, \{r(j,f(j) - u_j\} + \sum_{j=1}^{N} p_{ij}^{(n)}(f)\, t_j. \tag{1.41}$$

Met behulp van (1.13) en (1.27) volgt uit (1.40) en (1.41)

(1.42) $$t_i = v_i(n;f) - nx_i(f) + \sum_{j=1}^{N} p_{ij}^{(n)}(f)\, t_j \quad \text{voor } 1 \leqq i \leqq N;\ n \geqq 1.$$

Voor toestanden k en h uit een zelfde priemfuik geldt dat $x_k(f) = x_h(f)$ (zie (1.21)). Als deze priemfuik bovendien aperiodiek is dan geldt tevens dat (zie de stellingen 1.2.1a, 1.2.12 en 1.2.20)
$\lim_{n\to\infty} p_{kj}^{(n)}(f) = \lim_{n\to\infty} p_{hj}^{(n)}(f)$ voor $j = 1, \ldots, N$.
Deze relaties impliceren tezamen met (1.42) de relatie (1.39).

1.2 Aftelbare en eindige Markov-ketens.

Definitie 1.2.1

Een rij stochastische variabelen $\{\underline{x}_n,\ n = 0, 1, \ldots\}$ met een aftelbare of eindige toestandsruimte I is een (homogene) Markov-keten als getallen p_{ij}, $i, j \in I$ bestaan die voldoen aan

(1.42) $$\begin{cases} 0 \leqq p_{ij} \leqq 1 & \text{alle } i, j \in I \\ \sum_{j\in I} p_{ij} = 1 & \text{alle } i \in I, \end{cases}$$

en zodanig zijn dat

(1.43) $$P\{(\underline{x}_0,\ldots,\underline{x}_{n+1})=(i_0,\ldots,i_{n+1})\} = P\{(\underline{x}_0,\ldots,\underline{x}_n)=(i_0,\ldots,i_n)\}p_{i_n i_{n+1}}$$

voor alle $i_0,\ldots,i_{n+1} \in I$ en alle $n \geqq 0$.

Definitie 1.2.2

Voor alle $i, j \in I$ definieren wij

(1.44) $$p_{ij}^{(0)} = \begin{cases} 1 & \text{als } j = i \\ 0 & \text{als } j \neq i \end{cases}$$

en

$$(1.45) \qquad p_{ij}^{(n)} = \sum_{i_1,\ldots,i_{n-1}\in I} p_{ij_1}\, p_{i_1 i_2} \cdots p_{i_{n-2} i_{n-1}}\, p_{i_{n-1} j}$$

voor $n \geqq 2$.

De getallen $p_{ij}^{(n)}$ worden de n-stapsovergangskansen genoemd en kunnen geïnterpreteerd worden als de kans dat het systeem over n stappen in toestand j is gegeven dat het systeem nu in toestand i is.

In het hierna volgende zullen wij herhaaldelijk gebruik maken van enkele bekende stellingen uit de analyse, die wij bijeenbrengen in de volgende stelling.

Stelling 1.2.1

a) Als $\lim_{n\to\infty} a_n = a$ dan bestaat de Césarolimiet van de rij $\{a_n\}$ eveneens en is ook gelijk aan a, dus

$$(1.46) \qquad \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} a_k = a$$

b) Als de getallen $a_{nm} \geqq 0$ zijn voor alle m, $n \geqq 0$, dan geldt

$$(1.47) \qquad \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{nm} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_{nm}$$

c) Neem aan dat de getallen a_{nm} voldoen aan $\lim_{n\to\infty} a_{nm} = a_m$ voor $m = 0,1,\ldots$. Laten de getallen p_m niet-negatief zijn voor alle $m \geqq 0$. Als getallen b_m bestaan zodat $|a_{nm}| \leqq b_m$ voor $n = 0, 1, \ldots$ en $m = 0, 1, \ldots$ en $\sum_m b_m p_m < \infty$ is, dan geldt

$$(1.48) \qquad \lim_{n\to\infty} \sum_{m=0}^{\infty} a_{nm}\, p_m = \sum_{m=0}^{\infty} a_m\, p_m.$$

d) Laten $a_{nm} \geqq 0$ voor $m = 0, 1, \ldots, n = 0, 1, \ldots$ met $\lim_{n\to\infty} a_{nm} = a_m$ voor $m = 0, 1, \ldots$ en $\lim_{n\to\infty} \sum_m a_{nm}$ bestaat, dan geldt

$$(1.49) \qquad \lim_{n\to\infty} \sum_{m=0}^{\infty} a_{nm} \geqq \sum_{m=0}^{\infty} a_m .$$

Uit definitie 1.2.1 en stelling 1.2.1c volgt direct dat

$$(1.50) \qquad p_{ij}^{(m+n)} = \sum_{k\in I} p_{ik}^{(m)} p_{kj}^{(n)} \quad \text{voor alle } i, j \in I;\ m, n \geqq 0.$$

Definitie 1.2.3

Toestand j heet bereikbaar vanuit toestand i als een getal $n \geqq 1$ bestaat met $p_{ij}^{(n)} > 0$. Notatie $i \to j$ ($i \not\to j$ betekent dat j niet bereikbaar is vanuit i). De toestanden i en j communiceren als j bereikbaar is vanuit i en i bereikbaar is vanuit j. Notatie $i \leftrightarrow j$.

Stelling 1.2.2

Als k bereikbaar is vanuit i en j is bereikbaar vanuit k, dan is j ook bereikbaar vanuit i.

Bewijs

Getallen $m, n \geqq 1$ bestaan met $p_{ik}^{(m)} > 0$ en $p_{kj}^{(n)} > 0$. Uit (1.50) volgt

$$(1,51) \qquad p_{ij}^{(m+n)} \geqq p_{ik}^{(m)} p_{kj}^{(n)} > 0.$$

Definitie 1.2.4

Toestand i heet essentieel als uit $i \to j$ volgt $j \to i$, toestand i heet niet-essentieel als een toestand j bestaat met $i \to j$ en $j \not\to i$.

<u>Definitie 1.2.5</u>

Een niet-lege verzameling $S \subseteq I$ van toestanden noemen wij een <u>fuik</u> als vanuit elke toestand in S uitsluitend toestanden in S bereikbaar zijn, m.a.w. $p_{ij} = 0$ als $i \in S$ en $j \notin S$. Een fuik die geen twee of meer disjuncte deelfuiken bevat, heet een <u>priemfuik</u>. Een fuik die geen kleinere fuik bevat, heet een <u>kernfuik</u>.

Merk op dat de verzameling van alle toestanden een fuik is en dat een kernfuik een priemfuik is; een priemfuik behoeft echter geen kernfuik te zijn.

<u>Stelling 1.2.3</u>

Een verzameling van toestanden $S \subseteq I$ is een kernfuik dan en slechts dan als S een fuik is met de eigenschap dat voor elk paar $i, j \in S$ geldt dat $i \leftrightarrow j$.

<u>Bewijs</u>

$\Rightarrow$ Stel dat S een kernfuik is. Laat i een willekeurige toestand uit S zijn. Definieer $S(i) = \{j | i \to j\}$. Stel $j \in S(i)$ en $j \to k$. Als wij kunnen aantonen dat $k \in S(i)$, dan is S(i) een fuik. Uit $j \in S(i)$ volgt dat $i \to j$. Uit $i \to j$ en $j \to k$ volgt dat $i \to k$ (stelling 1.2.2). Dus $k \in S(i)$, oftewel S(i) is een fuik. Aangezien S een kernfuik is en de fuik $S(i) \subseteq S$, geldt dat $S(i) = S$. Uit de definitie van S(i) volgt nu dat alle toestanden in S communiceren.

$\Leftarrow$ Het andere deel van de stelling vinden wij met een bewijs uit het ongerijmde. Stel dat S geen kernfuik is. Dan bestaat een toestand $i \in S$ die tot een kleinere fuik van S behoort. Deze kleinere fuik bevat uiteraard S(i). Aangezien alle toestanden van S communiceren, geldt $S(i) = S$. Dit is een tegenspraak met het feit dat S(i) bevat is in een kleinere fuik van S. Dus S is een kernfuik.

<u>Stelling 1.2.4</u>

Een <u>eindige</u> fuik S bevat tenminste één essentiële toestand.

Bewijs

Het bewijs is uit het ongerijmde. Stel dat de fuik S uitsluitend niet-essentiële toestanden bevat. Laat $i_1 \in S$ (S is niet leeg). Uit definitie 1.2.4 volgt dat een toestand $i_2 \in S$ bestaat met $i_1 \to i_2$ en $i_2 \not\to i_1$. Dus $i_2 \neq i_1$. Bij toestand i_2 is een toestand $i_3 \in S$ te vinden met $i_2 \to i_3$, $i_3 \not\to i_2$. Voorts geldt $i_3 \not\to i_1$ (immers uit $i_3 \to i_1$ volgt $i_3 \to i_2$). De toestanden i_1, i_2 en i_3 zijn dus verschillend. Op deze wijze voortgaand, vinden wij dat de eindige fuik S een oneindige rij van verschillende toestanden $\{i_n\}$ bevat. Dit is een tegenspraak. Dus de stelling is juist.

Wij merken op dat stelling 1.2.4 voor een aftelbare fuik niet behoeft te gelden; een voorbeeld vormt de aftelbare Markov-keten met als overgangswaarschijnlijkheden $p_{n,n+1} = 1$ alle $n \geqq 1$ $(I = 1,2,\dots)$.

Stelling 1.2.5

Als

$$(1.52) \qquad I' \overset{\text{def}}{=} \{\text{verzameling van alle essentiële toestanden}\}$$

niet leeg is, dan is I' een fuik.

Bewijs

Zij $i \in I'$ en $i \to j$. Als wij kunnen aantonen dat $j \in I'$, dan is I' een fuik. Stel $j \notin I'$, dan bestaat een toestand k met $j \to k$ en $k \not\to j$. Uit $i \to j$ en stelling 1.2.2 volgt dat $i \to k$ en $k \not\to i$ moet gelden. Dit is in tegenspraak met $i \in I'$, dus de aanname $j \notin I'$ is onjuist, oftewel I' is een fuik.

De relatie "communiceren" is op I' een equivalentierelatie. Daartoe moeten wij aantonen dat voor alle $i, j, k \in I'$ geldt
a) $i \leftrightarrow i$ b) $i \leftrightarrow j \Leftrightarrow j \leftrightarrow i$ c) $i \leftrightarrow j$ en $j \leftrightarrow k \Rightarrow i \leftrightarrow k$.
De beweringen b) en c) zijn triviaal. Om bewering a) te bewijzen merken wij op dat een toestand h bestaat met $i \to h$. Omdat $i \in I'$ geldt tevens $h \to i$. Dus $i \to h \to i$, oftewel $i \leftrightarrow i$. Aangezien de relatie "communiceren" op I' een equivalentierelatie is, kunnen wij

I' splitsen in een aantal disjuncte equivalentieklassen I_1, I_2, ... (aangenomen dat I' niet leeg is). Elke I_ν is een kernfuik. Immers als $i \in I_\nu$ en $i \to j$ dan geldt ook $j \to i$, dus $i \leftrightarrow j$, oftewel $j \in I_\nu$. Dus I_ν is een fuik. Als $i, j \in I_\nu$, dan volgt uit de definitie van I_ν dat $i \leftrightarrow j$. Uit stelling 1.2.3 volgt nu dat I_ν een kernfuik is. Wij hebben nu de volgende stelling bewezen (vgl. stelling 1.2.3):

Stelling 1.2.6

Elke kernfuik is een fuik bestaande uit essentiële toestanden. De verzameling I' van alle essentiële toestanden is op unieke wijze te splitsen in een aantal kernfuiken I_1, I_2, ... (mits I' niet leeg is).

Stelling 1.2.7

Een eindige priemfuik bevat precies één kernfuik.

Bewijs

Een eindige fuik bevat tenminste één essentiële toestand (stelling 1.2.4), dus tenminste één kernfuik (stelling 1.2.6). Uit de definitie van priemfuik volgt, dat een priemfuik tenhoogste één kernfuik bevat. Waarmee de stelling bewezen is.

Definieer voor alle $n \geq 1$, $i, j \in I$

(1.53) $$f_{ij}^{(n)} = P\{\underline{x}_n = j, \underline{x}_m \neq j \text{ voor } 1 \leq m < n \mid \underline{x}_0 = i\}$$

De kans $f_{ij}^{(n)}$ kunnen wij interpreteren als de kans dat op $t = n$ voor het eerst een overgang naar toestand j plaatsvindt, gegeven dat het systeem op $t = 0$ in toestand i is.
Voorts definiëren wij voor alle $i, j \in I$

(1.54) $$f_{ij} = P\left(\bigcup_{n=1}^{\infty} \{\underline{x}_n = j\} \mid \underline{x}_0 = i\right).$$

De kans f_{ij} kunnen wij interpreteren als de kans dat ooit een overgang naar toestand j plaatsvindt, gegeven dat het systeem toestand i als begintoestand heeft.

Uit de definities (1.53) en (1.54) volgt op eenvoudige wijze

$$(1.55) \qquad f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)} \qquad \text{voor alle } i, j \in I.$$

Definieer voor alle i, j ∈ I

$$(1.56) \qquad \mu_{ij} = \begin{cases} \sum_{n=1}^{\infty} n f_{ij}^{(n)} & \text{als } f_{ij} = 1 \\ \\ \infty & \text{als } f_{ij} < 1. \end{cases}$$

Wij kunnen μ_{ij} interpreteren als de verwachting van het tijdstip waarop voor het eerst een overgang naar toestand j plaatsvindt, gegeven dat het systeem op t = 0 in toestand i is.

Uit de relatie

$$(1.57) \quad P\{\underline{x}_n = j \mid \underline{x}_0 = i\} = \sum_{k=1}^{n} P\{\underline{x}_k = j, \underline{x}_m \neq j, 1 \leq m < k, \underline{x}_n = j \mid \underline{x}_0 = i\} =$$

$$= \sum_{k=1}^{n} P\{\underline{x}_k = j, \underline{x}_m \neq j, 1 \leq m < k \mid \underline{x}_0 = i\} P\{\underline{x}_n = j \mid \underline{x}_k = j\}$$

volgt

<u>Stelling 1.2.8</u>

$$p_{ij}^{(n)} = \sum_{k=1}^{n} f_{ij}^{(k)} \, p_{jj}^{(n-k)} \qquad \text{voor alle } n \geq 1, \ i, j \in I.$$

<u>Stelling 1.2.9</u>

a) Als voor een toestand $j \in I$ geldt dat

(1.59) $$\lim_{n\to\infty} p_{jj}^{(n)} = p_{jj}^{*}$$

dan is

(1.60) $$\lim_{n\to\infty} p_{ij}^{(n)} = f_{ij}\, p_{ij}^{*} \qquad \text{voor alle } i \in I.$$

b) Als voor een toestand $j \in I$ geldt dat

(1.61) $$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{jj}^{(k)} = p_{jj}^{*}$$

dan is

(1.62) $$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)} = f_{ij}\, p_{ij}^{*} \qquad \text{voor alle } i \in I.$$

<u>Bewijs</u>

a) Zij $\varepsilon > 0$. Aangezien $f_{ij}^{(k)} \geqq 0$ is voor alle $k \geqq 1$ en $f_{ij} \leqq 1$ is, kunnen wij een getal n_0 vinden, zodat

(1.63) $$\sum_{k=n}^{\infty} f_{ij}^{(k)} \leqq \frac{\varepsilon}{3} \qquad \text{voor alle } n \geqq n_0.$$

Uit (1.59) volgt het bestaan van een getal $n_1 > n_0$, zodat voor elke $k = 1, \ldots, n_0$ geldt dat

(1.64) $$|p_{jj}^{(n-k)} - p_{jj}^{*}| \leqq \frac{\varepsilon}{3n_0} \qquad \text{voor alle } n \geqq n_1.$$

Uit (1.55), (1.58), (1.63), (1.64) en de afschatting $|p_{jj}^{(n)} - p_{jj}^{*}| \leqq 1$, volgt dat voor alle $n \geqq n_1$ geldt

(1.65) $$|p_{ij}^{(n)} - f_{ij}p_{jj}^{*}| = |\sum_{k=1}^{n} f_{ij}^{(k)}(p_{jj}^{(n-k)} - p_{jj}^{*}) - p_{jj}^{*} \sum_{k=n+1}^{\infty} f_{ij}^{(k)}| \leqq$$

$$\leq \sum_{k=1}^{n_0} f_{ij}^{(k)} |p_{jj}^{(n-k)} - p_{jj}^{*}| + \sum_{k=n_0+1}^{n} f_{ij}^{(k)} |p_{jj}^{(n-k)} - p_{jj}^{*}| + \frac{\varepsilon}{3} \leq$$

$$\leq n_0 \; \frac{\varepsilon}{3n_0} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

b) Definieer

(1.66) $$q_{ij}^{(n)} = \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)} \qquad \text{voor alle } i, j \in I \text{ en } n \geq 1.$$

Uit (1.58) en (1.66) volgt op eenvoudige wijze dat

(1.67) $$q_{ij}^{(n)} = \sum_{k=1}^{n-1} \frac{(n-k)}{n} f_{ij}^{(k)} q_{jj}^{(n-k)} + \frac{1}{n} \sum_{k=1}^{n} f_{ij}^{(k)}$$

voor alle $i, j \in I$; $n \geq 1$.

Op analoge wijze als onder punt a), kunnen wij nu aantonen dat $q_{ij}^{(n)}$ naar $f_{ij} \, p_{jj}^{*}$ convergeert voor alle $i \in I$.

Stelling 1.2.10

Laat $\{f_n, n \geq 1\}$ een rij van getallen zijn met $f_n \geq 0$ voor alle $n \geq 1$. Stel dat de rij van getallen $\{u_n, n \geq 0\}$ gedefinieerd is door

(1.68) $$u_0 = 1 \text{ en } u_n = \sum_{k=1}^{n} f_k \, u_{n-k} \qquad \text{voor } n \geq 1.$$

Dan kunnen de volgende uitspraken gedaan worden:

a) als

(1.69) $$\sum_{n=1}^{\infty} f_n < 1$$

dan geldt

(1.70) $$\sum_{n=0}^{\infty} u_n < \infty$$

b) als

$$(1.71) \qquad \sum_{n=1}^{\infty} f_n = 1$$

dan geldt

$$(1.72) \qquad \sum_{n=0}^{\infty} u_n = \infty$$

en

$$(1.73) \qquad \lim_{n\to\infty} u_{n\lambda} = \frac{\lambda}{\mu} ,$$

waarin λ het grootste gehele getal $t \geqq 1$ is met de eigenschap dat $u_n = 0$ als n niet deelbaar door t is en

$$(1.74) \qquad \mu \overset{\text{def}}{=} \sum_{n=1}^{\infty} nf_n .$$

Het bewijs van deze stelling laten wij achterwege. Het bewijs kan gevonden worden in W. Feller, An Introduction to Probability Theory and Its Applications, Volume I, second edition, chapter 13.

Wij merken op dat (1.68) de getallen u_n op ondubbelzinnige wijze vastlegt en dat $u_n \geqq 0$ is voor elke $n \geqq 0$. Als $\mu = \infty$ dan leze men $\lambda/\mu = 0$ in (1.73).

<u>Definitie 1.2.6</u>

Een toestand $j \in I$ heet een <u>terugkeertoestand</u> als $f_{jj} = 1$ en toestand j heet een doorgangstoestand als $f_{jj} < 1$ is. Een terugkeertoestand j wordt <u>positief</u> resp. een <u>nultoestand</u> genoemd als $\mu_{jj} < \infty$ is resp. $\mu_{jj} = \infty$ (vgl. (1.55) en (1.56)).

<u>Definitie 1.2.7</u>

De <u>periode</u> λ van een terugkeertoestand j is gedefinieerd als het <u>grootste</u> getal $t \geqq 1$ met de eigenschap dat $p_{jj}^{(n)} = 0$ als n niet deelbaar door t is.

Stelling 1.2.11

a) Een toestand j is een doorgangstoestand dan en slechts dan als

$$\sum_{n=0}^{\infty} p_{jj}^{(n)} < \infty . \tag{1.75}$$

b) Een terugkeertoestand j is een nultoestand dan en slechts dan als

$$\lim_{n\to\infty} p_{jj}^{(n)} = 0 . \tag{1.76}$$

c) Zowel voor een doorgangstoestand j als een nultoestand j geldt

$$\lim_{n\to\infty} p_{ij}^{(n)} = 0 \qquad \text{voor alle } i \in I . \tag{1.77}$$

d) Als j een terugkeertoestand met periode 1 is, dan geldt

$$\lim_{n\to\infty} p_{ij}^{(n)} = \frac{f_{ij}}{\mu_{jj}} \qquad \text{voor alle } i \in I, \tag{1.78}$$

waarin $1/\mu_{jj} = 0$ als $\mu_{jj} = \infty$.

e) Als j een terugkeertoestand met periode λ is, dan geldt

$$\lim_{n\to\infty} p_{jj}^{(n\lambda)} = \frac{\lambda}{\mu_{jj}} . \tag{1.79}$$

Bewijs

Wij zullen gebruik maken van stelling 1.2.10 met (vgl. (1.58))

$$f_n = f_{jj}^{(n)} \text{ voor } n \geqq 1 \text{ en } u_n = p_{jj}^{(n)} \text{ voor } n \geqq 0. \tag{1.80}$$

Uit de definitie van doorgangstoestand en stelling 1.2.9 volgt direct bewering a).

Als j een nultoestand is dan geldt $f_{jj} = 1$ en $\mu_{jj} = \infty$ en uit (1.73) volgt dat $p_{jj}^{(n\lambda)}$ naar nul convergeert voor $n \to \infty$, als λ de periode van j is. Voorts geldt dat $p_{jj}^{(n)}$ nul is als n niet deelbaar door λ is. Dus voor een nultoestand geldt (1.76). Stel nu dat j een terugkeertoestand is waarvoor (1.76) geldt. Laat λ de periode van j zijn. Aangezien $p_{jj}^{(n\lambda)}$ naar nul convergeert, volgt uit (1.73) dat $\mu_{jj} = \infty$. Dus j is een nultoestand. Waarmee bewering b) bewezen is.

Uit (1.75) volgt dat $p_{jj}^{(n)}$ naar nul convergeert als j een doorgangstoestand is. Bewering c) volgt nu met behulp van punt a) van stelling 1.2.9.

De beweringen d) en e) volgen direct met behulp van (1.73) en punt a) van stelling 1.2.9.

Stelling 1.2.12

a) Voor alle i, j $\in$ I bestaat

$$p_{ij}^{*} \overset{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)}. \tag{1.81}$$

b) Voor alle i, j $\in$ I geldt

$$p_{ij}^{*} = f_{ij}\, p_{jj}^{*}. \tag{1.82}$$

c) Zowel voor een doorgangstoestand j als een nultoestand j geldt

$$p_{ij}^{*} = 0 \qquad \text{voor alle } i \in I. \tag{1.83}$$

d) Uitsluitend als j een positieve terugkeertoestand is, geldt

$$p_{jj}^{*} = \frac{1}{\mu_{jj}} > 0. \tag{1.84}$$

Bewijs

Beschouw eerst het geval dat j geen positieve terugkeertoestand is. Uit punt a) van stelling 1.2.1 en punt c) van stelling 1.2.11 volgt direct dat bewering c) juist is en dat (1.81) juist is voor het geval j geen positieve terugkeertoestand is. Beschouw vervolgens het geval dat j een positieve terugkeertoestand is; dus $\mu_{jj} < \infty$. Zij λ de periode van toestand j. Uit de definitie van periode volgt dat voor elke $n \geq 1$ geldt

$$(1.85) \qquad \frac{1}{n}\sum_{k=1}^{n} p_{jj}^{(k)} = \frac{[n/\lambda]}{n}\;\frac{1}{|n/\lambda|}\sum_{k=1}^{[n/\lambda]} p_{jj}^{(k\lambda)} ,$$

waarin $[n/\lambda]$ het grootste gehele getal is dat $\leqq \frac{n}{\lambda}$ is. Uit (1.73) en punt a) van stelling 1.2.1 volgt nu

$$(1.86) \qquad \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} p_{jj}^{(k)} = \frac{1}{\lambda} \cdot \frac{\lambda}{\mu_{jj}} = \frac{1}{\mu_{jj}} .$$

Met behulp van punt b) van stelling 1.2.9 volgt direct dat bewering a) juist is. De bewering b) is ook een direct gevolg van punt b) van stelling 1.2.9 Uit (1.86) volgt dat voor een positieve terugkeertoestand j geldt dat $p_{jj}^{*} = 1/\mu_{jj} > 0$ is. Waaruit bewering d) volgt, omdat overeenkomstig bewering c) geldt $p_{jj}^{*} = 0$ als j geen positieve terugkeertoestand is.

Opmerking

Uit punt d) van stelling 1.2.1 en de gelijkheid

$$(1.87) \qquad \sum_{j\in I} \frac{1}{n}\sum_{k=1}^{n} p_{ij}^{(k)} = 1 \qquad \text{voor alle } i\in I \text{ en } n \geqq 1 ,$$

volgt dat

$$(1.88) \qquad \sum_{j\in I} p_{ij}^{*} \leqq 1 \qquad \text{voor alle } i \in I.$$

Als I eindig is, dan geldt het gelijkheidsteken (zie punt c) van stelling 1.2.1). De getallen p_{ij}^{*} zijn uiteraard niet-negatief.

Stelling 1.2.13

$$(1.89) \qquad \sum_{k\in I} p_{ik}^{*} p_{kj} = \sum_{k\in I} p_{ik} p_{kj}^{*} = \sum_{k\in I} p_{ik}^{*} p_{kj}^{*} = p_{ij}^{*}$$

voor alle i, j $\in$ I.

Bewijs

Voer ter afkorting in

(1.90) $$q_{ij}^{(n)} = \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)}$$ voor alle $i, j \in I$ en $n \geq 1$.

Uit (1.50) en (1.90) volgt op eenvoudige wijze dat voor alle $n \geq 1$ geldt

(1.91) $$\left| \sum_{k \in I} p_{ik}\, q_{kj}^{(n)} - q_{ij}^{(n)} \right| = \left| \frac{1}{n}(p_{ij}^{(n+1)} - p_{ij}) \right| \leq \frac{2}{n}$$

voor alle $i, j \in I$.

Aangezien $q_{kj}^{(n)}$ naar p_{kj}^{*} convergeert, volgt uit (1.91) dat

(1.92) $$\lim_{n \to \infty} \sum_{k \in I} p_{ik}\, q_{kj}^{(n)} = p_{ij}^{*}$$ voor alle $i, j \in I$.

Omdat $|q_{kj}^{(n)}| \leq 1$ voor alle $k, j \in I$, $n \geq 1$ en $\sum_{k \in I} p_{ik} = 1$ voor alle $i \in I$, volgt door punt c) van stelling 1.2.1 op (1,92) toe te passen dat

(1.93) $$p_{ij}^{*} = \sum_{k \in I} p_{ik} (\lim_{n \to \infty} q_{kj}^{(n)}) = \sum_{k \in I} p_{ik}\, p_{kj}^{*}$$

voor alle $i, j \in I$.

Uit de ongelijkheid (vgl. (1.90) en (1.50))

(1.94) $$\left| \sum_{k \in I} q_{ik}^{(n)}\, p_{kj} - q_{ij}^{(n)} \right| = \left| \frac{1}{n}(p_{ij}^{(n+1)} - p_{ij}) \right| \leq \frac{2}{n}$$

voor alle $i, j \in I$

volgt

(1.95) $$\lim_{n \to \infty} \sum_{k \in I} q_{ik}^{(n)}\, p_{kj} = p_{ij}^{*}$$ voor alle $i, j \in I$.

Aangezien $q_{ij}^{(n)}$, $p_{ij} \geqq 0$ voor alle i, j $\in$ I en $n \geqq 1$, volgt uit (1.95) en punt d) van stelling 1.2.1 dat

$$(1.96) \qquad \overset{*}{p}_{ij} \geqq \sum_{k\in I} (\lim_{n\to\infty} q_{ik}^{(n)})\, p_{kj} = \sum_{k\in I} \overset{*}{p}_{ik}\, p_{kj}$$

voor alle i, j $\in$ I.

Wij zullen nu aantonen dat in (1.96) het gelijkheidsteken moet gelden voor alle i, j $\in$ I. Stel dat voor een paar (i_0, j_0) het > teken in (1.96) geldt, dan volgt uit (1.96) en punt b) van stelling 1.2.1

$$(1.97) \qquad \sum_{j\in I} \overset{*}{p}_{i_0 j} > \sum_{j\in I} \sum_{k\in I} \overset{*}{p}_{i_0 k}\, p_{kj} = \sum_{k\in I} \overset{*}{p}_{i_0 k}.$$

Dit is een tegenspraak, dus

$$(1.98) \qquad \overset{*}{p}_{ij} = \sum_{k\in I} \overset{*}{p}_{ik}\, p_{kj} \qquad \text{voor alle i, j} \in \text{I.}$$

Door een herhaalde toepassing van de gelijkeid (1.98) vinden wij

$$(1.99) \qquad \overset{*}{p}_{ij} = \sum_{k\in I} \overset{*}{p}_{ik}\, p_{kj}^{(n)} \qquad \text{voor alle i, j} \in \text{I en alle } n \geqq 1.$$

Uit (1.90), (1.99) en punt b) van stelling 1.2.1 volgt dat

$$(1.100) \qquad \overset{*}{p}_{ij} = \sum_{k\in I} \overset{*}{p}_{ik}\, q_{kj}^{(n)} \qquad \text{voor alle i, j} \in \text{I en alle } n \geqq 1.$$

Aangezien $|q_{kj}^{(n)}| \leqq 1$ voor alle k, j $\in$ I en $n \geqq 1$ en $\sum_{k\in I} \overset{*}{p}_{ik} \leqq 1$ is (vgl. (1.88)) volgt door toepassing van punt c) van stelling 1.2.1 dat

$$(1.101) \qquad \overset{*}{p}_{ij} = \sum_{k\in I} \overset{*}{p}_{ik}\, \overset{*}{p}_{kj} \qquad \text{voor alle i, j} \in \text{I.}$$

Waarmee het bewijs van de stelling voltooid is.

Stelling 1.2.14

Als toestand j niet-essentieel is, dan is j een doorgangstoestand; m.a.w. elke terugkeertoestand is essentieel.

Bewijs

Wij moeten aantonen dat (zie (1.54) en definitie 1.2.6)

$$(1.102) \qquad f_{jj} = P\Big(\bigcup_{n=1}^{\infty}\{\underline{x}_n = j \mid \underline{x}_0 = j\}\Big) < 1.$$

De toestand j is niet-essentieel, dus een toestand $k \in I$ bestaat met $j \to k$ en $k \not\to j$. Uit $j \to k$ volgt dat een eindige $n \geqq 1$ bestaat met $p_{jk}^{(n)} > 0$. Zij r de kleinste n met deze eigenschap. Dan geldt dat $p_{jk}^{(r)}$ gelijk is aan de kans op een overgang van toestand j naar toestand k in r stappen, zonder dat tussentijds een overgang naar toestand j plaatsvindt. Deze betekenis van $p_{jk}^{(r)}$ impliceert tezamen met de relatie $k \not\to j$, dat

$$(1.103) \qquad 1 - f_{jj} = P\{\underline{x}_n \neq j \text{ voor alle } n \geqq 1 \mid \underline{x}_0 = j\} \geqq p_{jk}^{(r)} > 0.$$

Dus $f_{jj} < 1$; waarmee de stelling bewezen is.

Opmerking

Het omgekeerde van deze stelling behoeft niet te gelden. Voor een aftelbare Markov-keten kan gelden dat een doorgangstoestand essentieel is (voor een eindige Markov-keten geldt wel het omgekeerde van de stelling, zie stelling 1.2.16).

Stelling 1.2.15

In een kernfuik zijn alle toestanden van hetzelfde type; d.w.z. òf alle toestanden zijn doorgangstoestanden òf alle toestanden zijn positieve terugkeertoestanden met een zelfde periode òf alle toestanden zijn nultoestanden met een zelfde periode.

Bewijs

Laten i en j, $i \neq j$ twee vaste, doch willekeurig gekozen, toestanden uit de kernfuik zijn. Uit stelling 1.2.3 volgt dat $i \leftrightarrow j$; m.a.w. getallen $r, s \geq 1$ bestaan, zodat

$$(1.104) \qquad p_{ij}^{(r)} > 0 \text{ en } p_{ji}^{(s)} > 0 \text{ is.}$$

Uit (1.50) volgt dat voor elke $n \geq 0$ geldt

$$(1.105) \qquad p_{ii}^{(n+r+s)} \geq p_{ij}^{(r)}\, p_{jj}^{(n)}\, p_{ji}^{(s)} \text{ en } p_{jj}^{(n+r+s)} \geq p_{ji}^{(s)}\, p_{ii}^{(n)}\, p_{ij}^{(r)}.$$

Uit (1.105) volgt dat

$$(1.106) \qquad \sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty \Longleftrightarrow \sum_{n=0}^{\infty} p_{jj}^{(n)} < \infty.$$

Punt a) van stelling 1.2.11 leert ons nu dat i en j òf beide doorgangstoestanden òf beide terugkeertoestanden zijn.
Beschouw nu het geval dat i en j beide terugkeertoestanden zijn. Uit de eerste ongelijkheid van (1.105) volgt dat

$$(1.107) \qquad \limsup_{n \to \infty} p_{ii}^{(n)} \geq p_{ij}^{(r)}\, p_{ji}^{(s)} \limsup_{n \to \infty} p_{jj}^{(n)} \geq 0.$$

Stel dat i een nultoestand is. Dan geldt (zie punt b) van stelling 1.2.11) dat $\limsup p_{ii}^{(n)} = \lim p_{ii}^{(n)} = 0$. Uit (1.107) volgt dan $\limsup p_{jj}^{(n)} = 0$. Omdat $0 \leq \liminf p_{jj}^{(n)} \leq \limsup p_{jj}^{(n)} = 0$, geldt $\limsup p_{jj}^{(n)} = \liminf p_{jj}^{(n)} = \lim p_{jj}^{(n)} = 0$. Dus terugkeertoestand j is ook een nultoestand (punt b) van stelling 1.2.11). Rest ons nog te bewijzen dat de terugkeertoestanden i en j dezelfde periode hebben. Zij λ_i de periode van i en λ_j de periode van j. Uit de eerste ongelijkheid van (1.105) met $n = 0$ volgt dat $p_{ii}^{(r+s)} > 0$ is, derhalve is $r + s$ deelbaar door λ_i. Als voor een n geldt $p_{jj}^{(n)} > 0$, dan volgt uit de eerste ongelijkheid in (1.105) dat $p_{ii}^{(n+r+s)} > 0$ is; dus $n + r + s$ is deelbaar door λ_i. Omdat $r + s$ deelbaar door λ_i is, geldt dus dat n deelbaar door

λ_i is, als $p_{jj}^{(n)} > 0$ is. Uit de definitie van periode volgt nu dat $\lambda_i \leqq \lambda_j$. Op dezelfde wijze kunnen wij aantonen dat $\lambda_j \leqq \lambda_i$ is. Dus $\lambda_i = \lambda_j$. Waarmee het bewijs van de stelling voltooid is.

Stelling 1.2.16

Als toestand j een essentiële doorgangstoestand òf een nultoestand is, dan behoort de toestand j tot een kernfuik die uit oneindig veel toestanden bestaat.

Bewijs

Als j een nultoestand is, dan is toestand j essentieel (stelling 1.2.14). Dus òf j nu een nultoestand òf een essentiële doorgangstoestand is, toestand j behoort tot een ondubbelzinnig bepaalde kernfuik K (stelling 1.2.6). Stelling 1.2.15 leert ons dat alle toestanden van K van hetzelfde type zijn. Uit punt c) van stelling 1.2.11 volgt nu dat

$$\lim_{n\to\infty} p_{ik}^{(n)} = 0 \qquad \text{voor alle } i, k \in K. \tag{1.108}$$

Aangezien K een fuik is, geldt

$$\sum_{k\in K} p_{ij}^{(n)} = 1 \qquad \text{voor alle } i \in K \text{ en alle } n \geqq 1. \tag{1.109}$$

Als de kernfuik K eindig is, dan volgt met behulp van punt c) van stelling 1.2.1 uit (1.108) en (1.109)

$$1 = \lim_{n\to\infty} \sum_{k\in K} p_{ik}^{(n)} = \sum_{k\in K} (\lim_{n\to\infty} p_{ik}^{(n)}) = 0 \qquad \text{voor alle } i \in I. \tag{1.110}$$

Dit is een tegenspraak, dus K is een kernfuik die oneindig veel toestanden bevat.

Uit de stellingen 1.2.6, 1.2.14 en 1.2.16 volgt

Stelling 1.2.17

Een eindige kernfuik bevat uitsluitend positieve terugkeertoestanden.

Stelling 1.2.18

Als K een kernfuik is bestaande uit terugkeertoestanden, dan geldt

(1.111) $$f_{ij} = 1 \qquad \text{voor alle } i, j \in K$$

Bewijs

Volgens de definitie van terugkeertoestand geldt $f_{jj} = 1$ voor alle $j \in K$. Laten i en j, $i \neq j$ twee vaste, doch willekeurig gekozen, toestanden uit K zijn. Uit stelling 1.2.3 volgt dat $i \leftrightarrow j$. Dus een eindige $n \geqq 1$ bestaat met $p_{ji}^{(n)} > 0$. Zij nu

(1.112) $$r \overset{\text{def}}{=} \min\{n \mid p_{ji}^{(n)} > 0\}.$$

Uit de definitie van r volgt dat $p_{ji}^{(r)}$ de kans is op een overgang van toestand j naar toestand i in r stappen, zonder dat tussentijds een overgang naar toestand j plaatsvindt. Met behulp van de definitie van f_{hk} vinden wij nu

(1.113) $$1 - f_{jj} = P\{\underline{x}_n \neq j, \text{ alle } n \geqq 1 \mid \underline{x}_0 = j\} \geqq p_{ji}^{(r)}\,(1-f_{ij})$$

Aangezien $f_{jj} = 1$ en $p_{ji}^{(r)} > 0$ is, volgt uit (1.113) dat $f_{ij} = 1$.

Stelling 1.2.19

Als K een kernfuik is bestaande uit positieve terugkeertoestanden dan geldt

(1.114) $$\mu_{ij} < \infty \qquad \text{voor alle } i, j \in K$$

Bewijs

Volgens de definitie van positieve terugkeertoestanden geldt $\mu_{jj} < \infty$ voor alle $j \in K$. Laten i en j, $i \neq j$ twee vaste, doch willekeurig gekozen, toestanden uit K zijn. Laat r gedefinieerd zijn door (1.112). De definitie van r en de definitie (1.53) van $f_{hk}^{(n)}$ hebben tot gevolg

$$f_{jj}^{(n)} \geqq p_{ji}^{(r)} f_{ij}^{(n-r)} \qquad \text{voor } n > r. \tag{1.115}$$

Uit deze ongelijkheid volgt

$$\infty > \mu_{jj} = \sum_{n=1}^{\infty} nf_{jj}^{(n)} \geqq \sum_{n=r+1}^{\infty} (n-r)\, f_{jj}^{(n)} \tag{1.116}$$

$$\geqq \sum_{n=r+1}^{\infty} (n-r)\, p_{ji}^{(r)} f_{ij}^{(n-r)} = p_{ji}^{(r)} \mu_{ij}.$$

Omdat $p_{ji}^{(r)} > 0$ is, volgt uit (1.116) dat $\mu_{ij} < \infty$ is.

Stelling 1.2.20

Als K een kernfuik is, dan geldt

$$p_{ij}^{*} = p_{jj}^{*} \qquad \text{voor alle } i, j \in K. \tag{1.117}$$

Bewijs

Als de kernfuik òf alleen uit essentiële doorgangstoestanden òf alleen uit nultoestanden bestaat (vgl. stelling 1.2.15), dan is (1.117) een direct gevolg van punt c) van stelling 1.2.12. Als K uit positieve terugkeertoestanden bestaat, dan volgt (1.117) uit stelling 1.2.18 en punt b) van stelling 1.2.12.

Stelling 1.2.21.

Als een priemfuik P geen positieve terugkeertoestanden bevat of slechts een eindig aantal inessentiële toestanden bevat, dan geldt

$$(1.118) \qquad p^{*}_{ij} = p^{*}_{jj} \qquad \text{voor alle } i,\, j \in P.$$

Bewijs

Als P geen positieve terugkeertoestanden bevat dan volgt de bewering direct uit punt c) van stelling 1.2.12.
Beschouw nu het geval dat D de eindige verzameling van inessentiële toestanden uit P is en dat $K = P \setminus D$ niet leeg is. Dan bestaat K uit essentiële toestanden. Uit de definitie van priemfuik en stelling 1.2.6 volgt dat K een kernfuik is. Uit stelling 1.2.20 en punt c) van stelling 1.2.12 volgt direct dat bewering (1.118) juist is als i en j beide tot K behoren of als $j \in D$. Als wij kunnen aantonen dat $f_{ij} = 1$ voor $i \in D$ en $j \in K$, dan volgt uit punt b) van stelling 1.1.12 dat de bewering (1.118) ook juist is voor $i \in D$, $j \in K$.
Daartoe definiëren wij

$$(1.119) \qquad f_{iK} = P\Big(\bigcup_{n=1}^{\infty} \{\underline{x}_n \in K \mid \underline{x}_0 = i\}\Big) \qquad \text{voor } i \in D.$$

Omdat K een fuik is, geldt dat de eventualiteit $\{\underline{x}_{n+1} \notin K\}$ bevat is in $\{\underline{x}_n \notin K\}$, alle $n \geqq 1$. Dus de rij $A_n = \{\underline{x}_n \notin K \mid \underline{x}_0 = i\}$, $n \geqq 1$ van eventualiteiten is monotoon niet-stijgend en heeft derhalve als limiet $\bigcap_n A_n$. Dus

$$(1.120) \qquad 1 - f_{iK} = P\Big(\bigcap_{n=1}^{\infty} A_n\Big) = P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n) =$$

$$= \lim_{n\to\infty} \sum_{j\in D} p^{(n)}_{ij} \qquad \text{voor alle } i \in D.$$

Omdat D eindig is en uit doorgangstoestanden bestaat, volgt door toepassing van punt c) van de stellingen 1.2.1 en 1.2.11 dat

(1.121) $\qquad f_{iK} = 1 \qquad$ voor alle $i \in D$

Laten $i \in D$ en $j \in K$ vast gekozen zijn. M.b.v. (1.53) volgt

$$(1.122) \qquad f_{ij}^{(n)} = P\{\underline{x}_n = j,\ \underline{x}_m \notin K \text{ voor } 1 \leq m < n \mid \underline{x}_0 = i\} +$$

$$+ \sum_{k \in K, k \neq j} \sum_{r=0}^{n-1} P\{\underline{x}_r = k,\ \underline{x}_m \notin K \text{ voor } 1 \leq m < r \mid \underline{x}_0 = i\}\ f_{kj}^{(n-r)}.$$

Met behulp van de relatie $f_{kj} = 1$ voor alle $k \in K$(stelling 1.2.18), (1.119) en (1.121) volgt op eenvoudige wijze uit (1.222)

$$(1.123) \qquad f_{ij} = \sum_{n=1}^{\infty} P\{\underline{x}_n = j,\ \underline{x}_m \notin K \text{ voor } 1 \leq m < n \mid \underline{x}_0 = i\} +$$

$$+ \sum_{k \in K, k \neq j} \sum_{r=1}^{\infty} P\{\underline{x}_r = k,\ \underline{x}_m \notin K \text{ voor } 1 \leq m < r \mid \underline{x}_o = i\} = f_{iK} = 1.$$

Opmerking. De stelling geldt niet altijd als P naast tenminste één positieve terugkeertoestand oneindig veel inessentiële toestanden bevat. Voorbeeld, $P = \{1,2,\ldots\}$ met $p_{11} = 1$, $p_{21} = \frac{1}{2}$, $p_{23} = \frac{1}{2}$ en $p_{n,n+1} = 1$ voor $n \geq 3$.

Definitie 1.2.8.

Een kansverdeling $(a_j,\ j \in I)$ op I heet een invariante kansverdeling van de Markov-keten als

(1.124) $\qquad \sum_{j \in I} a_j = 1, \quad a_j \geq 0 \qquad$ voor alle $j \in I$

en

(1.125) $\qquad a_j = \sum_{k \in I} p_{kj}\, a_k \qquad$ voor alle $j \in I$.

Stelling 1.2.22.

Als de toestandsruimte I een kernfuik is bestaande uit positieve terugkeertoestanden, dan heeft de Markov-keten een unieke invariante kansverdeling $(a_j,\ j \in I)$, waarin

(1.126) $\qquad a_j = p_{jj}^{*} > 0 \qquad$ voor alle $j \in I$.

Bewijs

Uit punt d) van stelling 1.2.12 en uit stelling 1.2.20 volgt dat

(1.127) $\qquad p^*_{ij} = p^*_{jj} > 0 \qquad$ voor alle $i, j \in I$.

Met behulp van stelling 1.2.13 en relatie (1.127) vinden wij dat

(1.128) $\qquad p^*_{jj} = \sum_{k\in I} p^*_{jk}\, p^*_{kj} = p^*_{jj} \sum_{k\in I} p^*_{jk} = p^*_{jj} \sum_{k\in I} p^*_{kk}$

voor alle $j \in I$.

Aangezien $p^*_{jj} > 0$ voor $j \in I$, volgt uit (1.128) dat

(1.129) $\qquad \sum_{k\in I} p^*_{kk} = 1.$

Uit stelling 1.2.13 en relatie (1.127) volgt verder dat

(1.130) $\qquad p^*_{jj} = \sum_{k\in I} p^*_{jk}\, p_{kj} = \sum_{k\in I} p_{kj}\, p^*_{kk}$

voor alle $j \in I$.

De relaties (1.129) en (1.130) houden in dat $(p^*_{jj},\ j \in I)$ een invariante kansverdeling van de Markov-keten is.

Om aan te tonen dat deze kansverdeling uniek is, nemen wij aan dat $(a_j,\ j \in I)$ ook een invariante kansverdeling van de Markov-keten is. Dus de kansen a_j voldoen aan (1.124) en (1.125). Door de gelijkheid (1.125) herhaald toe te passen, vinden wij met behulp van (1.50) en punt b) van stelling 1.2.1 dat

(1.131) $\qquad a_j = \sum_{k\in I} p^{(n)}_{kj}\, a_k \qquad$ voor alle $j \in I$ en $n \geq 1$.

Uit (1.131) en punt b) van stelling 1.2.1 volgt dat

(1.132) $\qquad a_j = \sum_{k\in I} a_k \frac{1}{m} \sum_{n=1}^{m} p^{(n)}_{kj} \qquad$ voor alle $j \in I$ en $n \geq 1$.

Laten wij nu m naar ∞ gaan in (1.132) en passen wij punt c) stelling 1.2.1 toe, dan vinden wij met behulp van de relaties (1.124) en (1.127) dat

$$(1.333) \qquad a_j = \sum_{k\in I} a_k\, p^{*}_{kj} = p^{*}_{jj} \sum_{k\in I} a_k = p^{*}_{jj} \qquad \text{voor alle } j\in I.$$

Hiermee is de stelling bewezen.

Opmerking

Als de toestandsruimte I precies één kernfuik bevat, dan volgt uit de stellingen 1.2.6 en 1.2.14 dat de verzameling $I\setminus K$ (aangenomen dat deze niet leeg is) uit niet-essentiële doorgangstoestanden bestaat. Met behulp van punt c) van stelling 1.2.12 kan nu eenvoudig de volgende generalisatie van stelling 1.2.22 bewezen worden. Als de toestandsruimte I precies een kernfuik bevat en deze kernfuik bestaat uit positieve terugkeertoestanden, dan heeft de Markov-keten $(p^{*}_{jj},\ j\in I)$ als unieke invariante kansverdeling, waarin $p^{*}_{jj} > 0$ dan en slechts dan als j tot de kernfuik behoort.

Voor alle $i, j\in I$ definiëren wij

$$(1.134) \qquad {}_ip^{(n)}_{ij} = P\{\underline{x}_n = j,\ \underline{x}_m \neq i \text{ voor } 1 \le m < n \mid \underline{x}_0 = i\} \text{ voor } n \ge 1$$

en

$$(1.135) \qquad e_{ij} = \sum_{n=1}^{\infty} {}_ip^{(n)}_{ij}.$$

Uit deze definities volgt dat voor alle $i\in I$ geldt

$$(1.136) \qquad {}_ip^{(n)}_{ii} = f^{(n)}_{ii} \quad , \quad e_{ii} = f_{ii}\ .$$

Dus e_{ii} is de kans dat het systeem uitgaande van toestand i ooit in deze toestand terugkeert. Als $j \neq i$, kunnen wij e_{ij} interpreteren als het verwachte aantal keren dat een overgang naar toestand j plaatsvindt alvorens een overgang naar toestand i plaatsvindt, gegeven dat het systeem toestand i als begintoestand heeft. Deze interpretatie kunnen wij als volgt toelichten. Definieer $\underline{y}_n = 1$ als $\underline{x}_n = j$ en $\underline{x}_m \neq i$ voor $1 \leq m < n$ en definieeer $\underline{y}_n = 0$ anders $(n \geq 1)$.

Dan is $\sum_{n=1}^{\infty} \underline{y}_n$ het aantal keren dat een overgang naar toestand j plaatsvindt alvorens een overgang naar toestand i plaatsvindt. Omdat $E(\underline{y}_n | \underline{x}_0 = i) = P\{\underline{y}_n = 1 | \underline{x}_0 = i\} = {}_i p_{ij}^{(n)}$, vinden wij $E(\sum_{n=1}^{\infty} \underline{y}_n | \underline{x}_0 = i) = \sum_{n=1}^{\infty} E(\underline{y}_n | \underline{x}_0 = i) = e_{ij}$.

Uit de definitie van e_{ij} kan op eenvoudige wijze afgeleid worden

(1.137) $\qquad e_{ij} = 0 \qquad$ als $i \not\to j$.

Stelling 1.2.22a

Als i een terugkeertoestand is, dan geldt

(1.138) $$\sum_{j \in I} e_{ij} = \mu_{ii} .$$

Bewijs

Met behulp van punt b) van stelling 1.2.1 en de definities (1.134) en (1.135) vinden wij

(1.139) $$\sum_{j \in I} e_{ij} = \sum_{j \in I} \sum_{n=1}^{\infty} {}_i p_{ij}^{(n)} = \sum_{n=1}^{\infty} \sum_{j \in I} {}_i p_{ij}^{(n)}$$

$$= \sum_{n=1}^{\infty} P\{\underline{x}_m \neq i \text{ voor } 1 \leq m < n | \underline{x}_0 = i\}$$

Omdat i een terugkeertoestand is (dus $f_{ii} = 1$), geldt dat $P\{\underline{x}_m \neq i \text{ voor } 1 \leq m < n | \underline{x}_0 = i\} = \sum_{k=n}^{\infty} f_{ii}^{(k)}$. Hieruit volgt (vgl.(1.56))

(1.140) $$\sum_{j \in I} e_{ij} = \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} f_{ii}^{(k)} = \sum_{k=1}^{\infty} k f_{ii}^{(k)} = \mu_{ii} .$$

Wij merken op dat de relatie (1.138) intuitief direct duidelijk is uit de physische interpretatie van de getallen e_{ij}.

Stelling 1.2.226

Als K een kernfuik is bestaande uit positieve terugkeerteostanden, dan geldt

(1.141) $$e_{ij} = \frac{p^{*}_{jj}}{p^{*}_{ii}} \qquad \text{voor alle } i,\ j \in K.$$

Bewijs

Eerst merken wij op dat uit definitie (1.134) volgt

(1.142) $$\sum_{k \in I, k \neq i} {}_ip^{(n)}_{ik}\ p_{kj} = {}_ip^{(n+1)}_{ij} \qquad \text{voor alle } i,\ j \in I \text{ en } n \geq 1.$$

Zij i een vaste, doch willekeurig gekozen, toestand uit K. Wij zullen nu aantonen dat $(e_{ij},\ j \in I)$ een oplossing vormt van het stelsel vergelijkingen

(1.143) $$u_j = \sum_{k \in I} p_{kj}\ u_k \qquad \text{voor alle } j \in I.$$

Door gebruik te maken van punt b) van stelling 1.2.1 en relatie (1.142) vinden wij dat

(1.144) $$\sum_{k \in I} e_{ik}\ p_{kj} = \sum_{k \in I} p_{kj} \sum_{n=1}^{\infty} {}_ip^{(n)}_{ik} =$$

$$= \sum_{n=1}^{\infty} \sum_{k \in I} {}_ip^{(n)}_{ik}\ p_{kj} =$$

$$= \sum_{n=1}^{\infty} ({}_ip^{(n+1)}_{ij} + {}_ip^{(n)}_{ii}\ p_{ij}) =$$

$$= e_{ij} - {}_ip^{(1)}_{ij} + p_{ij}\ e_{ii} \qquad \text{voor alle } j \in I.$$

Uit de definities (1.134) en (1.135) volgt dat ${}_ip^{(1)}_{ij} = p_{ij}$ en $e_{ii} = f_{ii}$. Toestand i is een terugkeertoestand, dus $f_{ii} = 1$. Uit (1.144) volgt nu dat

(1.145) $$e_{ij} = \sum_{k \in I} e_{ik}\ p_{kj} \qquad \text{voor alle } j \in I.$$

Door een herhaalde toepassing van de gelijkheid (1.145), vinden wij dat

(1.146) $$e_{ij} = \sum_{k\in I} e_{ik}\, p_{kj}^{(n)} \qquad \text{voor alle } j\in I \text{ en } n \geq 1.$$

Uit (1.146) volgt dat

(1.147) $$e_{ij} = \sum_{k\in I} e_{ik} \frac{1}{m} \sum_{n=1}^{m} p_{kj}^{(n)} \quad \text{voor alle } j\in I \text{ en } m \geq 1.$$

Toestand i is een positieve terugkeertoestand, dus $\sum_{k\in I} e_{ik} < \infty$ (vgl. (1.138)). Passen wij nu punt c) van stelling 1.2.1 toe, dan vinden wij

(1.148) $$e_{ij} = \sum_{k\in I} e_{ik}\, p_{kj}^{*} \qquad \text{voor alle } j\in I.$$

Omdat K een kernfuik is, geldt voor elke toestand $k\in K$ dat $i \not\rightarrow k$, dus $e_{ik} = 0$ (vgl. (1.137)). Als $j\in K$ dan geldt $p_{kj}^{*} = p_{jj}^{*}$ voor alle $k\in K$ (stelling 1.2.20). Uit (1.84), (1.139) en (1.148) volgt nu

(1.149) $$e_{ij} = \sum_{k\in I} e_{ik}\, p_{kj}^{*} = \sum_{k\in K} e_{ik}\, p_{jj}^{*} =$$

$$= p_{jj}^{*} \sum_{k\in K} e_{ik} = p_{jj}^{*} \sum_{k\in I} e_{ik} =$$

$$= p_{jj}^{*}\, \mu_{ii} = \frac{p_{jj}^{*}}{p_{ii}^{*}} \qquad \text{voor alle } j\in I.$$

Zij i een terugkeertoestand. Wij kunnen dan een stochastische variabele $\underline{\nu}_i$ invoeren, die als volgt gedefinieerd wordt. Laat ω een realisering van het Markov-proces zijn. Dan definiëren wij

(1.150) $$\nu_i(\omega) = \min\,(n \mid n \geq 1,\ x_n(\omega) = i).$$

Stelling 1.2.22c

Zij K een kernfuik bestaande uit terugkeertoestanden en laat f een reële funktie op K zijn. Voor elke begintoestand $i \in K$ (d.w.z. $\underline{x}_0 = i$) geldt dan

$$(1.151) \qquad E\Big(\sum_{n=0}^{\underline{\nu}_i - 1} j(\underline{x}_n)\Big) = \sum_{j \in K} e_{ij}\, f(j),$$

aangenomen dat het rechterlid van (1.151) absoluut convergeert.

Het bewijs van deze stelling laten wij achterwege. Een bewijs kan gevonden worden in hoofdstuk 14 (stelling 5) van het boek van K.L. Chung, Markov-chains with stationary transition probabilities.

Ook de volgende belangrijke ergodenstelling bewijzen wij niet; voor een bewijs (berustend op stelling 1.2.22c en de sterke wet der grote aantallen) wordt de lezer verwezen naar hoofdstuk 15(stelling 1) van het hierboven genoemde boek van Chung.

Stelling 1.2.22d

Zij K een kernfuik bestaande uit terugkeertoestanden en laten f en g twee reële funkties op K zijn. Voor elke begintoestand $i \in K$ (d.w.z. $\underline{x}_0 = i$) geldt dan

$$(1.152) \qquad \lim_{n\to\infty} \frac{\sum_{k=0}^{n-1} f(\underline{x}_k)}{\sum_{k=0}^{n-1} g(\underline{x}_k)} = \frac{\sum_{j\in K} e_{ij}\, f(j)}{\sum_{j\in K} e_{ij}\, g(j)} \quad \text{met kans 1,}$$

aangenomen dat de noemer en de teller uit het rechterlid van (1.152) absoluut convergent zijn en niet beide nul zijn. Voorts geldt dat het rechterlid van (1.152) niet van de begintoestand i afhangt.

Deze stelling heeft als gevolg:

Stelling 1.2.22e

Zij K een kerfuik bestaande uit positieve terugkeertoestanden en laat f een reële funktie op K zijn. Voor elke begintoestand $i \in K$ (d.w.z. $\underline{x}_0 = i$) geldt dan

$$(1.153) \qquad \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\underline{x}_k) = \sum_{j\in K} p_{jj}^{*} f(j) \quad \text{met kans } 1,$$

aangenomen dat het rechterlid van (1.153) absoluut convergeert.

Bewijs

Neem 9 $\equiv$ 1 in stelling 1.2.22d en bedenk dat $e_{ij} = p_{jj}^{*} / p_{ii}^{*}$ voor alle $i, j \in K$ (stelling 1.2.22b) en dat $\sum_{j\in K} p_{jj}^{*} = 1$ (vgl. de afleiding van relatie (1.129)).

In het boek van J. Doob, Stochastic Processes, wordt in hoofdstuk 5 (stelling 6.2) de volgende stelling bewezen.

Stelling 1.2.23

Zij de toestandsruimte I eindig en laat f een eindige reële funktie op I zijn. Dan geldt voor elke begintoestand $i \in I$ dat de stochastische variabele

$$(1.154) \qquad \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\underline{x}_k)$$

met kans 1 bestaat en als verwachting heeft

$$(1.155) \qquad \sum_{j\in I} p_{ij}^{*} f(j).$$

Als de begintoestand i tot een priemfuik behoort, dan geldt tevens dat de stochastische variabele (1.154) met kans de waarde (1.155) aanneemt.

1.3 Niet-negatieve, vierkante matrices.

In deze paragraaf beschouwen wij vierkante matrices, waarvan de elementen niet-negatief zijn. Voor de matrices zullen wij de letters A, B en I gebruiken, waarbij I de eenheidsmatrix is. De elementen van de matrix A resp. B zijn de getallen a_{ij} resp. b_{ij}. De elementen van de matrix B^n (het n-voudig matrixproduct van B met zichzelf) geven wij aan met $b_{ij}^{(n)}$. De matrix B^0 identificeren wij met I.

Voor wij de benodigde stellingen over niet-negatieve matrices formuleren, bewijzen wij eerst enkele eenvoudige lemmas.

Lemma 1.3.1

Zij A een verzameling van reële getallen. Dan geldt

$$\inf_{a\in A} e^a = e^{\inf_{a\in A} a} . \tag{1.156}$$

Bewijs.

Stel ter afkorting, $a_0 = \inf_{a\in A} a$. Uit de monotonie van de exponentiële funktie volgt $e^{a_0} \leqq e^a$ voor alle $a\in A$, dus

$$e^{a_0} \leqq \inf_{a\in A} e^a . \tag{1.157}$$

Door een bewijs uit het ongerijmde tonen wij aan dat in (1.157) het gelijkheidsteken geldt. Stel dat het "< teken" geldt. Uit het continu zijn van de exponentiële funktie volgt dat dan een getal $\delta > 0$ bestaat, zodat

$$e^{a_0+\delta} < \inf_{a\in A} e^a . \tag{1.158}$$

Uit de definitie van infimum volgt dat een getal $a_1 \in A$ bestaat met $a_1 < a_0 + \delta$. Uit de monotonie van de exponentiële funktie en (1.158) volgt dat voor $a_1 \in A$ geldt $e^{a_1} < \inf_{a\in A} e^a$. Dit is een tegenspraak. Dus de relatie (1.156) is juist.

Lemma 1.3.2

Zij de machtreeks $\sum_{n=0}^{\infty} a_n x^n$ convergent voor $0 < x < 1$ en zij $a_n \geq 0$ voor alle $n \geq 0$. Als

(1.159) $$\lim_{x \uparrow 1} \sum_{n=0}^{\infty} a_n x^n = a$$

(waarin a oneindig mag zijn), dan geldt

(1.160) $$\sum_{n=0}^{\infty} a_n = a.$$

Bewijs

Voor alle x, $0 < x < 1$ geldt dat

(1.161) $$\sum_{n=0}^{\infty} a_n x^n \leq \sum_{n=0}^{\infty} a_n .$$

Uit (1.161) volgt direct dat de stelling juist is voor het geval $a = \infty$.

Beschouw nu het geval $a < \infty$. Met behulp van (1.161) vinden wij

(1.162) $$a = \lim_{x \uparrow 1} \sum_{n=0}^{\infty} a_n x^n \leq \sum_{n=0}^{\infty} a_n .$$

Omdat $a_n \geq 0$ voor alle $n \geq 0$, geldt

(1.163) $$\sum_{n=0}^{m} a_n x^n \leq \sum_{n=0}^{\infty} a_n x^n \qquad \text{voor alle } 0 < x < 1 \text{ en } m \geq 1 .$$

Met behulp van (1.163) vinden wij

(1.164) $$\sum_{n=0}^{m} a_n = \lim_{x \uparrow 1} \sum_{n=0}^{m} a_n x^n \leq \lim_{x \uparrow 1} \sum_{n=0}^{\infty} a_n x^n = a \qquad \text{voor alle } m \geq 1 .$$

Uit (1.164) en het feit dat $a_n \geq 0$ voor alle $n \geq 0$ volgt

(1.165) $$\sum_{n=0}^{\infty} a_n = \lim_{m \to \infty} \sum_{n=0}^{m} a_n \leq a.$$

Uit (1.162) en (1.165) volgt de bewering (1.160).

Definitie 1.3.1

Zij $B = (b_{ij})$ een N x N-matrix met niet-negatieve elementen. De norm van B is het (niet-negatieve) reële getal,

(1.166) $$||B|| = \max_{1 \leq i \leq N} \sum_{j=1}^{N} b_{ij} .$$

Lemma 1.3.3

Laten A en B twee N x N-matrices zijn met niet-negatieve elementen. Dan geldt

(1.167) $$||A\, B\,|| \leq ||A||\ \ ||B||.$$

Bewijs

Geven wij de elementen van het matrixprodukt AB aan met c_{ij}, dan geldt

(1.168) $$\sum_{j=1}^{N} c_{ij} = \sum_{j=1}^{N} \sum_{k=1}^{N} a_{ik} b_{kj} = \sum_{k=1}^{N} a_{ik} \sum_{j=1}^{N} b_{ij} \leq \sum_{k=1}^{N} a_{ik}\ ||B|| \leq$$

$$\leq ||A||\ \ ||B|| \qquad \text{voor alle } i = 1,\ldots,N .$$

Stelling 1.3.1

Laat B een N x N-matrix zijn met niet-negatieve elementen. Dan geldt dat $||B^n||^{\frac{1}{n}}$ een limiet heeft voor $n \to \infty$, welke gegeven wordt door

(1.169) $$\lim_{n\to\infty} ||B^n||^{\frac{1}{n}} = \inf_{n \geq 1} ||B^n||^{\frac{1}{n}} .$$

Bewijs

Als B^k gelijk aan de nulmatrix is voor een of andere k, dan is ook $B^n = 0$ voor $n \geq k$. De stelling is dan triviaal. Neem nu aan dat B^n voor alle n ongelijk aan de nulmatrix is. Dus $||B^n|| > 0$ voor alle $n \geqq 1$. Definieer

$$(1.170 \qquad a_n = \log ||B^n|| \qquad \text{voor alle } n \geqq 1.$$

Uit $||B^{n+k}|| = ||B^n \, B^k|| \leqq ||B^n|| \; ||B^k||$ (lemma 1.3.3) en de monotonie van de log-funktie volgt

$$(1.171) \qquad a_{n+k} \leqq a_n + a_k \qquad \text{voor alle } n,\, k \geqq 1 \, .$$

Zij m een vast, doch willekeurig gekozen natuurlijk getal. Ieder natuurlijk getal n kunnen wij schrijven als

$$(1.172) \qquad n = m\, q_n + r_n \, ,$$

waarin $q_n \geqq 0$ en $r_n \geqq 0$ gehele getallen zijn met $0 \leqq r_n \leqq m - 1$. Zij $M = \max (a_0, \ldots, a_{m-1})$. Door (1.171) toe te passen vinden wij met behulp van de representatie (1.172) dat

$$(1.173) \qquad a_n \leqq a_m \, q_n + M \qquad \text{voor alle } n \geqq 1 \, .$$

Uit (1.172) en (1.173) volgt

$$(1.174) \qquad \frac{a_n}{n} \leqq \frac{mq_n}{mq_n + r_n} \; \frac{a_m}{m} + \frac{M}{n} \qquad \text{voor alle } n \geqq 1 \, .$$

Uit (1.174) volgt dat

$$(1.175) \qquad \limsup_{n\to\infty} \frac{a_n}{n} \leqq \frac{a_m}{m} \, .$$

Aangezien m willekeurig gekozen is, geldt (1.175) voor elke $m \geqq 1$. Dus

$$(1.176) \qquad \limsup_{n\to\infty} \frac{a_n}{n} \leqq \inf_{n \geqq 1} \frac{a_n}{n} \, .$$

Omdat $\liminf\limits_{n\to\infty} \frac{a_n}{n} \geqq \inf\limits_{n\geqq 1} \frac{a_n}{n}$ en $\liminf\limits_{n\to\infty} \frac{a_n}{n} \leqq \limsup\limits_{n\to\infty} \frac{a_n}{n}$ volgt

uit (1.176) dat de getallen a_n/n een limiet bezitten, die gelijk is aan

$$(1.177)\qquad \lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\geq 1} \frac{a_n}{n} .$$

Uit (1.170), de continuïteit van de e-funktie, (1.177) en lemma 1.2.1 volgt dat

$$(1.178)\qquad \lim_{n\to\infty} ||B^n||^{\frac{1}{n}} = \lim_{n\to\infty} e^{\frac{a_n}{n}} = e^{\lim\limits_{n\to\infty} \frac{a_n}{n}} =$$

$$= \inf_{n\geq 1} e^{\frac{a_n}{n}} = \inf_{n\geq 1} ||B^n||^{\frac{1}{n}} .$$

<u>Stelling 1.3.2</u>

Zij B een N x N-matrix met niet-negatieve elementen. Stel verder dat $||B|| \leq 1$ is. De volgende zes uitspraken zijn dan equivalent

(i) $\lim\limits_{n\to\infty} ||B^n||^{\frac{1}{n}} < 1$,

(ii) $||B^n|| < 1$ voor tenminste één n,

(iii) $||B^N|| < 1$,

(iv) $\sum\limits_{n=0}^{\infty} B^n$ convergeert ,

(v) $(I - B)^{-1}$ bestaat ,

(vi) $\lim\limits_{n\to\infty} B^n = 0$.

<u>Bewijs</u>

<u>(i)→(ii)</u>

Uit stelling 1.2.1 en (i) volgt dat $\inf\limits_{n\geq 1} ||B^n||^{\frac{1}{n}} < 1$ is. Dus voor tenminste één n geldt $||B^n||^{\frac{1}{n}} < 1$, oftewel, $||B^n|| < 1$ voor tenminste één n.

(ii)→(iii)

Wij geven een kansrekenkundig bewijs. Een Markov-keten kan gedefinieerd worden met als toestandsruimte $\{1,\dots, N+1\}$ en met als één-stapsovergangswaarschijnlijkheden

$$(1.179)\qquad \begin{cases} p_{ij} = b_{ij} & \text{voor } i, j = 1,\dots, N, \\ p_{i,N+1} = 1 - \sum_{j=1}^{N} b_{ij} & \text{voor } i = 1,\dots, N, \\ p_{N+1,N+1} = 1,\ p_{N+1,i} = 0 & \text{voor } i = 1,\dots, N\,. \end{cases}$$

Omdat $p_{N+1,N+1} = 1$ geldt dat de k-stapsovergangswaarschijnlijkheden $p_{ij}^{(k)}$ gelijk zij aan $b_{ij}^{(k)}$ voor alle $i,j = 1,\dots, N$ en $k \geq 1$. Dus

$$(1.180)\qquad 1 = \sum_{j=1}^{N+1} p_{ij}^{(k)} = \sum_{j=1}^{N} b_{ij}^{(k)} + p_{i,N+1}^{(k)} \quad \text{voor alle } i = 1,\dots,N \text{ en } k \geq 1.$$

Uit (1.180) volgt dat voor alle $i = 1,\dots, N$ en $k \geq 1$ geldt

$$(1.181)\qquad \sum_{j=1}^{N} b_{ij}^{(k)} < 1 \iff p_{i,N+1}^{(k)} > 0\,.$$

Uit (ii) volgt dat een getal n, stel $n = m$, bestaat met $\sum_{j=1}^{N} b_{ij}^{(m)} < 1$ voor alle $i = 1,\dots, N$, oftewel, $p_{i,N+1}^{(m)} > 0$ voor $i = 1,\dots, N$. Hieruit kunnen wij concluderen dat toestand N+1 vanuit elke toestand $i = 1,\dots, N$ bereikbaar is. Omdat het aantal toestanden van de Markov-keten gelijk aan N+1 is en toestand N+1 vanuit elke toestand i bereikbaar is, kunnen wij bij elke toestand i een getal $n_i \leq N$ vinden met $p_{i,N+1}^{(n_i)} > 0$, zoals eenvoudig valt in te zien (vgl. 1.45). Als $p_{i,N+1}^{(k)} > 0$ is, dan is $p_{i,N+1}^{(n)} > 0$ voor alle $n \geq k$ (immers $p_{N+1,N+1} > 0$). Omdat $n_i \leq N$ voor alle $i = 1,\dots, N$, geldt dat $p_{i,N+1}^{(N)} > 0$ voor alle $i = 1,\dots, N$. Uit (1.171) en (1.181) volgt nu dat $||B^N|| < 1$ is.

(iii) → (iv)

Wij moeten, uitgaande van (iii), aantonen dat $\sum_{n=0}^{\infty} b_{ij}^{(n)}$ convergeert voor alle i, j. Omdat $b_{ij}^{(n)} \leqq ||B^n||$ kunnen wij volstaan met te bewijzen dat $\sum_{n=0}^{\infty} ||B^n||$ convergeert ($B^0 = I$).

Voor elke $n \geqq 0$ geldt

(1.182) $$\sum_{k=0}^{n} ||B^k|| = \sum_{k=0}^{N-1} ||B^k|| + \sum_{k=N}^{2N-1} ||B^k|| + \dots + \sum_{k=[n/N]}^{n} ||B^k||$$

Aangezien $||B^{k+mN}|| \leqq ||B^k||\, ||B^N||^m$ (lemma 1.3.3), volgt uit (1.182)

(1.183) $$\sum_{k=0}^{n} ||B^k|| \leqq \sum_{k=0}^{N-1} ||B^k|| \left(1 + ||B^N|| + ||B^N||^2 + \dots + ||B^N||^{[n/N]}\right)$$

$$\leqq \frac{1}{1-||B^N||} \sum_{k=0}^{N-1} ||B^k|| < \infty \quad \text{voor alle } n \geqq 0.$$

(iv) → (v)

Uit de convergentie van $\sum_{n=0}^{\infty} B^n$ volgt direct dat $\lim_{n\to\infty} B^n = 0$ (de nulmatrix). Uit de identiteit

(1.184) $$(I-B)\left(\sum_{k=0}^{n} B^k\right) = \left(\sum_{k=0}^{n} B^k\right)(I-B) = I - B^{n+1} \quad \text{voor } n \geqq 0$$

en $B^n \to 0$ voor $n \to \infty$ volgt dat de matrix $I - B$ een inverse bezit, welke gegeven wordt door

(1.185) $$(I-B)^{-1} = \sum_{n=0}^{\infty} B^n .$$

(v) → iv

Omdat $||B|| \leqq 1$ (merk op dat deze aanname hier voor het eerst gebruikt wordt) geldt $||\lambda B|| < 1$ voor elke $0 < \lambda < 1$. Dus (bedenk (ii) → (iii) → (iv) → (v))

(1.186) $$\sum_{n=0}^{\infty} \lambda^n B^n = (I-\lambda B)^{-1} \qquad \text{voor } 0 < \lambda < 1 .$$

Volgens (v) geldt dat $(I-B)^{-1}$ bestaat, dus de determinant van de matrix I - B is ongelijk aan nul. Met behulp van de uit de matrixtheorie welbekende formule $A^{-1} = \frac{1}{\det A} A^{+}$, waarbij A^{+} de cofactorenmatrix van de matrix A is, vinden wij vervolgens de relatie

$$(1.187) \qquad \lim_{\lambda \uparrow 1} (I-\lambda B)^{-1} = (I-B)^{-1} .$$

Uit (1.186), (1.187) en lemma 1.3.2 volgt dat $\sum_{n=0}^{\infty} B^n = (I-B)^{-1}$. Hiermee is (iv) aangetoond.

<u>(iv) → (vi)</u>

Uit de convergentie van $\sum_{n=0}^{\infty} B^n$ volgt direct (vi).

<u>(vi) → (i)</u>

Uit $\lim_{n\to\infty} B^n = 0$ volgt het bestaan van een getal n_0 zodat

$$(1.188) \qquad \|B^n\| < 1 \qquad \text{voor alle } n \geqq n_0 .$$

Uit (1.188) en stelling 1.2.1 volgt dat

$$(1.189) \qquad \lim_{n\to\infty} \|B^n\|^{\frac{1}{n}} = \inf_{n\geqq 1} \|B^n\|^{\frac{1}{n}} < 1 .$$

Hiermee is het bewijs van de stelling voltooid.

<u>Opmerking</u>

Uit het bewijs van deze stelling blijkt dat als een van de uitspraken (i)-(vi) geldt, dat dan

$$(1.190) \qquad (I-B)^{-1} = \sum_{n=0}^{\infty} B^n .$$

Hoofdstuk 2. Eindige Markovbeslissingsproblemen.

2.1 Definities

De meetbare ruimten die we gebruiken zullen steeds zó zijn dat de verzameling van elementen een borelverzameling is in een eindig-dimensionale Euclidische ruimte (notatie: $\mathbb{R}^k$, k natuurlijk getal), zeg X, en de σ-algebra der meetbare verzamelingen de borelverzamelingen zijn die in X bevat zijn. We zullen zowel de verzameling van elementen als de meetbare ruimte met hetzelfde symbool aangeven.

De verzameling van alle kansmaten op een meetbare ruimte X duiden we aan met $\mathbb{P}(X)$. Een voorwaardelijke kansmaat op een meetbare ruimte Y t.o.v. een meetbare ruimte X is een funktie $q(.|.)$ zó dat voor iedere $x \in X$, $q(.|x) \in \mathbb{P}(Y)$ en voor iedere borelverzameling $B \subset Y$ $q(B|.)$ een meetbare funktie op X is. $\mathbb{P}(Y|X)$ is de aanduiding voor de verzameling van alle voorwaardelijke kansmaten op Y t.o.v. X.

Een <u>Markovbeslissingsproces</u> wordt gedefinieerd door drie objekten: S, A, q, waarin: S en A meetbare ruimten, $q \in \mathbb{P}(S|SA)$ met SA de produktruimte van S en A.
S wordt de toestandsruimte en A de aktieruimte (beslissingsruimte) genoemd; q noemen we de bewegingswet.

Een <u>strategie</u> (beslissingsregel) R is een rij $(\pi_0, \pi_1, \pi_2, \ldots)$ met $\pi_n \in \mathbb{P}(A_n|S_0A_0S_1A_1 \ldots S_{n-1}A_{n-1}S_n)$ waarin S_i en A_i $(i = 0,\ldots,n)$ copieën van S en A zijn.
Elementen uit $S_0A_0S_1A_1\ldots S_{n-1}A_{n-1}S_n$ duiden we aan met $(x_0,y_0,x_1,y_1,\ldots,x_{n-1},y_{n-1},x_n)$ en we zullen x_i resp. y_i $(i = 0,\ldots,n)$ de toestand resp. aktie op de i-de dag noemen.
De verzameling van alle strategieën duiden we aan met C.
Strategieën zó dat van iedere n en $B \subset A_n$ $\pi_n(B|x_0,y_0,x_1,y_1,\ldots,x_{n-1},y_{n-1},x_n)$ alleen afhankelijk is van x_n worden gerandomiseerde Markovstrategieën genoemd. De verzameling van deze strategieën duiden we aan met C_{rm}.
Voor een gerandomiseerde Markovstrategie, zeg $(\pi_0, \pi_1, \pi_2, \ldots)$, geldt: $\pi_n \in \mathbb{P}(A|S)$ voor n = 0, 1, 2, ... Indien bovendien geldt dat $\pi_n = \pi$ met $\pi \in \mathbb{P}(A|S)$ voor n = 0, 1, 2, ..., dan noemen we de strategie een gerandomiseerde stationaire strategie. De verzameling van gerandomi-

seerde stationaire strategieën duiden we aan met C_{rs} en elementen uit C_{rs} worden ook wel als π^{∞} genoteerd.
De klasse van alle meetbare afbeeldingen van S naar A duiden we aan met F. Een $\pi \in \mathbb{P}(A|S)$ noemen we ontaard indien er een $f \in F$ bestaat zó dat $\pi(\{f(s)\}|s) = 1$ voor alle $s \in S$. Een gerandomiseerde Markovstrategie $(\pi_0, \pi_1, \pi_2, \ldots)$ waarvoor geldt dat π_n ontaard is voor $n = 0,1,2,\ldots$, noemen we een Markovstrategie. De verzameling van Markovstrategieën duiden we aan met C_m en elementen uit C_m worden ook wel genoteerd als $(f_0, f_1, f_2, \ldots)$.
Een Markovstrategie $(f_0, f_1, f_2, \ldots)$ waarvoor $f_n = f \in F$ voor $n = 0,1,2,\ldots$ wordt een stationaire strategie genoemd. De verzameling van alle stationaire strategieën duiden we aan met C_s en elementen uit C_s worden ook wel genoteerd als f^{∞}.
Als we $\supset$ noteren als $\longrightarrow$ dan vinden we voor de klassen van strategieën de volgende graaf.

$$\begin{tikzcd}
 & C \arrow[d] & \\
 & C_{rm} \arrow[dl] \arrow[dr] & \\
C_m \arrow[dr] & & C_{rs} \arrow[dl] \\
 & C_s &
\end{tikzcd}$$

Een Markovbeslissingsproces gecombineerd met een $R \in C$ en een beginverdeling $p \in \mathbb{P}(S)$ induceert een kansmaat op $S_0A_0S_1A_1\ldots$.
Deze kansmaat, die we aanduiden met $P_{p,R}$, wordt eenduidig vastgelegd indien de kans op de meetbare verzamelingen in $S_0A_0S_1A_1\ldots$ van het type $BS_{n+1}A_{n+1}\ldots$, waarin B een borelverzameling in $S_0A_0S_1A_1\ldots S_nA_n$ en n een willekeurig natuurlijk getal is, wordt gegeven door:

$$P_{p,R}(B_nS_{n+1}A_{n+1}\ldots) = \int\limits_{S_0A_0}\cdots\int\limits_{S_nA_n} p(dx_0)\pi_0(dy_0|x_0)q(dx_1|x_0,y_0) \ldots$$

$$\ldots q(dx_n|x_{n-1},y_{n-1})\pi_n(dy_n|x_0,y_0,\ldots,x_{n-1},y_{n-1},x_n)\chi_B(x_0,y_0,\ldots,x_n,y_n). \quad 1)$$

Voor B meetbare verzameling in $S_0A_0S_1A_1\ldots$ geldt dat $P_{p,R}(B|x_0=s_0)$ onafhankelijk is van de beginverdeling p; daarom laten we voor deze kansen de index p in $P_{p,R}$ weg.

1) $\chi_B(x_0,y_0,\ldots,x_n,y_n) = 1$ als $(x_0,y_0,\ldots,x_n,y_n) \in B$
$\chi_B(x_0,y_0,\ldots,x_n,y_n) = 0$ als $(x_0,y_0,\ldots,x_n,y_n) \notin B$

De toestand resp. aktie op de n-de dag (x_n resp. y_n) kunnen we zien als de projekties van $S_0A_0S_1A_1\dots$ op S_n resp. A_n. Door de kansmaat $P_{p,R}$ op $S_0A_0S_1A_1\dots$ worden deze meetbare afbeeldingen stochastische variabelen. De rij stochastische variabelen x_0, y_0, x_1, y_1, ... vormt dus een stochastisch proces en een Markovbeslissingsproces kan gezien worden als een klasse van stochastische processen; immers voor iedere R,p is er een stochastisch proces gedefinieerd.

Indien $R = (\pi_0, \pi_1, \dots)$ een gerandomiseerde Markovstrategie is, dan geldt:

$$P_R(x_{n_1} \in B_1, x_{n_2} \in B_2, \dots, x_{n_k} \in B_k | x_0 = s_0) =$$

$$= \int_{B_1} P_R(dx_{n_1} | x_0 = s_0) \int_{B_2} P_R(dx_{n_2} | x_{n_1}) \dots \int_{B_k} P_R(dx_{n_k} | x_{n_{k-1}}) ,$$

waarin $n_1 < n_2 < \dots < n_k$ en B_{n_i} borelverzameling in S_{n_i}.
Uit bovenstaande gelijkheid volgt dat x_0, x_1, x_2, ... een Markovproces vormt; dit verklaart de naam Markovbeslissingsproces.

We introduceren verder een reëelwaardige meetbare funktie op SAS' , zeg $r^*(s,a,s')$, die we de onmiddellijke-opbrengstfunktie zullen noemen. Als de toestand op een bepaalde dag s is, de aktie die op die dag wordt ondernomen a is en de toestand op de volgende dag s' zal zijn, dan wordt volgens dit model een onmiddellijke opbrengst $r^*(s,a,s')$ ontvangen.
We definiëren r(s,a) als $\int_S q(ds'|s,a)r^*(s,a,s')$; dit is de verwachting van de opbrengst op bepaalde dag indien de toestand op die dag s is en de ondernomen aktie a. Deze verwachting van de opbrengst op bepaalde dag, gegeven de toestand en de aktie, is onafhankelijk van de gebruikte strategie.

Met een begrensde r kunnen we nu een aantal opbrengstfunkties definiëren:

1. $V_{\beta,R}(s) = \sum_{n=0}^{\infty} \beta^n E_R(r(x_n,y_n)|x_0 = s)$, waarin $0 < \beta < 1$ en

   $E_R(r(x_n,y_n)|x_0 = s)$ de verwachting van de opbrengst op de n-de dag, als we starten in s en strategie R toepassen.

   $V_{\beta,R}(s)$ is de verwachte verdisconteerde opbrengst.

2. $V_R^N(s) = \sum_{n=0}^{N} E_R(r(x_n,y_n)|x_0 = s)$

   Dit is de verwachte opbrengst van de eerste N+1 dagen.

3. $V_{1,R}(s\ ) = \lim_{N\to\infty} V_R^N(s\ )$, in die gevallen waarin de limiet bestaat (eventueel $\pm\infty$).

   $V_{1,R}(s\ )$ is de <u>totale verwachte opbrengst</u>.

4. $g_R(s\ ) = \liminf_{N \to \infty} \frac{1}{N+1} V_R^N(s\ )$.

5. $G_R(s\ ) = \limsup_{N \to \infty} \frac{1}{N+1} V_R^N(s\ )$.

   Indien $g_R(s\ ) = G_R(s\ )$ dan spreken we van de <u>gemiddelde verwachte opbrengst</u>.

Een Markovbeslissingsproces gecombineerd met een opbrengstfunktie geeft een <u>Markovbeslissingsprobleem</u>.
Zo is voor de opbrengstfunktie $V_{\beta,R}(s)$ de vraag: bestaat er een $R_0 \in C$ waarvoor

$$* \qquad V_{\beta,R_0}(s\ ) = \sup_{R\in C} V_{\beta,R}(s) \qquad \text{voor alle } s \in S.$$

Een R_0 die aan * voldoet, noemen we optimaal t.o.v. $V_{\beta,R}$.

Naast deze existentievraag is er dan bovendien het probleem om een algorithme aan te geven dat in een eindig aantal stappen de optimale strategie, zo die bestaat, bepaalt.

Sommige auteurs beschouwen i.p.v. opbrengsten kosten; in dat geval wordt het probleem het vinden van een strategie die minimaliseert. Deze problemen zijn equivalent met de onze en zijn dan ook analoog op te lossen.

Markovbeslissingsproblemen waarvoor geldt dat S en A eindige verzamelingen zijn noemen we <u>eindige</u> Markovbeslissingsproblemen.
Bij eindige Markovbeslissingsproblemen, waartoe we ons in dit hoofdstuk zullen beperken, kunnen we integratie door sommatie vervangen. We vinden dan:

$$V_{\beta,R}(s_0) = \sum_{n=0}^{\infty} \beta^n \sum_{s,a} P_R(x_n=s,\ y_n=a \mid x_0=s_0)\ r(s,a)$$

en $$V_R^N(s_0) = \sum_{n=0}^{\infty} \sum_{s,a} P_R(x_n=s,\ y_n=a \mid x_0=s_0)\ r(s,a).$$

2.2. Verwachte verdisconteerde opbrengst.

We zullen in deze paragraaf bewijzen dat er een $f \in F$ bestaat zódat f^∞ optimaal t.o.v. $V_{\beta,R}$. In deze paragraaf is β een vast getal met $0 < \beta < 1$.

Stelling 1 (Derman en Strauch)
Voor iedere $R \in C$ en $s_0 \in S$ is er een $R^* \in C_{rm}$ zó dat

(1) $$P_{R^*}(x_n=s, y_n=a/x_0=s_0) = P_R(x_n=s, y_n=a/x_0=s_0) \quad \text{voor alle } s,a,n.$$

Bewijs
Voor $n = 0,1,2,\ldots$ definiëren we $\pi_n^* \in \mathbb{P}(A/S)$ zó dat

$$\pi_n^*(a/s) = P_R(y_n=a/x_0=s_0, x_n=s).$$

Zij $R^* = (\pi_0^*, \pi_1^*, \ldots)$ dan $R^* \in C_{rm}$.

Met volledige induktie bewijzen we dat (1) juist is voor $n = 0,1,2,\ldots$.

i) Voor $n = 0$.

Indien $s \neq s_0$ dan $P_R(x_0=s, y_0=a/x_0=s_0) = 0 = P_{R^*}(x_0=s, y_0=a/x_0=s_0)$.

Indien $s = s_0$ dan
$$\begin{aligned} P_R(x_0=s_0, y_0=a/x_0=s_0) &= P_R(y_0=a/x_0=s_0) \\ &= \pi_0^*(a/s_0) \\ &= P_{R^*}(y_0=a/x_0=s_0) \\ &= P_{R^*}(x_0=s_0, y_0=a/x_0=s_0) \end{aligned}$$

ii) Stel (1) is juist voor n gelijk aan N, dan

$$\begin{aligned} P_R(x_{N+1}=s, y_{N+1}=a/x_0=s_0) &= P_R(y_{N+1}=a/x_0=s_0, x_{N+1}=s).P_R(x_{N+1}=s/x_0=s_0) \\ &= \pi_{N+1}^*(a/s) \sum_{s',a'} P_R(x_N=s', y_N=a'/x_0=s_0) q(s/s',a') \end{aligned}$$

$$= \pi^{*}_{N+1}(a/s) \sum_{s',a'} P_{R^{*}}(x_N=s',y_N=a'/x_0=s_0)q(s/s',a')$$

$$= \pi^{*}_{N+1}(a/s)P_{R^{*}}(x_{N+1}=s/x_0=s_0)$$

$$= P_{R^{*}}(x_{N+1}=s,y_{N+1}=a/x_0=s_0).$$

Stelling 2

Indien R_0 optimaal in C_{rm} (t.o.v. $V_{\beta,R}$) dan R_0 optimaal in C.

Bewijs

Zij R een willekeurig element uit C en s_0 een willekeurig element uit S. Bij R,s_0 is er volgens stelling 1 een $R^{*} \in C_{rm}$ zó dat

$$P_{R^{*}}(x_n=s,y_n=a/x_0=s_0) = P_R(x_n=s,y_n=a/x_0=s_0) \quad \text{voor alle } s,a,n$$

$$\Longrightarrow V_{\beta,R}(s_0) = V_{\beta,R^{*}}(s_0).$$

Daar $R^{*} \in C_{rm}$ volgt uit het gegeven dat $V_{\beta,R^{*}}(s_0) \leq V_{\beta,R_0}(s_0)$. En dus $V_{\beta,R}(s_0) \leq V_{\beta,R_0}(s_0)$.

In het vervolg van deze paragraaf tot aan stelling 5 zullen we ons beperken tot de klasse van strategieën C_{rm}.

Notatie 1

Zeg k is het aantal toestanden in S. We noteren de vektor in $\mathbb{R}^k$ waarvan de s-de component gelijk is aan $V_{\beta,R}(s)$ met V_R.

Zij $\pi \in \mathbb{P}(A/S)$ en $f \in F$ dan is $r(\pi)$ resp. $r(f)$ de notatie van de vektoren met s-de component

$$\sum_a \pi(a/s)r(s,a) \quad \text{resp.} \quad r(s,f(s))$$

en $P(\pi)$ resp. $P(f)$ de notatie voor de k bij k matrices met (s,s')-de element

$$\sum_a \pi(a/s)q(s'/s,a) \quad \text{resp.} \quad q(s'/s,f(s)).$$

Indien $u,v \in \mathbb{R}^k$ en $u(s) \leq v(s)$ voor alle s dan noteren we $u \leq v$. Is $u \leq v$ en geldt voor minstens één s $u(s) < v(s)$ dan notatie $u < v$.

<u>Definitie 1</u>

Voor $\pi \in \mathbb{P}(A/S)$ zij $L(\pi)$ de afbeelding van $\mathbb{R}^k$ naar $\mathbb{R}^k$ met

$$L(\pi)v = r(\pi) + \beta P(\pi)v, \quad v \in \mathbb{R}^k$$

voor $f \in F$ zij $L(f)$ de afbeelding van $\mathbb{R}^k$ naar $\mathbb{R}^k$ met

$$L(f)v = r(f) + \beta P(f)v, \quad v \in \mathbb{R}^k.$$

Bij een $f \in F$ is er een $\pi \in \mathbb{P}(A/S)$ met $\pi(f(s)/s) = 1$ voor alle s. Voor deze combinatie van f en π geldt $L(f) \equiv L(\pi)$. Dit betekent dat eigenschappen die gelden voor iedere $L(\pi)$ met $\pi \in \mathbb{P}(A/S)$ zeker gelden voor iedere $L(f)$ met $f \in F$.

<u>Lemma 1</u>

De afbeeldingen $L(\pi)$ zijn monotoon d.w.z. uit $u \leq v$ volgt $L(\pi)u \leq L(\pi)v$.

<u>Bewijs</u>

Daar alle elementen uit $P(\pi)$ niet negatief zijn volgt uit $u \leq v$ dat

$$\begin{aligned} & P(\pi)u \leq P(\pi)v \\ \Longrightarrow\ & \beta P(\pi)u \leq \beta P(\pi)v \\ \Longrightarrow\ & r(\pi) + \beta P(\pi)u \leq r(\pi) + \beta P(\pi)v \ . \end{aligned}$$

<u>Lemma 2</u>

Voor $R = (\pi_0, \pi_1, \ldots)$ een element uit C_{rm} geldt

i) de matrix met (s_0,s)-de element $P_R(x_n=s/x_0=s_0)$ is gelijk aan $P(\pi_0)P(\pi_1)\ldots P(\pi_{n-1})$

ii) $V_R = r(\pi_0) + \sum\limits_{n=1}^{\infty} \beta^n P(\pi_0)\ldots P(\pi_{n-1})r(\pi_n)$

iii) $V_R = L(\pi_0)\ldots L(\pi_N)V_{(\pi_{N+1},\pi_{N+2},\ldots)}$ voor ieder natuurlijk getal N.

Bewijs

De bewering i) is equivalent met te zeggen dat het proces een niet-stationair Markov proces is met overgangsmatrices $P(\pi_n)$. We bewijzen i) met behulp van volledige inductie. We tonen alleen de inductiestap aan.

Stel i) is juist voor n gelijk aan N dan

$$P_R(x_{N+1}=s/x_0=s_0) = \sum_{s',a} P_R(x_{N+1}=s/x_0=s_0, x_N=s', y_N=a) . P_R(x_N=s', y_N=a/x_0=s_0)$$

$$= \sum_{s',a} q(s/s',a) P_R(x_N=s'/x_0=s_0)\pi_N(a/s')$$

$$= \sum_{s'} P_R(x_N=s'/x_0=s_0) \sum_a \pi_N(a/s') q(s/s',a)$$

$$= ((P(\pi_0)\ldots P(\pi_{N-1}))P(\pi_N))_{s_0,s} .$$

ii) $$V_R(s_0) = \sum_{n=0}^{\infty} \beta^n \sum_{s,a} P_R(x_n=s, y_n=a/x_0=s_0) r(s,a)$$

$$= \sum_{n=0}^{\infty} \beta^n \sum_{s,a} P_R(x_n=s/x_0=s_0)\pi_n(a/s) r(s,a) \qquad (R \in C_{rm})$$

$$= \sum_{n=0}^{\infty} \beta^n \sum_{s} P_R(x_n=s/x_0=s_0)(\sum_a \pi_n(a/s) r(s,a))$$

$$= \sum_{s} P_R(x_0=s/x_0=s_0)(r(\pi_0))_s +$$

$$+ \sum_{n=1}^{\infty} \beta^n \sum_{s} (P(\pi_0)\ldots P(\pi_{n-1}))_{s_0,s}(r(\pi_n))_s$$

$$= (r(\pi_0))_{s_0} + \sum_{n=1}^{\infty} \beta^n (P(\pi_0)\ldots P(\pi_{n-1}) r(\pi_n))_{s_0} .$$

iii) $$V_R = r(\pi_0) + \sum_{n=1}^{\infty} \beta^n P(\pi_0)\ldots P(\pi_{n-1}) r(\pi_n)$$

$$= r(\pi_0) + \beta P(\pi_0) r(\pi_1) + \sum_{n=2}^{\infty} \beta^n P(\pi_0)\ldots P(\pi_{n-1}) r(\pi_n)$$

$$= r(\pi_0) + \beta P(\pi_0)\{r(\pi_1) + \sum_{n=2}^{\infty} \beta^{n-1}P(\pi_1)\ldots P(\pi_{n-1})r(\pi_n)\}$$

$$= r(\pi_0) + \beta P(\pi_0)V_{(\pi_1,\pi_2,\ldots)}$$

$$= L(\pi_0)V_{(\pi_1,\pi_2,\ldots)} \ .$$

Passen we deze gelijkheid toe op $V_{(\pi_1,\pi_2,\ldots)}$ dan vinden we $V_{(\pi_1,\pi_2,\ldots)} = L(\pi_1)V_{(\pi_2,\pi_3,\ldots)}$ waaruit volgt: $V_{(\pi_0,\pi_1,\ldots)} =$ $= L(\pi_0).L(\pi_1)V_{(\pi_2,\pi_3,\ldots)}$. Zo is in (N+1)-stappen te bewijzen dat $V_R = L(\pi_0)\ldots L(\pi_N)V_{(\pi_{N+1},\pi_{N+2},\ldots)}$.

Lemma 3

Voor iedere $v \in \mathbb{R}^k$ en $R = (\pi_0,\pi_1,\ldots) \in C_{rm}$ geldt:

$$V_R = \lim_{N\to\infty} L(\pi_0)\ldots L(\pi_N)v \ .$$

Bewijs

$$L(\pi_N)v = r(\pi_N) + \beta P(\pi_N)v$$

$$\Longrightarrow L(\pi_{N-1})L(\pi_N)v = r(\pi_{N-1}) + \beta P(\pi_{N-1})r(\pi_N) + \beta^2 P(\pi_{N-1})P(\pi_N)v$$

....

$$\Longrightarrow L(\pi_0)\ldots L(\pi_N)v = r(\pi_0) + \sum_{n=1}^{N} \beta^n P(\pi_0)\ldots P(\pi_{n-1})r(\pi_n) +$$

$$+ \beta^{N+1}P(\pi_0)\ldots P(\pi_N)v \ .$$

Uit de definitie van V_R volgt dat $V_R = \lim_{N\to\infty}\{r(\pi_0) + \sum_{n=1}^{N} \beta^n P(\pi_0)\ldots P(\pi_{n-1})r(\pi_n)\}$.

Zij $M = \max_s |v_s|$ dan $|\beta^{N+1}(P(\pi_0)\ldots P(\pi_N)v)_s| \leq \beta^{N+1}M$ voor alle s

$\Longrightarrow \lim_{N\to\infty} \beta^{N+1}(P(\pi_0)\ldots P(\pi_N)v)_s = 0$ voor alle s.

Lemma 4

Voor iedere $v \in \mathbb{R}^k$ en $\pi \in \mathbb{P}(A/S)$ geldt, dat

i) $L(\pi)v \leq v \Longrightarrow V_{\pi^\infty} \leq v$

ii) $L(\pi)v > v \Longrightarrow V_{\pi^\infty} > v$.

Bewijs

Achtereenvolgens N maal toepassen van $L(\pi)$ op de vektor v noteren we als $L^N(\pi)v$.

i) Uit $L(\pi)v \leq v$ volgt m.b.v. de monotonie van de afbeelding $L(\pi)$ (lemma 1) dat $L^2(\pi)v = L(\pi).L(\pi)v \leq L(\pi)v \leq v$. Passen we dit N maal toe dan vinden we $L^N(\pi)v \leq v$.

Dus voor ieder natuurlijk getal N geldt dat $L^N(\pi)v \leq v$.

Hieruit volgt m.b.v. lemma 3 $V_{\pi^\infty} = \lim_{N\to\infty} L^N(\pi)v \leq v$.

ii) Uit $L(\pi)v > v$ volgt door N maal toepassen van de operatie $L(\pi)$ dat $L^N(\pi)v \geq L(\pi)v > v$.

Hieruit volgt dat $V_{\pi^\infty} = \lim_{N\to\infty} L^N(\pi)v \geq L(\pi)v > v$ en dus $V_{\pi^\infty} > v$.

Definitie 2

$$G(s,f) = \{a \in A | r(s,a) + \beta p(s,a)V_{f^\infty} > V_{f^\infty}(s)\} \quad \text{waarin } p(s,a)$$

de vektor is met s'-de component $q(s'|s,a)$.

Lemma 5

Voor $f,g \in F$ geldt dat

i) $g(s) = f(s) \Longrightarrow V_{(g,f,f,\ldots)}(s) \overset{\text{NOT.}}{=} V_{(g,f^\infty)}(s) = V_{f^\infty}(s)$

ii) $g(s) \in G(s,f) \Longleftrightarrow V_{(g,f^\infty)}(s) > V_{f^\infty}(s)$.

Bewijs

Volgens de opmerking na definitie 1 en lemma 2 geldt dat

$$V_{(g,f^\infty)} = L(g)V_{f^\infty} = r(g) + \beta P(g)V_{f^\infty}$$

$$\Longrightarrow V_{(g,f^\infty)}(s) = r(s,g(s)) + \beta p(s,g(s))V_{f^\infty}$$

i) indien $f(s) = g(s)$ dan

$$V_{(g,f^\infty)}(s) = r(s,g(s)) + \beta p(s,g(s))V_{f^\infty}$$

$$= r(s,f(s)) + \beta p(s,f(s))V_{f^\infty}$$

$$= V_{(f,f^\infty)}(s) = V_{f^\infty}(s).$$

ii) Volgens de definitie van $G(s,f)$ geldt dat $g(s) \in G(s,f)$ dan en slechts dan als

$$V_{(g,f^\infty)}(s) = r(s,g(s)) + \beta p(s,g(s))V_{f^\infty} > V_{f^\infty}(s).$$

Lemma 6

Indien voor $f,g \in F$ en $S_0 \subset S$ met $S_0 \neq \emptyset$ geldt dat

$$g(s) \in G(s,f) \quad \text{voor } s \in S_0$$

$$\text{en } g(s) = f(s) \quad \text{voor } s \notin S_0$$

dan $V_{g^\infty} > V_{f^\infty}$.

Bewijs

Passen we lemma 5 toe dan vinden we dat

$$V_{(g,f^\infty)}(s) > V_{f^\infty}(s) \quad \text{voor } s \in S_0$$

$$\text{en } V_{(g,f^\infty)}(s) = V_{f^\infty}(s) \quad \text{voor } s \notin S_0.$$

Hieruit volgt dat $L(g)V_{f^\infty} = V_{(g,f^\infty)} > V_{f^\infty}$.

Volgens lemma 4 (zie de opmerking na definitie 1) geldt dan dat $V_{g^\infty} > V_{f^\infty}$.

Stelling 3

Er is een $f \in F$ waarvoor $G(s,f) = \emptyset$ voor alle s.

Bewijs

We construeren inductief een rij $\{f_n\}_{n=1}^{\infty} \subset F$ door

i) kies f_1 willekeurig uit F

ii) indien $G(s,f_n) = \emptyset$ dan nemen we $f_{n+1}(s)$ gelijk aan $f_n(s)$

iii) indien $G(s,f_n) \neq \emptyset$ kies dan voor $f_{n+1}(s)$ een element uit $G(s,f_n)$

Volgens lemma 6 geldt dat, indien $G(s,f_n) \neq \emptyset$ voor alle s dan
$V_{f_{n+1}^{\infty}} > V_{f_n^{\infty}}$.

Nemen we aan dat er geen enkele n is waarvoor $G(s,f_n) = \emptyset$ voor alle s dan vinden we dat $V_{f_1^{\infty}} < V_{f_2^{\infty}} < \ldots$. Dit zou betekenen dat alle elementen uit de rij $\{f_n\}_{n+1}^{\infty}$ verschillend zijn. Echter dit is in tegenspraak met het feit dat F slechts eindig veel elementen bevat.

Lemma 7

Indien $G(s,f) = \emptyset$ voor alle s dan geldt dat

$$L(\pi)V_{f^{\infty}} \leq V_{f^{\infty}} \quad \text{voor alle } \pi \in \mathbb{P}(A/S).$$

Bewijs

We moeten bewijzen dat $V_{(\pi,f^{\infty})} = L(\pi)V_{f^{\infty}} = r(\pi) + \beta P(\pi)V_{f^{\infty}} \leq V_{f^{\infty}}$.
Deze "vektor" vergelijking zullen we bewijzen door de ongelijkheid voor een willekeurige component te bewijzen.
In het bewijs wordt gebruik gemaakt van het feit dat

$$r(s,a) + \beta p(s,a)V_{f^{\infty}} \leq V_{f^{\infty}}(s) \quad \text{voor alle a.}$$

Dit is een onmiddellijk gevolg van $G(s,f) = \emptyset$ voor alle s.

$$\begin{aligned} V_{(\pi,f^{\infty})}(s) &= \sum_a \pi(a/s)r(s,a) + \beta \sum_a \pi(a/s)p(s,a)V_{f^{\infty}} \\ &= \sum_a \pi(a/s)\{r(s,a) + \beta p(s,a)V_{f^{\infty}}\} \\ &\leq \sum_a \pi(a/s)V_{f^{\infty}}(s) = V_{f^{\infty}}(s). \end{aligned}$$

Stelling 4

Indien voor $f \in F$ geldt dat $G(s,f) = \emptyset$ voor alle s dan is f^∞ optimaal (t.o.v. $V_{\beta,R}$) in C_{rm}.

Bewijs

Te bewijzen is dat f^∞ optimaal in C_{rm} d.w.z. we moeten bewijzen dat

(2) $$V_{f^\infty} \geq V_R \quad \text{voor iedere } R \in C_{rm}.$$

We zullen dit doen door voor een willekeurige $R \in C_{rm}$ ongelijkheid (2) te bewijzen.

Zij $R = (\pi_0, \pi_1, \ldots) \in C_{rm}$.
Uit de lemma's 1 en 7 volgt dat voor ieder natuurlijk getal N geldt dat

$$L(\pi_0)\ldots L(\pi_N)V_{f^\infty} \leq V_{f^\infty}.$$

Hieruit volgt $V_R = \lim_{N\to\infty} L(\pi_0)\ldots L(\pi_N)V_{f^\infty} \leq V_{f^\infty}$.

Volgens stelling 2 geldt dan ook dat:

Stelling 5

Indien voor $f \in F$ geldt dat $G(s,f) = \emptyset$ voor alle s dan is f^∞ optimaal (t.o.v. $V_{\beta,R}$) in C.

In stelling 3 wordt aangetoond dat er een $f \in F$ bestaat zó dat f^∞ optimaal is. Indien we bij de constructie van de rij $\{f_n\}_{n=1}^\infty$ in het bewijs van stelling 3 onder iii) voor $f_{n+1}(s)$ niet een willekeurig element uit $G(s,f_n)$ hadden gekozen maar een a die $\{r(s,a) + \beta p(s,a)V_{f_n^\infty}\}$ maximaal maakt dan was de rij geconstrueerd volgens Howard's iteratiemethode (vergelijk Hoofdstuk 1, paragraaf 1). Dus Howard's iteratiemethode leidt in een eindig aantal stappen tot een optimale en stationaire strategie.

We besluiten deze paragraaf met enkele kriteria voor optimaliteit (t.o.v. $V_{\beta,R}$).

Stelling 6

f^∞ is optimaal d.e.s.d. als $G(s,f) = \emptyset$ voor alle s.

Bewijs

Volgens stelling 5 geldt dat f^∞ optimaal is indien $G(s,f) = \emptyset$ voor alle s. Indien $G(s,f) \neq \emptyset$ voor alle s dan is er volgens lemma 6 een $g \in F$ met $V_{g^\infty} > V_{f^\infty}$ en dus is f^∞ dan niet optimaal.

Definitie 3

$$G(s,R) = \{a \in A \mid r(s,a) + \beta p(s,a)V_R > V_R(s)\}.$$

Stelling 7

$R \in C$ is optimaal d.e.s.d. als $G(s,R) = \emptyset$ voor alle s.

Bewijs

Volgens de stellingen 3 en 5 is er een $f_0 \in F$ met f_0^∞ optimaal. Volgens stelling 6 geldt dan dat $G(s,f_0) = \emptyset$ voor alle s.
Indien R optimaal is dan moet gelden dat $V_R(s) = V_{f_0^\infty}(s)$ voor alle s.
Hieruit volgt dat $G(s,R) = G(s,f_0) = \emptyset$ voor alle s.
Is omgekeerd $G(s,R) = \emptyset$ voor alle s dan kan geheel analoog aan lemma 7 bewezen worden dat $L(\pi)V_R \leq V_R$ voor alle $\pi \in \mathbb{P}(A/S)$. In het bijzonder geldt dan dat $L(f_0)V_R \leq V_R$, uit lemma 4 volgt dan dat $V_{f_0^\infty} \leq V_R$ en dus ook R is optimaal.

Stelling 8

$R \in C$ is optimaal d.e.s.d. als $V_R(s) = \max_a \{r(s,a) + \beta p(s,a)V_R\}$ voor alle s.

Bewijs

Zij f_0 als in het bewijs van stelling 7.
Indien R optimaal is dan geldt $V_R(s) = V_{f_0^\infty}(s)$ voor alle s.
Voor $V_{f_0^\infty}$ geldt dat $V_{f_0^\infty}(s) = \max_a\{r(s,a)+\beta p(s,a)V_{f_0^\infty}\}$ voor alle s en dus ook $V_R(s) = \max_a \{r(s,a)+\beta p(s,a)V_R\}$ voor alle s.
Omgekeerd indien $V_R(s) = \max_a \{r(s,a)+\beta p(s,a)V_R\}$ voor alle s dan volgt dat $G(s,R) = \emptyset$ voor alle s en volgens stelling 7 is R dan optimaal.

2.3. Totale verwachte opbrengst

In het eerste gedeelte van deze paragraaf zullen we veronderstellen dat de verwachte opbrengst van de eerst N+1 dagen ($V_R^N(s)$) voor $N \to \infty$ een limiet heeft (eventueel $\pm\infty$) voor alle $R \in C$ en alle $s \in S$. We kunnen dan opbrengstfunktie $V_{1,R}(s)$ (= $\lim_{N\to\infty} V_R^N(s)$) als optimaliteitskriterium hanteren. Onder de aanname dat $V_{1,R}(s)$ steeds bestaat zullen we bewijzen (stelling 1) dat er een $f \in F$ is zó dat f^∞ optimaal (t.o.v. $V_{1,R}$) is, d.w.z. $V_{1,R}(s) \leq V_{1,f^\infty}(s)$ voor alle $R \in C$ en alle $s \in S$.
In het tweede gedeelte zullen we voorwaarden geven waaronder $\lim_{N\to\infty} V_R^N(s)$ bestaat voor alle $R \in C$ en alle $s \in S$ (stellingen 2 en 4).

Lemma 1

Zij $\sum_{n=0}^{\infty} a_n x^n$ een machtreeks met reële coëfficiënten en een convergentiestraal groter dan of gelijk aan 1.
Zij $S_n = \sum_{m=0}^{n} a_m$ en voor $-1 < x < +1$ $S(x) = \sum_{n=0}^{\infty} a_n x^n$.
Indien geldt dat $\lim_{n\to\infty} S_n = S$ (S eventueel $\pm\infty$), dan volgt dat $\lim_{X\uparrow 1} S(x) = S$.

Bewijs

Voor $-1 < x < +1$ zijn de machtreeksen $\sum_{n=0}^{\infty} a_n x^n$ en $\sum_{n=0}^{\infty} x^n$ absoluut convergent. Hieruit volgt dat voor $-1 < x < +1$

$$\frac{S(x)}{1-x} = \sum_{n=0}^{\infty} a_n x^n \cdot \sum_{n=0}^{\infty} x^n = \sum_{n,m=0}^{\infty} a_n x^{n+m} =$$

$$= \sum_{N=0}^{\infty} \sum_{n+m=N} a_n x^{n+m} = \sum_{N=0}^{\infty} \sum_{n=0}^{N} a_n x^N = \sum_{n=0}^{\infty} S_n x^n$$

We onderscheiden in het vervolg van het bewijs drie mogelijkheden voor $\lim_{n\to\infty} S_n$, nl. i) $-\infty < S < +\infty$ ii) $S = +\infty$ iii) $S = -\infty$

i) Door de machtreeksen $\sum_{n=0}^{\infty} S_n x^n$ en $\sum_{n=0}^{\infty} S x^n$ termsgewijs af te trekken vinden we dat voor $-1 < x < +1$ geldt

$$\frac{S(x) - S}{1 - x} = \sum_{n=0}^{\infty} (S_n - S)\, x^n$$

Kies $\varepsilon > 0$. Uit $\lim_{n\to\infty} S_n = S$ volgt dat er een natuurlijk getal N bestaat zó dat $|S_n - S| < \frac{\varepsilon}{2}$ voor $n > N$. Zij $K = \max |S_n - S|$ voor $n = 0,1,\ldots,N$.

Voor $0 < x < 1$ geldt dan dat

$$\frac{|S(x)-S|}{1-x} = \left|\sum_{n=0}^{\infty} (S_n - S)\, x^n\right|$$

$$\leq \sum_{n=0}^{N} |S_n - S|\, x^n + \sum_{n=N+1}^{\infty} |S_n - S|\, x^n$$

$$< K \sum_{n=0}^{N} x^n + \frac{\varepsilon}{2} \sum_{n=0}^{\infty} x^n$$

$$= K\, \frac{1 - x^{N+1}}{1-x} + \frac{\varepsilon}{2}\, \frac{1}{1-x} \quad .$$

Hieruit volgt dat $|S(x)-S| < K(1-x^{N+1}) + \frac{\varepsilon}{2}$ voor $0 < x < 1$.
Zij $C > 0$ zó dat $K(1-x^{N+1}) < \frac{\varepsilon}{2}$ voor $C < x < 1$, dan volgt dat voor $C < x < 1$ geldt dat $|S(x)-S| < \varepsilon$.

ii) Kies $A > 0$. Uit $\lim_{n\to\infty} S_n = +\infty$ volgt dat er een natuurlijk getal N bestaat zó dat $S_n > 4A$ voor $n > N$. Zij $K = \min S_n$ voor $n = 0,1,\ldots,N$.

Voor $0 < x < 1$ geldt dan dat

$$\frac{S(x)}{1-x} = \sum_{n=0}^{\infty} S_n\, x^n$$

$$= \sum_{n=0}^{N} S_n\, x^n + \sum_{n=N+1}^{\infty} S_n\, x^n$$

$$> K \sum_{n=0}^{N} x^n + 4A \sum_{n=N+1}^{\infty} x^n$$

$$= K\, \frac{1 - x^{N+1}}{1-x} + 4A\, \frac{x^{N+1}}{1-x}$$

Hieruit volgt dat $S(x) > K(1-x^{N+1}) + 4A\ x^{N+1}$ voor $0 < x < 1$. Zij $C > 0$ zó dat $K(1-x^{N+1}) > -A$ en $x^{N+1} > \frac{1}{2}$ voor $C < x < 1$. Dan geldt voor $C < x < 1$ dat $S(x) > A$.

iii) laat zich analoog aan ii) bewijzen.

Definitie 1

Voor $f \in F$ zij $B(f) = \{\beta \mid 0 < \beta < 1,\ V_{\beta,f^\infty}(s) \geq V_{\beta,R}(s)$ voor alle $R \in C$ en alle $s \in S\}$

Lemma 2

Er is een $f \in F$ zó dat $\beta = 1$ verdichtingspunt is van $B(f)$.

Bewijs

Volgens de stellingen 2.2.3 en 2.2.5 is er voor iedere $0 < \beta < 1$ een $f \in F$ zó dat $\beta \in B(f)$. Hieruit volgt dat $\bigcup_{f \in F} B(f) = (0,1)$ en dus is $\beta = 1$ verdichtingspunt van $\bigcup_{f \in F} B(f)$. Omdat F eindig veel elementen bevat is er een f zó dat $\beta = 1$ verdichtingspunt is van $B(f)$.

Stelling 1

Indien $\lim_{N\to\infty} V_R^N(s)$ bestaat (eventueel $\pm\infty$) voor alle $R \in C$ en alle $s \in S$ en voor $f \in F$ geldt dat $\beta = 1$ verdichtingspunt is van $B(f)$ dan volgt dat f^∞ optimaal is t.o.v. $V_{1,R}$.

Bewijs

We moeten bewijzen dat $V_{1,R}(s) \leq V_{1,f^\infty}(s)$ voor alle $R \in C$ en alle $s \in S$. We zullen dit doen door bovenstaande ongelijkheid voor willekeurig gekozen R en s_0 te bewijzen

Zij $\quad a_n = \sum_{s,a} P_R(x_n=s, y_n=a \mid x_0=s_0)\ r(s,a) \qquad$ en

$$b_n = \sum_{s,a} P_{f^\infty}(x_n=s, y_n=a \mid x_0=s_0)\ r(s,a)\ .$$

Dan geldt dat $V_{\beta,R}(s_0) = \sum_{n=0}^{\infty} a_n \beta^n$ en $V_{\beta,f^\infty} = \sum_{n=0}^{\infty} b_n \beta^n$.
Volgens lemma 1 geldt dat $\lim_{\beta\uparrow 1} \sum_{n=0}^{\infty} a_n \beta^n = V_{1,R}(s_0)$ en
$\lim_{\beta\uparrow 1} \sum_{n=0}^{\infty} b_n \beta^n = V_{1,f^\infty}(s_0)$.
Omdat $\beta = 1$ verdichtingspunt van $B(f)$ is, bestaat er een rij $\{\beta_k\}_{k=1}^{\infty} \subset B(f)$ zó dat $\beta_k \to 1$ voor $k \to \infty$.
Volgens de definitie van $B(f)$ geldt dat $\sum_{n=0}^{\infty} a_n \beta_k^n \leq \sum_{n=0}^{\infty} b_n \beta_k^n$ voor $k = 1,2,\ldots$. Waaruit volgt dat

$$V_{1,R}(s_0) = \lim_{\beta\uparrow 1} \sum_{n=0}^{\infty} a_n \beta^n = \lim_{k\to\infty} \sum_{n=0}^{\infty} a_n \beta_k^n$$

$$\leq \lim_{k\to\infty} \sum_{n=0}^{\infty} b_n \beta_k^n = \lim_{\beta\uparrow 1} \sum_{n=0}^{\infty} b_n \beta^n = V_{1,f^\infty}(s_0) .$$

Dat $\lim_{N\to\infty} V_R(s)$ niet altijd bestaat toont het volgende voorbeeld aan.
Zij $S = \{s_1,s_2\}$, $A = \{a\}$, $q(s_2|s_1,a) = 1$, $q(s_1|s_2,a) = 1$, $r(s_1,a) = +1$ en $r(s_2,a) = -1$. Omdat A slechts één element bevat is er slechts één strategie en deze strategie is stationair. Voor deze strategie R geldt: $V_R^N(s_1) = V_R^N(s_2) = 0$ voor N is oneven en $V_R^N(s_1) = +1$, $V_R^N(s_2) = -1$ voor N even.

Stelling 2

Indien $r(s,a) \geq 0$ voor alle $s \in S$ en alle $a \in A$ of $r(s,a) \leq 0$ voor alle $s \in S$ en alle $a \in A$ dan bestaat $\lim_{N\to\infty} V_R^N(s)$ (eventueel $\pm\infty$) voor alle $R \in C$ en alle $s \in S$.

Bewijs

Indien $r(s,a) \geq 0$ voor alle $s \in S$ en alle $a \in A$ dan volgt dat $\sum_{s,a} P_R(x_n=s,y_n=a|x_0=s_0)\ r(s,a) \geq 0$ voor alle $R \in C$, alle $s_0 \in S$ en voor $n = 0,1,2,\ldots$. Hieruit volgt dat $\{V_R^N(s_0)\}_{N=0}^{\infty}$ een monotoon niet dalende rij is voor alle $R \in C$ en alle $s_0 \in S$. Dus $\lim_{N\to\infty} V_R^N(s_0)$ bestaat voor alle $R \in C$ en $s_0 \in S$.

Indien $r(s,a) \leq 0$ voor alle $s \in S$ en alle $a \in A$ dan is $\{V_R^N(s_0)\}_{N=0}^{\infty}$ een monotoon niet stijgende rij voor alle $R \in C$ en alle $s_0 \in S$.

Definitie 2

Zij de verzameling van absorberende toestanden

$$S_{ab} = \{s \mid q(s|s,a) = 1 \text{ voor alle } a \in A\}$$

en zij de complementaire verzameling $S_c = S \setminus S_{ab}$.

Definitie 3

Zij $P^n(R)$ de matrix met (s_0,s)-de element gelijk aan $P_R(x_n=s|x_0=s_0)$. Zij $\widetilde{P^n(R)}$ de deelmatrix van $P^n(R)$ die verkregen wordt uit $P^n(R)$ door iedere rij en kolom met een index $s \in S_{ab}$ te schrappen.
Voor $\pi \in \mathbb{P}(A|S)$ zij $\widetilde{P(\pi)}$ de overeenkomstige deelmatrix van $P(\pi)$.

Lemma 3

Voor $\pi_1,\pi_2 \in \mathbb{P}(A|S)$ geldt dat $\widetilde{P(\pi_1)P(\pi_2)} = \widetilde{P(\pi_1)}.\widetilde{P(\pi_2)}$.

Bewijs

Voor $s_0,s_1 \in S_c$ geldt

$$(\widetilde{P(\pi_1)\ P(\pi_2)})_{s_0,s_1} = (P(\pi_1)\ P(\pi_2))_{s_0,s_1}$$

$$= \sum_{s' \in S} \sum_{a} \pi_1(a|s_0)q(s'|s_0,a) \sum_{a'} \pi_2(a'|s')q(s_1|s',a')$$

indien $s' \in S_{ab}$ dan $q(s_1|s',a') = 0$ voor alle a'

$$= \sum_{s' \in S_c} (P(\pi_1))_{s_0,s'}\ (P(\pi_2))_{s',s_1}$$

$$= (\widetilde{P(\pi_1)}.\widetilde{P(\pi_2)})_{s_0,s_1}$$

Definitie 4

Een strategie R waarvoor $\sum_{n=0}^{\infty} P_R(x_n=s|x_0=s_0) < \infty$ voor alle $s_0, s \in S_c$ noemen we een doorgangsstrategie.

Stelling 3

f^{∞} doorgangsstrategie $\Longleftrightarrow$ alle toestanden in S_c zijn doorgangstoestanden voor de Markovketen met overgangsmatrix $P(f)$.

Bewijs

Uit stelling 1.2.8 volgt

$$\sum_{n=0}^{\infty} P_{f^{\infty}}(x_n=s|x_0=s_0) < \infty \text{ voor alle } s_0 \in S \Longleftrightarrow \sum_{n=0}^{\infty} P_{f^{\infty}}(x_n=s|x_0=s) < \infty.$$

Volgens stelling 1.2.11 geldt

$$\sum_{n=0}^{\infty} P_{f^{\infty}}(x_n=s|x_0=s) < \infty \Longleftrightarrow s \text{ is doorgangstoestand voor}$$

de Markovketen met overgangsmatrix $P(f)$.
Waaruit volgt dat

$$f^{\infty} \text{ doorgangsstrategie} \Longleftrightarrow \sum_{n=0}^{\infty} P_{f^{\infty}}(x_n=s|x_0=s_0) < \infty \text{ voor alle } s_0, s \in S_c$$

$$\Longleftrightarrow \sum_{n=0}^{\infty} P_{f^{\infty}}(x_n=s|x_0=s_0) < \infty \text{ voor alle } s \in S_c \text{ en } s_0 \in S$$

$$\Longleftrightarrow \sum_{n=0}^{\infty} P_{f^{\infty}}(x_n=s|x_0=s) < \infty \text{ voor alle } s \in S_c$$

$\Longleftrightarrow$ alle toestanden in S_c zijn doorgangstoestanden

van de Markovketen met overgangsmatrix $P(f)$.

Lemma 4

Voor ieder natuurlijk getal n is er een $R_n \in C_m$ zó dat

$$\sum_{s \in S_c} P_R(x_n=s|x_0=s_0) \leq \sum_{s \in S_c} P_{R_n}(x_n=s|x_0=s_0) \text{ voor alle } R \in C \text{ en alle } s_0 \in S.$$

Bewijs

Met volledige inductie.

Voor n = 0 geldt $\sum_{s \in S_c} P_R(x_0=s|x_0=s_0) = 1$ resp. 0 voor alle $R \in C$ en alle $s_0 \in S_c$ resp. $s_0 \in S_{ab}$ en is de bewering dus juist.

Neem aan $\sum_{s \in S_c} P_R(x_m=s|x_0=s_0) \leq \sum_{s \in S_c} P_{R_m}(x_m=s|x_0=s_0)$ voor alle $R \in C$ en alle $s_0 \in S$.

Zij dan $f_0 \in F$ zó dat voor alle $s_0 \in S$

$$\sum_{s'} q(s'|s_0,f_0(s_0)) \sum_{s \in S_c} P_{R_m}(x_m=s|x_0=s') =$$

$$= \max_a \sum_{s'} q(s'|s_0,a) \sum_{s \in S_c} P_{R_m}(x_m=s|x_0=s')$$

Volgens de inductieaanname is R_m een Markovstrategie en dus (f_0,R_m) is ook een Markovstrategie.

Zij R willekeurig element uit C en s_0 willekeurig element uit S. Dan volgt uit stelling 2.2.1 dat er $R^* \in C_{rm}$ bestaat zó dat

$$\sum_{s \in S_c} P_R(x_{m+1}=s|x_0=s_0) = \sum_{s \in S_c} P_{R^*}(x_{m+1}=s|x_0=s_0) .$$

Indien $R^* = (\pi_0,\pi_1,\ldots)$ dan noteren we voor $(\pi_1,\pi_2,\ldots)$ R^{**}

$$\sum_{s \in S_c} P_{R^*}(x_{m+1}=s|x_0=s_0) = \sum_{s'} (P(\pi_0))_{s_0,s'} \sum_{s \in S_c} P_{R^{**}}(x_m=s|x_0=s')$$

$$\leq \sum_{s'} (P(\pi_0))_{s_0,s'} \sum_{s \in S_c} P_{R_m}(x_m=s|x_0=s')$$

volgens de inductieaanname

$$= \sum_{s'} \sum_{a} \pi_0(a|s_0) q(s'|s_0,a) \sum_{s \in S_c} P_{R_m}(x_m=s|x_0=s')$$

$$\leq \sum_{a} \pi_0(a|s_0) \sum_{s'} q(s'|s_0,f_0(s_0)) \sum_{s \in S_c} P_{R_m}(x_m=s|x_0=s')$$

$$= \sum_{s \in S_c} P_{(f_0,R_m)}(x_{m+1}=s|x_0=s_0)$$

Uit de definitie (1.3.1) van de norm van een matrix met niet-negatieve elementen volgt dan m.b.v. lemma 4.

Lemma 5

Voor ieder natuurlijk getal n is er een $R_n \in C_m$ zó dat

$$\sup_{R \in C} ||\widetilde{P^n(R)}|| \leq ||\widetilde{P^n(R_n)}||$$

Lemma 6

De volgende beweringen zijn equivalent.

i) Iedere stationaire strategie is een doorgangsstrategie

ii) Iedere strategie is een doorgangsstrategie

iii) Er is een m waarvan $||P^m(R)|| < 1$ voor iedere strategie $R \in C$

iv) $||P^N(R)|| < 1$ voor iedere strategie $R \in C$, als N het aantal elementen in S_c aangeeft.

Bewijs

iv) $\Longrightarrow$ iii) triviaal

iii) $\Longrightarrow$ ii) neem aan dat iii) geldt voor m gelijk aan n. Passen we lemma 5 toe dan vinden we

$$\sup_{R \in C} ||\widetilde{P^n(R)}|| \leq ||\widetilde{P^n(R_n)}|| < 1 \quad \text{voor een } R_n \in C_m$$

Stel $a = ||\widetilde{P^n(R_n)}||$ dan is $0 \leq a < 1$.

We moeten bewijzen dat $\sum\limits_{n=0}^{\infty} P_R(x_n=s|x_0=s_0) < \infty$ voor alle $R \in C$ en alle $s_0, s \in S_c$. Kies $s_0, s \in S_c$ en R willekeurig dan is er volgens stelling 2.2.1 een $R_0 \in C_{rm}$ zó dat

$$P_R(x_n=s|x_0=s_0) = P_{R_0}(x_n=s|x_0=s_0) \quad \text{voor } n = 0,1,2,\ldots$$

Daar de reeks $\sum\limits_{n=0}^{\infty} P_{R_0}(x_n=s|x_0=s_0)$ gemajoreerd wordt door $\sum\limits_{n=0}^{\infty} ||\widetilde{P^n(R_0)}||$ kunnen we volstaan met te bewijzen dat $\sum\limits_{n=0}^{\infty} ||\widetilde{P^n(R_0)}|| < \infty$.

Indien $R_0 = (\pi_0, \pi_1, \ldots)$ definieer dan $R_k = (\pi_{kn}, \pi_{kn+1}, \ldots)$ voor $k = 1,2,\ldots$.

Met behulp van lemma 1 volgt dat

$$||\widetilde{P^{kn+l}(R_0)}|| = ||\widetilde{P^n(R_0)}\ \widetilde{P^n(R_1)} \ldots \widetilde{P^n(R_{k-1})}\ \widetilde{P^l(R_k)}|| \leq a^k ||\widetilde{P^l(R_k)}||$$

Hieruit volgt dat

$$\sum_{l=0}^{\infty} ||\widetilde{P^l(R_0)}|| = \sum_{l=0}^{n-1} ||\widetilde{P^l(R_0)}|| + \sum_{l=0}^{n-1} ||\widetilde{P^{n+l}(R_0)}|| + \sum_{l=0}^{n-1} ||\widetilde{P^{2n+l}(R_0)}|| + \ldots$$

$$\leq \sum_{l=0}^{n-1} ||\widetilde{P^l(R_0)}|| + a \sum_{l=0}^{n-1} ||\widetilde{P^l(R_1)}|| + a^2 \sum_{l=0}^{n-1} ||\widetilde{P^l(R_2)}|| + \ldots$$

$$\leq n(1 + a + a^2 + \ldots) = \frac{n}{1-a} < \infty$$

ii) $\Longrightarrow$ i) triviaal

i) $\Longrightarrow$ iv) we moeten bewijzen dat $||\widetilde{P^N(R)}|| < 1$ voor alle $R \in C$. Dit is equivalent met $\sum\limits_{s \in S_c} P_R(x_N=s|x_0=s_0) < 1$ voor alle $s_0 \in S_c$ en alle $R \in C$.

Kies $s_0 \in S_c$ en $R \in C$ willekeurig dan is er volgens stelling 2.2.1 een $R_0 \in C_{rm}$, zij $R_0 = (\pi_0, \pi_1, \ldots)$, zó dat $\sum\limits_{s \in S_c} P_{R_0}(x_N=s|x_0=s_0) =$
$= \sum\limits_{s \in S_c} P_R(x_N=s|x_0=s_0)$. We zullen bewijzen dat

$$\sum_{s \in S_c} P_{R_0}(x_N=s|x_0=s_0) < 1. \tag{2.3.1}$$

We voeren de toestand $s_\oplus$ in en definiëren $P_R(x_n=s_\oplus|x_0=s_0) = 1 - \sum_{s\in S_c} P_R(x_n=s|x_0=s_0)$ voor alle $s_0 \in S_c$ en $q(s_\oplus|s_\oplus,a) = 1$ voor alle $a\in A$.

In het vervolg van het bewijs beperken we ons tot de verzameling toestanden $S_c \cup \{s_\oplus\}$.

We definiëren $T_n = \{s \mid P_{R_0}(x_n=s|x_0=s_0) > 0\}$ voor $n = 0,1,2,\ldots$.

Daar $s_\oplus$ een absorberende toestand is volgt dat

$$s_\oplus \in T_n \quad \text{met } n \leq N \Longrightarrow s_\oplus \in T_N .$$

Hieruit volgt dat (2.3.1) geldt indien $s_\oplus \in \bigcup_{i=0}^{N} T_i$.

We bewijzen eerst dat uit i) volgt dat

$$(2.3.2) \qquad s_\oplus \notin \bigcup_{i=0}^{n} T_i \Longrightarrow T_{n+1} \not\subset \bigcup_{i=0}^{n} T_i \quad \text{voor } n = 0,1,2,\ldots .$$

Veronderstellen we dat (2.3.2) niet juist is dan is er een n waarvoor geldt

$$(2.3.3) \qquad s_\oplus \notin \bigcup_{i=0}^{n} T_i \quad \text{en} \quad T_{n+1} \subset \bigcup_{i=0}^{n} T_i .$$

Zij $f\in F$ waarvoor geldt dat

$$\pi_j(f(s)|s) > 0 \quad \text{als } s \in T_j \setminus \bigcup_{i=0}^{j-1} T_i .$$

Definieer $T_k^* = \{s \mid P_{f^\infty}(x_k=s|x_0=s_0) > 0\}$ voor $k = 0,1,2,\ldots$ dan is met behulp van volledige inductie te bewijzen dat

$$(2.3.4) \qquad T_k^* \subset \bigcup_{i=0}^{k} T_i \quad k = 0,1,2,\ldots .$$

Uit (2.3.4) en $T_{n+1} \subset \bigcup_{i=0}^{n} T_i$ is weer via volledige inductie te bewijzen dat

$$T_k^* \subset \bigcup_{i=0}^{n} T_i \quad \text{voor } k = 0,1,2,\ldots .$$

We vinden dat dat uit (2.3.3) volgt dat voor alle k

$$\left.\begin{array}{l} T_k^{*} \subset \bigcup\limits_{i=0}^{n} T_i \\ \text{en} \\ s_{\oplus} \notin \bigcup\limits_{i=0}^{n} T_i \end{array}\right\} \Longrightarrow \text{vanuit } s_0 \text{ is } s_{\oplus} \text{ onder strategie } f^{\infty} \text{ niet bereikbaar}$$

$$\Longrightarrow \sum_{s \in S_c} P_{f^{\infty}}(x_m = s \mid x_0 = s_0) = 1 \quad \text{voor } m = 0,1,2,\ldots$$

$$\Longrightarrow \|(\widetilde{P(f)})^m\| \qquad \text{voor } m = 0,1,2,\ldots$$

$$\Longrightarrow \sum_{m=0} (\widetilde{P(f)})^m \quad \text{divergeert (vergelijk stelling 1.3.2)}$$

$\Longrightarrow$ f^{∞} is geen doorgangsstrategie. Dit is in tegenspraak met i) dus (2.3.2) is juist.

Daar S_c precies N elementen bevat volgt uit (2.3.2) dat $s_{\oplus} \in \bigcup\limits_{i=0}^{N} T_i$.

Stelling 4

Indien voor iedere $f \in F$ geldt dat alle elementen uit S_c doorgangstoestanden zijn voor de Markovketen met overgangsmatrix P(f) en als bovendien geldt dat $r(s,a) = 0$ voor alle $s \in S_{ab}$ en alle $a \in A$, dan bestaat $\lim\limits_{N\to\infty} V_R^N(s)$ voor alle $R \in C$ en alle $s \in S$ en is bovendien eindig.

Bewijs

Met stelling 3 volgt uit het gegeven dat alle stationaire strategieën doorgangsstrategieën zijn.

Indien $s_0 \in S_{ab}$ dan volgt dat $V_R^N(s_0) = 0$ voor $N = 0,1,2,\ldots$ en alle $R \in C$.

Voor $s_0 \in S_c$ en $R \in C$ is er weer volgens stelling 2.2.1 een $R_0 \in C_{rm}$ zeg $R_0 = (\pi_0, \pi_1, \ldots)$ zó dat

$$V_R^N(s_0) = V_{R_0}^N(s_0) \qquad \text{voor } N = 0,1,2,\ldots\ .$$

Zij $M = \max_{s,a} |r(s,a)|$ dan volgt

$$\left|\sum_{s,a} P_{R_0}(x_n=s,y_n=a|x_0=s_0)r(s,a)\right| \leq \sum_{s\in S_c} P_{R_0}(x_n=s|x_0=s_0) \sum_a \pi_n(a|s)\ |r(s,a)|$$

$$\leq M \sum_{s\in S_c} P_{R_0}(x_n=s|x_0=s_0)$$

$$\leq M\ ||\widetilde{P^n(R_0)}||$$

Uit lemma 6 volgt (zie bewijs van iii) $\Longrightarrow$ ii)) dat $\sum_{n=0}^{\infty} ||\widetilde{P^n(R_0)}|| < \infty$.
Dus $\sum_{n=0}^{\infty} \sum_{s,a} P_{R_0}(x_n=s,y_n=a|x_0=s_0)r(s,a)$ wordt absoluut gemajoreerd door een convergente reeks.

2.4. Gemiddelde verwachte opbrengst

In deze paragraaf worden de optimaliteitskriteria $g_R(s)$ $(= \liminf_{N\to\infty} \frac{1}{N+1} V_R^N(s)$, $V_R^N(s)$ is de verwachte opbrengst van de eerste $n+1$ dagen) en $G_R(s)$ $(= \limsup_{N\to\infty} \frac{1}{N+1} V_R^N(s))$ beschouwd.
We zullen bewijzen dat er een stationaire strategie bestaat die optimaal is t.o.v. G_R in de klasse van alle strategieën (stelling 6). Omdat voor iedere gerandomiseerde stationaire strategie R geldt dat $g_R(s) = G_R(s)$ voor alle s (stelling 1) volgt dan dat een stationaire strategie die optimaal is t.o.v. G_R tevens optimaal is t.o.v. g_R. Het existentiebewijs voor een optimale stationaire strategie zal evenals in paragraaf 2 weer algorithmisch zijn d.w.z. uit het existentiebewijs is een algorithme te construeren dat in een eindig aantal stappen een optimale strategie bepaalt (zie opmerking 2).

Lemma 1

Zij $\sum_{n=0}^{\infty} a_n x^n$ een machtreeks met reële coëfficiënten.
Zij $S_n = \sum_{k=0}^{n} a_k$ en $S(x) = \sum_{n=0}^{\infty} a_n x^n$.
Indien $\lim_{n\to\infty} \frac{S_n}{n+1} = S$ $(-\infty<S<\infty)$ dan geldt dat $\sum_{n=0}^{\infty} a_n x^n$ een convergentiestraal heeft die groter dan of gelijk aan 1 is en $\lim_{x\uparrow 1} (1-x)\, S(x) = S$.

Bewijs
De machtreeks $\sum_{n=0}^{\infty} a_n x^n$ heeft een convergentiestraal gelijk aan $R = (\limsup_{n\to\infty} \sqrt[n]{|a_n|})^{-1}$.
Daar $\lim_{n\to\infty} \frac{S_n}{n+1}$ bestaat en eindig is volgt dat de rij $\{\frac{|S_n|}{n+1}\}$ begrensd is, zeg $\frac{|S_n|}{n+1} \leq K$ (met $0<K<\infty$) voor alle n. Hieruit volgt dat

$|a_n| = |S_n - S_{n-1}| \leq (n+1)\,K + nK = (2n+1)K$ en dus $\limsup_{n\to\infty} \sqrt[n]{|a_n|} \leq \lim_{n\to\infty} \sqrt[n]{(2n+1)K} = 1$. Waaruit volgt dat R groter dan of gelijk aan 1 is.

Dus voor $-1 < x < +1$ geldt (vergelijk lemma 2.3.1) $\frac{S(x)}{1-x} = \sum_{n=0}^{\infty} S_n x^n$.

Door de machtreeks $\frac{1}{1-x} = \sum_{n=0}^{\infty} Sx^n$ termsgewijs te differentiëren vinden we dat $\frac{S}{(1-x)^2} = \sum_{n=0}^{\infty} (n+1)\, Sx^n$.

Trekken we de twee machtreeksen van elkaar af dan vinden we dat

$$\frac{S(x)}{1-x} - \frac{S}{(1-x)^2} = \sum_{n=0}^{\infty} (S_n-(n+1)S)x^n.$$

Kies $\varepsilon > 0$. Uit $\lim_{n\to\infty} \frac{S_n}{n+1} = S$ volgt dat er een natuurlijk getal N bestaat zó dat $\left|\frac{S_n}{n+1} - S\right| < \frac{\varepsilon}{2}$ voor $n > N$. Zij $K = \max |S_n-(n+1)S|$ voor $n = 0, 1, \ldots, N$.

Voor $0 < x < 1$ geldt dan dat

$$\left|\frac{(1-x)S(x)-S}{(1-x)^2}\right| \leq \sum_{n=0}^{N} |S_n-(n+1)S|x^n + \sum_{n=N+1}^{\infty} |S_n-(n+1)S|x^n$$

$$< K \sum_{n=0}^{N} x^n + \frac{\varepsilon}{2} \sum_{n=0}^{\infty} (n+1)x^n$$

$$= K\,\frac{1-x^{N+1}}{1-x} + \frac{\varepsilon}{2}\,\frac{1}{(1-x)^2}\,.$$

Hieruit volgt dat $|(1-x)S(x)-S| \leq K(1-x)(1-x^{N+1}) + \frac{\varepsilon}{2}$.
Zij $c > 0$ zó dat $K(1-x)(1-x^{N+1}) < \frac{\varepsilon}{2}$ voor $c < x < 1$, dan volgt dat voor $c < x < 1$ geldt dat $|(1-x)S(x)-S| < \varepsilon$.

Lemma 2

Indien voor de matrix P met reële elementen geldt dat

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} P^k = \text{nulmatrix} \overset{\text{not.}}{=} 0$$

dan volgt dat $(I-P)^{-1}$ bestaat, bovendien geldt dat

$$\lim_{x\uparrow 1} \sum_{n=0}^{\infty} P^n x^n = (I-P)^{-1}.$$

Bewijs

Zij $S(x) = \sum_{n=0}^{\infty} P^n x^n$. Dan volgt uit lemma 1 dat de machtreeks een convergentiestraal heeft groter dan of gelijk aan 1. Hieruit volgt dat voor $-1 < x < +1$ geldt dat $\lim_{n\to\infty} P^n x^n = 0$ en dus

$$(I-xP)\ S(x) = \sum_{n=0}^{\infty} P^n x^n - \sum_{n=0}^{\infty} P^{n+1} x^{n+1} = \lim_{n\to\infty} (I-P^n x^n) = I.$$

Dus geldt voor $-1 < x < +1$ dat

$$(2.4.1) \qquad (I-P)\ S(x) = I - (1-x)\ PS(x).$$

Uit lemma 1 volgt dat $\lim_{x\uparrow 1} (1-x)\ S(x) = 0 \Rightarrow \lim_{x\uparrow 1} (1-x)\ PS(x) = 0$.

Veronderstellen we dat determinant $(I-P) = 0$ dan volgt uit (2.4.1) dat $(I-(1-x)PS(x)) = 0$ voor $-1 < x < +1$. Dit leidt tot een tegenspraak immers dan volgt dat $0 = \lim_{x\uparrow 1} \det(I-(1-x)PS(x)) = \det(I-\lim_{x\uparrow 1}(1-x)PS(x)) = \det(I) = 1$. De veronderstelling dat $\det(I-P) = 0$ is dus onjuist oftewel $\det(I-P) \neq 0$ waaruit volgt dat $(I-P)^{-1}$ bestaat.

We zagen reeds dat $(I-xP)\ S(x) = I$ dus $S(x) = (I-xP)^{-1}$. Uit $\det(I-P) \neq 0$ volgt dan dat $\lim_{x\uparrow 1}(I-xP) = (I-P)^{-1}$ en dus ook $\lim_{x\uparrow 1} S(x) = (I-P)^{-1}$.

Notatie 1

Voor $\pi \in \mathbb{P}(A|S)$ zij $P^*(\pi) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} P^k(\pi)$.

Volgens stelling 1.2.12 bestaat deze limiet en volgens 1.2.13 geldt dat

$$P(\pi)\ P^*(\pi) = P^*(\pi)\ P(\pi) = P^*(\pi)\ P^*(\pi) = P^*(\pi).$$

Lemma 3

Voor $\pi \in \mathbb{P}(A|S)$ geldt dat $(P(\pi)-P^*(\pi))^n = P^n(\pi) - P^*(\pi)$ voor $n = 1, 2, \ldots$

Bewijs

Met volledige inductie:

Neem aan dat $(P(\pi)-P^*(\pi))^k = P^k(\pi) - P^*(\pi)$ dan volgt dat

$$(P(\pi)-P^*(\pi))^{k+1} = (P(\pi)-P^*(\pi))\ (P^k(\pi)-P^*(\pi)) = P^{k+1}(\pi) - P^*(\pi).$$

Lemma 4

Voor $\pi \in \mathbb{P}(A|S)$ geldt dat $(I-P(\pi)+P^*(\pi))^{-1}$ bestaat.

Zij $H(\pi,x) = \sum_{n=0}^{\infty} (P^n(x)-P^*(\pi))x^n$ deze reeks heeft een convergentiestraal groter dan of gelijk aan 1.

Zij $H(\pi) = (I-P(\pi)+P^*(\pi))^{-1} - P^*(\pi)$.

De volgende relaties gelden

i) $\lim_{x\uparrow 1} H(\pi,x) = H(\pi)$

ii) $P^*(\pi)\ H(\pi,x) = H(\pi,x)\ P^*(\pi) = P^*(\pi)\ H(\pi) = H(\pi)\ P^*(x) = 0$

iii) $(I-P(\pi))\ H(\pi) = H(\pi)\ (I-P(\pi)) = I - P^*(\pi)$.

Bewijs

i) Volgens lemma 3 geldt dat

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} (P(\pi)-P^*(\pi))^k = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} (P^k(\pi)-P^*(\pi)) = 0$$

en

$$H(\pi,x) + P^*(\pi) = \sum_{n=0}^{\infty} (P^n(\pi)-P^*(\pi))x^n + P^*(\pi) = \sum_{n=0}^{\infty} (P(\pi)-P^*(\pi))^n x^n.$$

Uit lemma 2 volgt dan dat

$$\lim_{x\uparrow 1} (H(\pi,x)+P^*(\pi)) = (I-P(\pi)+P^*(\pi))^{-1}.$$

ii) $$P^*(\pi)\ H(\pi,x) = P^*(\pi) \sum_{n=0}^{\infty} (P^n(\pi)-P^*(\pi))x^n = \sum_{n=0}^{\infty} (P^*(\pi)-P^*(\pi))x^n = 0$$

en

$$P^*(\pi)\ H(\pi) = P^*(\pi) \lim_{x\uparrow 1} H(\pi,x) = \lim_{x\uparrow 1} P^*(\pi)\ H(\pi,x) = 0.$$

iii) $$\begin{aligned}(I-P(\pi))\ H(\pi) &= (I-P(\pi)) \lim_{x\uparrow 1} H(\pi,x)\\ &= \lim_{x\uparrow 1} \sum_{n=0}^{\infty} (I-P(\pi))\ (P^n(\pi)-P^*(\pi))x^n\\ &= \lim_{x\uparrow 1} \sum_{n=0}^{\infty} (P^n(\pi)-P^{n+1}(\pi))x^n\\ &= \lim_{x\uparrow 1} (I + \sum_{n=1}^{\infty} P^n(\pi)x^n - \sum_{n=0}^{\infty} P^{n+1}(\pi)x^n)\\ &= \lim_{x\uparrow 1} (I - (1-x) \sum_{n=0}^{\infty} P^{n+1}(\pi)x^n).\end{aligned}$$

Daar $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} P^{n+1}(\pi) = P^*(\pi)$ volgt uit lemma 1

$$= I - P^*(\pi).$$

<u>Notatie 2</u>

$V_{\beta,R}$ respectievelijk V_R^N zijn de vectoren met s-de component $V_{\beta,R}(s)$ respectievelijk $V_R^N(s)$. Voor $\pi \in \mathbb{P}(A|S)$ zij $x(\pi) = P^*(\pi)\ r(\pi)$ en $y(\pi) = H(\pi)\ r(\pi)$.

<u>Stelling 1</u>

Voor $\pi \in \mathbb{P}(A|S)$ geldt dat

$$\liminf_{N\to\infty} \frac{1}{n+1} V^N_{\pi\ \infty} = \limsup_{N\to\infty} \frac{1}{N+1} V^N_{\pi\ \infty} = x(\pi).$$

<u>Bewijs</u>

$$\lim_{N\to\infty} \frac{1}{N+1} V^N_{\pi\ \infty} = \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} P^n(\pi)r(\pi) = P^*(\pi)r(\pi) = x(\pi).$$

Stelling 2

Voor $0 < \beta < 1$ en $\pi \in \mathbb{P}(A|S)$ geldt dat

$$(2.4.2) \qquad V_{\beta,\pi^\infty} = \frac{x(\pi)}{1-\beta} + y(\pi) + \varepsilon(\pi,\beta) \quad \text{met } \lim_{\beta\uparrow 1} \varepsilon(\pi,\beta) = 0.$$

Bewijs

$$V_{\beta,\pi^\infty} = \sum_{n=0}^{\infty} \beta^n P^n(\pi) r(\pi)$$

$$= \sum_{n=0}^{\infty} \beta^n (P^n(\pi)-P^*(\pi))r(\pi) + \sum_{n=0}^{\infty} \beta^n P^*(\pi)r(\pi)$$

$$= H(\pi,\beta)r(\pi) + \frac{P^*(\pi)r(\pi)}{1-\beta}$$

$$= \frac{x(\pi)}{1-\beta} + y(\pi) + \underbrace{(H(\pi,\beta)r(\pi)-H(\pi)r(\pi))}_{\varepsilon(\pi,\beta)}$$

Volgens lemma 4 deel i) geldt nu dat $\lim_{\beta\uparrow 1} \varepsilon(\pi,\beta) = 0$.

Stelling 3

Voor $0 < \beta < 1$ en $\pi_1, \pi_2 \in \mathbb{P}(A|S)$ geldt dat

$$(2.4.3) \quad V_{\beta,(\pi_1,\pi_2^\infty)} = \frac{P(\pi_1)x(\pi_2)}{1-\beta} + r(\pi_1) + P(\pi_1)(y(\pi_2)-x(\pi_2)) + \varepsilon(\pi_1,\pi_2,\beta)$$

$$\text{met } \lim_{\beta\uparrow 1} \varepsilon(\pi_1,\pi_2,\beta) = 0.$$

Bewijs

Volgens lemma 2.2.2 deel iii) geldt dat

$$V_{\beta,(\pi_1,\pi_2^\infty)} = r(\pi_1) + \beta P(\pi_1)\, V_{\beta,\pi_2^\infty}.$$

Substitueren we hierin (2.4.2) dan vinden we

$$= r(\pi_1) + \beta P(\pi_1)\,\frac{x(\pi_2)}{1-\beta} + \beta P(\pi_1)y(\pi_2) + \beta P(\pi_1)\varepsilon(\pi_2,\beta)$$

$$= \frac{P(\pi_1)x(\pi_2)}{1-\beta} - (1-\beta)\,\frac{P(\pi_1)x(\pi_2)}{1-\beta} + r(\pi_1) + P(\pi_1)y(\pi_2) -$$

$$- (1-\beta)P(\pi_1)y(\pi_2) + \beta P(\pi_1)\varepsilon(\pi_2,\beta)$$

$$= \frac{P(\pi_1)x(\pi_2)}{1-\beta} + r(\pi_1) + P(\pi_1)(y(\pi_2)-x(\pi_2)) +$$

$$+ \underbrace{\beta P(\pi_1)\varepsilon(\pi_2,\beta)-(1-\beta)P(\pi_1)y(\pi_2)}_{\varepsilon(\pi_1,\pi_2,\beta)}$$

Het is duidelijk dat dan $\lim\limits_{\beta\uparrow 1} \varepsilon(\pi_1,\pi_2,\beta) = 0$.

Definitie 1

Voor $s \in S$ en $f \in F$ zij $A(s,f)$ die deelverzameling van A waarvoor geldt dat

$$\text{òf} \quad p(s,a)x(f) > x_s(f)$$

$$\text{òf} \begin{cases} p(s,a)x(f) = x_s(f) \text{ én} \\ r(s,a) + p(s,a)y(f) > x_s(f) + y_s(f). \end{cases}$$

Opmerking 1

Onder de uitdrukking "bewering B geldt voor β groot genoeg" verstaan we er is een c met $0 < c < 1$ zodat bewering B geldt voor $c < \beta < 1$.

Lemma 5

Voor f, $g \in F$ geldt dat

i) $g(s) = f(s) \Longrightarrow V_{\beta,(g,f^\infty)}(s) = V_{\beta,f^\infty}(s)$ voor alle $0 < \beta < 1$

ii) $g(s) \in A(s,f) \Longrightarrow V_{\beta,(g,f^\infty)}(s) > V_{\beta,f^\infty}(s)$ voor β groot genoeg.

Bewijs

i) is reeds bewezen voor willekeurige $0 < \beta < 1$ in lemma 2.2.5 deel i).

ii) volgens (2.4.2) en (2.4.3) geldt dat

$$V_{\beta,(g,f^\infty)}(s)-V_{\beta,f^\infty}(s) = \underbrace{\frac{p(s,g(s))x(f)-x_s(f)}{1-\beta}}_{\frac{t_1}{1-\beta}} +$$

$$+ \underbrace{r(s,g(s))+p(s,g(s))(y(f)-x(f))-y_s(f)}_{t_2} +$$

$$+ \underbrace{\varepsilon_s(g,f,\beta)-\varepsilon_s(f,\beta)}_{t_3}$$

nu geldt dat $\lim\limits_{\beta\uparrow 1} t_3 = 0$ en als $g(s) \in A(s,f)$ dan

òf $t_1 > 0$

òf $t_1 = 0$ en dus $p(s,g(s)\ x(f) = x_s(f)$ èn

$$r(s,g(s)) + p(s,g(s))\ y(f) > x_s(f) + y_s(f).$$

Hieruit volgt dan dat $t_2 > 0$.

Dus in beide gevallen geldt dat $\lim\limits_{\beta\uparrow 1} (\frac{t_1}{1-\beta} + t_2 + t_3) > 0$.

<u>Lemma 6</u>

Indien voor f, g $\in$ F en $S_0 \subset S$ met $S_0 \neq \emptyset$ geldt dat

$$g(s) \in A(s,f) \text{ voor } s \in S_0$$

$$g(s) = f(s) \quad \text{voor } s \notin S_0$$

dan geldt dat $V_{\beta,g^\infty} > V_{\beta,f^\infty}$ voor β groot genoeg.

<u>Bewijs</u>

Passen we lemma 5 toe dan vinden we dat

$$V_{\beta,(g,f^\infty)}(s) > V_{\beta,f^\infty}(s) \text{ voor } s \in S_0 \text{ en } \beta \text{ groot genoeg en}$$

$$V_{\beta,(g,f^\infty)}(s) = V_{\beta,f^\infty}(s) \text{ voor } s \notin S_0.$$

Hieruit volgt dat $V_{\beta,(g,f^\infty)} > V_{\beta,f^\infty}$ voor β groot genoeg.

Uit lemma 2.2.4 deel ii) volgt dan dat $V_{\beta,g^\infty} > V_{\beta,f^\infty}$ voor β groot genoeg.

<u>Stelling 4</u>

Er is een $f \in F$ waarvoor $A(s,f) = \emptyset$ voor alle s.

<u>Bewijs</u>

We construeren inductief een rij $\{f_n\}_{n=1}^\infty \subset F$ door

i) kies f_0 willekeurig

ii) indien $A(s,f_n) = \emptyset$ dan nemen we $f_{n+1}(s)$ gelijk aan $f_n(s)$

iii) indien $A(s,f_n) \neq \emptyset$ kies dan voor $f_{n+1}(s)$ een element uit $A(s,f_n)$

Volgens lemma 6 geldt dan dat indien $A(s,f_n) \neq \emptyset$ voor alle s dan volgt dat $V_{\beta,f_{n+1}^\infty} > V_{\beta,f_n^\infty}$ voor β groot genoeg.

Nemen we aan dat er geen enkele n is waarvoor $A(s,f_n) = \emptyset$ voor alle s dan geldt voor ieder nat. getal N dat $V_{\beta,f_1^\infty} < V_{\beta,f_2^\infty} < \ldots < V_{\beta,f_N^\infty}$ voor β groot genoeg. Dit zou betekenen dat alle elementen uit de rij $\{f_n\}_{n=1}^\infty$ verschillend zijn. Echter dat is in tegenspraak met het feit dat F slechts eindig veel elementen bevat.

<u>Lemma 7</u>

Voor n = 0, 1, 2, ... is er een $R_n \in C_m$ zó dat

(2.4.4) $$\sup_{R \in C} V_R^n(s) = V_{R_n}^n(s) \text{ voor alle s} \quad \text{, en}$$

(2.4.5) $$\begin{cases} V_{R_0}^0(s) = \max_a r(s,a) \text{ en} \\ V_{R_{n+1}}^{n+1}(s) = \max_a \left(r(s,a) + p(s,a)V_{R_n}^n\right) \text{ voor } n = 0, 1, 2, \ldots \end{cases}$$

Bewijs

Met volledige inductie.

Zij $f_0 \in F$ zó dat $r(s,f_0(s)) = \max_a r(s,a)$ voor alle s, dan voldoet $R_0 = (f_0,R)$ met R een willekeurige markovstrategie aan (2.4.5). Verder geldt dat

$$\sup_{R \in C} V_R^0(s) = \underset{\pi \in \mathbb{P}(A|S)}{\text{supremum}} \sum_a \pi(a|s) r(s,a) \leq \max_a r(s,a) = r(s,f_0(s)) = V_{R_0}^0(s),$$

dus ook aan (2.4.4) is voldaan.

Neem aan dat $R_k \in C_m$ voldoet aan (2.4.4) en

$$V_{R_k}^k(s) = \max_a \left(r(s,a) + p(s,a)\, V_{R_{k-1}}^{k-1}\right) \text{ voor alle s.}$$

Kies dan $f_{k+1} \in F$ zó dat

$$r(s,f_{k+1}(s)) + p(s,f_{k+1}(s))\, V_{R_k}^k = \max_a \left(r(s,a) + p(s,a)\, V_{R_k}^k\right)$$

voor alle s,

dan volgt onmiddellijk dat $R_{k+1} = (f_{k+1},R_k)$ voldoet aan (2.4.5).

Dat ook aan (2.4.4) voldaan wordt zullen we bewijzen door aan te tonen dat voor alle $R \in C$ en $s \in S$ geldt dat $V_R^{k+1}(s) \leq V_{R_{k+1}}^{k+1}(s)$.

Kies R en s willekeurig dan volgens stelling 2.2.1 bestaat er een $R^* \in C_{rm}$ zeg $R^* = (\pi_0,\pi_1,\ldots)$ zó dat $V_{R^*}^n(s) = V_R^n(s)$ voor alle n.

$$\begin{aligned}
V_R^{k+1}(s) = V_{R^*}^{k+1}(s) &= \left(r(\pi_0) + P(\pi_0)\, V_{(\pi_1,\pi_2,\ldots)}^k\right)_s \\
&= \sum_a \pi_0(a|s)\left(r(s,a) + p(s,a)\, V_{(\pi_1,\pi_2,\ldots)}^k\right) \\
&\leq \sum_a \pi_0(a|s)\left(r(s,a) + p(s,a)\, V_{R_k}^k\right) \\
&\leq r(s,f_{k+1}(s)) + p(s,f_{k+1}(s))\, V_{R_k}^k = V_{R_{k+1}}^{k+1}(s).
\end{aligned}$$

Stelling 5

Indien voor $f \in F$ geldt dat $A(s,f) = \emptyset$ voor alle s dan is f^∞ optimaal in C t.o.v. G_R d.w.z.

$$\limsup_{N\to\infty} \frac{1}{N+1} V_R^N(s) \leq \lim_{N\to\infty} \frac{1}{N+1} V_{f^\infty}^N(s) \text{ voor alle } R \in C \text{ en alle } s \in S.$$

Bewijs

Indien $A(s,f) = \emptyset$ en $r(s,a) + p(s,a)\, y(f) > x_s(f) + y_s(f)$ dan moet gelden dat $p(s,a)\, x(f) < x_s(f)$. Hieruit volgt dat er dan een natuurlijk getal $N(s,a)$ bestaat zó dat voor $n \geq N(s,a)$

$$r(s,a) + p(s,a)y(f) - x_s(f) - y_s(f) \leq n(x_s(f) - p(s,a)x(f)). \tag{2.4.6}$$

Indien $A(s,f) = \emptyset$ en $r(s,a) + p(s,a)y(f) \leq x_s(f) + y_s(f)$ dan geldt (2.4.6) voor ieder natuurlijk getal omdat $x_s(f) - p(s,a)x(f) \geq 0$.
Dus indien $A(s,f) = \emptyset$ voor alle s dan is er bij iedere combinatie (s,a) een nat. getal $N(s,a)$ zó dat (2.4.6) geldt voor $n \geq N(s,a)$.
Zij $N = \max_{s,a} N(s,a)$ dan geldt voor $n \geq N$ en iedere paar (s,a) (2.4.6) die we herschrijven in

$$r(s,a) + p(s,a)y(f) + np(s,a)x(f) \leq (n+1)x_s(f) + y_s(f). \tag{2.4.7}$$

Zij $M = \max(\sup_{R\in C} V_R^n(s) - nx_s(f) - y_s(f))$ voor $n = 0, 1, \ldots, N$ en alle s en zij v de vektor met s-de component $v_s = M$ voor alle s.
Uit lemma 7 volgt dat $\sup_{R\in C} V_R^n$ bestaat voor $n = 0, 1, 2, \ldots$ en dat geldt

$$\sup_{R\in C} V_R^{n+1}(s) = \max_a (r(s,a) + p(s,a) \sup_{R\in C} V_R^n). \tag{2.4.8}$$

We bewijzen nu m.b.v. volledige inductie dat geldt dat

$$\sup_{R\in C} V_R^n \leq nx(f) + y(f) + v \quad \text{voor } n = 0, 1, 2, \ldots \tag{2.4.9}$$

Voor $n = 0, 1, \ldots, N$ volgt (2.4.9) onmiddellijk uit de definitie van de vektor v.

Neem aan dat (2.4.9) juist is voor $k \geq N$ dan geldt volgens (2.4.8) dat

$$\sup_{R \in C} V_R^{k+1}(s) = \max_a (r(s,a) + p(s,a) \sup_{R \in C} V_R^k)$$

en volgens de inductieaanname geldt dan

$$\leq \max_a (r(s,a) + p(s,a)(kx(f) + y(f) + v))$$

$$= \max_a (r(s,a) + p(s,a)\, y(f) + kp(s,a)x(f)) + v$$

en volgens (2.4.7)

$$\leq (k+1)x(f) + y(f) + v.$$

Uit (2.4.9) volgt dat voor willekeurige $R \in C$ geldt dat

$$G_R = \limsup_{N \to \infty} \frac{1}{N+1} V_R^N \leq \limsup_{N \to \infty} \frac{1}{N+1} (\sup_{R \in C} V_R^N) \leq$$

$$\leq \limsup_{N \to \infty} \frac{1}{N+1} (Nx(f) + y(f) + v)$$

$$= x(f) = G_{f^\infty}.$$

Stelling 6

Er is een $f \in F$ zó dat f^∞ optimaal is t.o.v. G_R.

Bewijs

De bewering is een onmiddellijk gevolg van de stellingen 4 en 5.

Stelling 7

Er is een $f \in F$ zó dat f^∞ optimaal is t.o.v. g_R.

Bewijs

Volgens stelling 6 is er een f^∞ optimaal t.o.v. G_R, voor deze f en willekeurige $R \in C$ geldt dat

$$g_R = \liminf_{N \to \infty} \frac{1}{N+1} V_R^N \leq \limsup_{N \to \infty} \frac{1}{N+1} V_R^N \leq G_{f^\infty} = g_{f^\infty} = x(f) \text{ (stelling 1).}$$

Lemma 8

Indien de vectoren x en y een oplossing zijn van

$$(2.4.10) \qquad P(f)x = x$$

$$(2.4.11) \qquad r(f) + P(f)y = x+y$$

dan geldt dat

$$x = x(f) \text{ en } y - P^*(f)y = y(f).$$

Bewijs

Als $P(f)x = x$ dan volgt dat $P^k(f)x = x$ voor iedere nat. getal k.
$\Rightarrow P^*(f)x = x$. Vermenigvuldigen we (2.4.11) met $P^*(f)$ dan vinden we dat $P^*(f)r(f) + P^*(f)P(f)y = P^*(f)x + P^*(f)y \Rightarrow P^*(f)r(f) = P^*(f)x$
$\Rightarrow x = P^*(f)r(f) = x(f)$.
Vermenigvuldigen we (2.4.11) met $H(f)$ en substitueren we $x(f)$ voor x, dan vinden we dat $H(f)r(f) + H(f)P(f)y = H(f)x(f) + H(f)y$. Uit lemma 4 deel iii) volgt dat $H(f)P(f) = H(f) - I + P^*(f)$, substitueren we dit in bovenstaande vergelijking dan vinden we dat

$$H(f)r(f) - y + P^*(f)y = H(f)P^*(f)r(f) = 0 \text{ (lemma 4 deel ii)}$$

dus

$$y - P^*(f)y = H(f)r(f) = y(f).$$

Opmerking 1

Dat $x = x(f)$ is reeds bewezen op pag. 9 e.v.

Lemma 9

Het stelsel vergelijkingen (2.4.10) en (2.4.11) gecombineerd met

$$(2.4.12) \qquad P^*(f)y = 0$$

heeft de unieke oplossing $x = x(f)$ en $y = y(f)$.

Bewijs

Uit lemma 8 en (2.4.12) volgt dat iedere oplossing voldoet aan $x = x(f)$ en $y = y(f)$. Dat $x(f)$, $y(f)$ een oplossing vormen van (2.4.10), (2.4.11) en (2.4.12) is m.b.v. lemma 4 te verifiëren.

Opmerking 2

Met behulp van dit lemma is een algorithme te construeren om een stationaire strategie te vinden die optimaal is ten opzichte van opbrengstfunktie G_R.

De n-de stap van het algorithme verloopt dan als volgt:

Met de in de (n-1)-ste stap verkregen funktie f_{n-1} bepalen we $x(f_{n-1})$ en $y(f_{n-1})$ door het stelselvergelijkingen $\begin{cases} P(f_{n-1})x = x \\ r(f_{n-1}) + P(f_{n-1})y = x+y \\ P^*(f_{n-1})y = 0 \end{cases}$

op te lossen. Met behulp van $x(f_{n-1})$ en $y(f_{n-1})$ bepalen we dan volgens definitie 1 $A(s,f_{n-1})$. Kies nu voor $f_n(s)$ een element uit $A(s,f_{n-1})$ indien $A(s,f_{n-1}) \neq \emptyset$. Is $A(s,f_{n-1}) = \emptyset$ neem dan $f_n(s)$ gelijk aan $f_{n-1}(s)$. Einde n-de stap. Volgens de stellingen 4 en 5 verkrijgen we dan in een eindig stappen een f zódat f^∞ optimaal is t.o.v. G_R.

Opmerking 3

Het in opmerking 2 beschreven algorithme is niet hetzelfde als het in hoofdstuk 1 paragraaf 1 beschreven algorithme van Howard.

Dat de algorithmes verschillend zijn tonen we met het volgende rekenvoorbeeld aan.

Neem aan dat S drie toestanden bevat zeg s_1, s_2 en s_3 en dat er een $f \in F$ bestaat zódat

$$P(f) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{en } r(f) = \begin{pmatrix} 0 \\ 1 \\ \frac{1}{2} \end{pmatrix}$$

Dan volgt uit $P^2(f) = P(f)$ dat $P^*(f) = P(f)$. Dus $x(f) = P^*(f)r(f) = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$.

We zien dat $\{s_1, s_2\}$ en $\{s_3\}$ de kernfuiken zijn dus de y-vektor die in het algorithme van Howard wordt bepaald (dit is de vektor t_i van pag. 9 e.v.) zeg y^H moet voldoen aan:

$$r(f) + P(f)y^H = x(f) + y^H,$$

met de extra voorwaarde dat $y_3^H = 0$ en y_1^H of $y_2^H = 0$; laten we kiezen $y_2^H = 0$.
Lossen we y_1^H hieruit op dan vinden we $y_1^H = -\frac{1}{2}$.
$y(f) = H(f)r(f)$ moet volgens lemma 9 voldoen aan

$$r(f) + P(f)y(f) = x(f) + y(f)$$

en

$$P^*(f)y(f) = 0.$$

Daar $P(f) = P^*(f)$ volgt hieruit dat $y(f) = r(f) - x(f)$. Dus $y(f) = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}$.
Veronderstel verder dat er een aktie a is zódat

$$p(s_2,a) = (\tfrac{1}{2},0,\tfrac{1}{2}) \text{ en } r(s_2,a) = 1.$$

Omdat $x_1(f) = x_2(f) = x_3(f)$ geldt dan $p(s_2,a)x(f) = x_2(f)$. Verder geldt dat

$$r(s_2,a) + p(s_2,a)y^H = 1 - \tfrac{1}{4} \quad \text{en} \quad x_2(f) + y_2^H = \tfrac{1}{2}$$

en

$$r(s_2,a) + p(s_2,a)y(f) = 1 - \tfrac{1}{4} \quad \text{en} \quad x_2(f) + y_2(f) = \tfrac{1}{2} + \tfrac{1}{2}.$$

Hieruit zien we dat aktie a in toestand s_2 in aanmerking komt om de strategie f^∞ te verbeteren volgens het algorithme van Howard echter niet volgens het in opmerking 2 beschreven algorithme.

<u>Stelling 8</u>
Indien voor $f \in F$ geldt dat voor alle s

$$(2.4.13) \qquad \max_{a \in A} \{p(s,a)x(f)\} = x_s(f)$$

en bovendien geldt dat

$$(2.4.14) \qquad \max\{r(s,a) + p(s,a)y(f)\} = x_s(f) + y_s(f)$$

voor de a's waarvoor $p(s,a)x(f) = x_s(f)$

dan is f^∞ optimaal t.o.v. G_R.

Bewijs

Indien voor f bovenstaande voorwaarden vervuld zijn dan geldt dat $A(s,f) = \emptyset$ voor alle s. Volgens stelling 5 is f^∞ dan optimaal t.o.v. G_R.

Opmerking 4

Het omgekeerde van bovenstaande stelling geldt niet. D.w.z. niet voor iedere $f \in F$ met f^∞ optimaal t.o.v. G_R geldt dat (2.4.13) en (2.4.14) vervuld zijn.

Om dit aannemelijk te maken veronderstellen we dat het Markovbeslissingsproces de volgende structuur heeft. Er is een toestand zeg s_0 die absorberend is d.w.z. $q(s_0|s_0,a) = 1$ en $r(s_0,a) = 0$ voor alle a en verder veronderstellen we dat $q(s_0|s,a) = 1-\beta$ met $0 < \beta < 1$ voor alle $s \neq s_0$ en alle a. Voor iedere $f \in F$ geldt dan dat $\{s_0\}$ de enige kernfuik is voor overgangsmatrix P(f) en dus $P^*_{s,s_0}(f) = 1$ voor alle s. Hieruit volgt dat $x(f) = P^*(f)r(f) = 0$ en uit de voorwaarde dat $P^*(f)y(f) = 0$ volgt dat $y_{s_0}(f) = 0$. Uit $x(f) = 0\ \forall f$ en stelling 6 volgt dan dat iedere stationaire strategie optimaal is t.o.v. G_R. Verder is (2.4.13) op triviale wijze vervuld voor iedere $f \in F$. Echter i.h.a. zal (2.4.14) niet voor iedere $f \in F$ vervuld zijn. Om dit in te zien reduceren we de toestandsverzameling S tot $\tilde{S} = S\backslash\{s_0\}$. Verder definiëren we $\tilde{q}(s_2|s_1,a) = \frac{1}{\beta}\, q(s_2|s_1,a)$ met $s_1, s_2 \in \tilde{S}$ en zullen we overeenkomstige objecten in $\tilde{S}$ met een "~" versieren. Dan gaat (2.4.14) over in

$$(2.4.15) \qquad \max_{a\in A} \{\tilde{r}(s,a) + \beta\tilde{p}(s,a)\tilde{y}(f)\} = \tilde{y}_s(f).$$

Daar y(f) voldoet aan $r(f) + P(f)y(f) = x(f) + y(f)$ vinden we dat $\tilde{r}(f) + \beta\tilde{P}(f)\tilde{y}(f) = \tilde{y}(f)$. Hieruit volgt m.b.v. lemma 2.2.3 dat $\tilde{V}_{\beta,f^\infty} = \tilde{y}(f)$.

Substitueren we dit in (2.4.15) dan vinden we

$$(2.4.16)\qquad \max_{a\in A} \{\tilde{r}(s,a) + \beta\tilde{p}(s,a)\tilde{V}_{\beta,f^\infty}\} = \tilde{V}_{\beta,f^\infty}.$$

Veronderstellen we nu dat (2.4.14) geldt voor iedere $f \in F$ dan volgt dat (2.4.16) geldt voor iedere $f \in \tilde{F}$. Volgens stelling 2.2.8 betekent dit dat iedere $f \in \tilde{F}$ een optimale stationaire strategie levert t.o.v. $\tilde{V}_{\beta,R}$. Echter dit is i.h.a. zeker niet juist.

Opmerking 5

C. Derman gebruikt in zijn onlangs verschenen boek "Finite State Markovian Decision Processes" een intuïtief argument uit een van zijn artikelen. Het argument is als volgt.

Indien R_0 een optimale stationaire strategie is voor een Markovbeslissingsproces t.o.v. g_R, dan blijft R_0 een optimale strategie voor die uitbreidingen van het Markovbeslissingsproces die we verkrijgen door nieuwe beslissingen toe te laten mits we aan het nemen van die nieuwe beslissingen maar voldoend kleine opbrengsten (d.w.z. in absolute waarde groot en negatief) verbinden.

Dit argument is onjuist zoals het volgende rekenvoorbeeld laat zien. Veronderstel dat er drie toestanden zijn, zeg 0, 1, 2 en één beslissing a_1 zódat

$$\begin{array}{lll} q(0|0,a_1) = 0 & q(1|0,a_1) = \frac{1}{2} & q(2|0,a_1) = \frac{1}{2} \\ q(0|1,a_1) = 0 & q(1|1,a_1) = 1 & q(2|1,a_1) = 0 \\ q(0|2,a_1) = 0 & q(1|2,a_1) = 0 & q(2|2,a_1) = 1. \end{array}$$

Zij verder $r(1,a_1) > r(2,a_1)$.

De verwachte gemiddelde opbrengst als we starten in 0 zal zijn $\frac{1}{2}(r(1,a_1) + r(2,a_1))$. Voeren we nu een beslissing a_2 in, zódat $q(0|i,a_2) = 1 \quad i = 0, 1, 2$ met opbrengst $r(i,a_2)$ $i = 0, 1, 2,$, dan zal de beslissingsregel bepaald door a_1 te kiezen in 0 en 1 en a_2 te kiezen in 2, de verwachte opbrengst startend vanuit 0 doen toenemen tot $r(1,a_1)$.

Hoofdstuk 3. Markovbeslissingsproblemen met een aftelbare toestandsruimte en een hoogstens aftelbare actieruimte.

3.1. De totale verwachte verdisconteerde opbrengst.

De toestandsruimte S is aftelbaar verondersteld. Derhalve mogen wij zonder beperking aannemen dat S de verzameling van de natuurlijke getallen is en dat toestand s het natuurlijke getal s is.

Het optimaliteitskriterium is de totale verwachte verdisconteerde opbrengst. In eerste instantie gaan wij uit van een hoogstens aftelbare actieruimte A (hoogstens aftelbaar betekent: aftelbaar of eindig). In sectie 3.1.2. zullen wij ons beperken tot het geval dat A eindig is.

Wij zullen laten zien dat als een optimale strategie bestaat, ook een optimale stationaire strategie bestaat en dat deze gevonden kan worden door een funktionaalvergelijking op te lossen. Aan de hand van een tegenvoorbeeld tonen wij aan dat een optimale strategie niet behoeft te bestaan als A aftelbaar is. Echter wij kunnen wel aantonen dat een optimale strategie altijd bestaat als A eindig is. De hiervoor genoemde resultaten worden bewezen onder de aanname dat de opbrengstfunktie r(s,a) uniform begrensd is. Wij zullen laten zien dat deze aanname een weinig verzwakt kan worden, doch in het algemeen niet verzwakt kan worden tot r(s,a) is eindig voor alle s,a.

3.1.1. De actieruimte A is hoogstens aftelbaar.

Wij maken de volgende modelveronderstelling. Er bestaat een eindige positieve constante M zo dat

$$|r(s,a)| \leq M \qquad \text{voor alle } s \in S \text{ en alle } a \in A.$$

Wij gaan uit van een vast gekozen verdisconteringsfactor β, $0 < \beta < 1$. Aangezien de opbrengstfunktie r(s,a) uniform begrensd is, geldt voor elke $R \in C$ en elke begintoestand s_0 dat $V_{\beta,R}(s_0)$ bestaat en in absolute waarde kleiner dan of gelijk aan $M/(1-\beta)$ is.

Men gaat eenvoudig na dat de stellingen 1 en 2 uit paragraaf 2.2 geldig blijven. De volgende stelling is derhalve geldig

Stelling 1. Voor iedere $R \in C$ en $s_0 \in S$ bestaat een $R^* \in C_{rm}$ zó dat

$$V_{\beta,R^*}(s_0) = V_{\beta,R}(s_0).$$

Vervolgens voeren wij enige definities en notaties in.

Definities en notaties

Laat $B^{(\infty)}$ de verzameling zijn van alle begrensde reële funkties gedefinieerd op S *). Wij noteren de ∞-dimensionale vektor waarvan de s^{de} component gelijk is aan $V_{\beta,R}(s)$ met V_R. Merk op dat $V_R \in B^{(\infty)}$ voor alle $R \in C$. Zij $\pi \in P(A|S)$ resp. $f \in F$ dan is $r(\pi)$ resp. $r(f)$ de ∞-dimensionale vektor met als componenten

$$\sum_{a \in A} \pi(a|s)\ r(s,a) \quad \text{resp.} \quad r(s,f(s)).$$

en $P(\pi)$ resp. $P(f)$ de ∞-dimensionale matrix met als (s,s')-de element

$$\sum_{a \in A} \pi(a|s)\ q(s'|s,a) \quad \text{resp.} \quad q(s'|s,\ f(s)).$$

De vektor met als s'-de element $q(s'|s,a)$ geven wij aan met $p(s,a)$. Laten $v,u \in B^{(\infty)}$. Dan is vu gedefinieerd door het reële getal

$$vu = \sum_{s \in S} v(s)\ u(s).$$

Laat $v \in B^{(\infty)}$ en laat M een ∞-dimensionale kansenmatrix zijn met $m(s,s')$ als $(s,s')^{de}$ element. Dan is Mv een vektor met als s^{de} component

$$\sum_{s' \in S} m(s,s')\ v(s').$$

*) De elementen van $B^{(\infty)}$ geven wij weer door ∞-dimensionale vektoren $v = (v(1),\ v(2),\ldots)$; $u = (u(1),\ u(2),\ldots)$; etc.

Zij $\pi \in P(A|S)$ dan is de afbeelding $L(\pi)$ van $B^{(\infty)}$ in $B^{(\infty)}$ als volgt gedefinieerd

$$L(\pi)v = r(\pi) + \beta P(\pi)v \,, \qquad v \in B^{(\infty)}.$$

Voor $f \in F$ is $L(f)$ de afbeelding van $B^{(\infty)}$ in $B^{(\infty)}$ met

$$L(f)v = r(f) + \beta P(f)v \,, \qquad v \in B^{(\infty)}.$$

Het behoeft geen betoog dat eigenschappen die gelden voor iedere $L(\pi)$, $\pi \in P(A|S)$, zeker gelden voor iedere $L(f)$, $f \in F$.

De vektorongelijkheden $u \leq v$ en $u < v$, waarbij $u, v \in B^{(\infty)}$, zijn als volgt gedefinieerd: $u \leq v \Longleftrightarrow u(s) \leq v(s)$ voor alle $s \in S$, en $u < v \Longleftrightarrow u \leq v$ én $u(s) < v(s)$ voor tenminste één $s \in S$.

Men gaat eenvoudig na dat de lemmas 1 t/m 4 uit paragraaf 2.2 geldig blijven. Wij kunnen dus de volgende stelling formuleren.

<u>Stelling 2</u>.

(a) Voor elke $\pi \in P(A|S)$ en alle $u, v \in B^{(\infty)}$ met $u \leq v$ geldt dat $L(\pi)u \leq L(\pi)v$.

(b) Voor elke $R = (\pi_0, \pi_1, \ldots) \in C_{rm}$ en elke $v \in B^{(\infty)}$ geldt dat

$$\lim_{N\to\infty} L(\pi_0)\ldots L(\pi_N)v = V_R.$$

(c) Voor elke $\pi \in P(A|S)$ en elke $u \in B^{(\infty)}$ geldt dat $V_{\pi^{(\infty)}} \leq u$ is als $L(\pi)u \leq u$. Deze bewering blijft geldig als wij het $\leq$ teken vervangen door het $\geq$ teken.

Wij kunnen nu de volgende belangrijke stelling bewijzen.

<u>Stelling 3</u>.

De funktionaalvergelijking

$$(1) \qquad u(s) = \sup_{a \in A} \{r(s,a) + \beta p(s,a)u\}, \qquad s \in S,$$

heeft als unieke begrensde oplossing

(2) $$u^* = \sup_{R\in C} V_R .$$

Voorts geldt dat bij elke $\varepsilon > 0$ een $f \in F$ te vinden is zó dat

(3) $$V_{\beta,f^{(\infty)}}(s) \geq u^*(s) - \varepsilon \qquad \text{voor alle } s \in S.$$

Bewijs.
Het bewijs valt in twee delen uiteen. In deel a) zullen wij aantonen dat de funktionaalvergelijking (1) een begrensde oplossing u^* bezit. In deel b) tonen wij aan dat u^* uniek is en als bijprodukt van het bewijs vinden wij de relaties (2) en (3).
(a) Wij zullen eerst onder de aanname dat $r(s,a) \geq 0$ is voor alle s,a, aantonen dat (1) een begrensde oplossing bezit.
Definieer

(4) $$u_0(s) = 0 \qquad \text{voor alle } s \in S,$$

en definieer voor $n \geq 1$ de funkties u_n door de recursieve betrekking

(5) $$u_n(s) = \sup_{a\in A} \{r(s,a) + \beta p(s,a)u_{n-1}\} \qquad \text{voor } s \in S.$$

Omdat $M \geq r(s,a) \geq 0$ voor alle s,a, volgt uit (4) en (5) dat $M \geq u_1(s) \geq 0 = u_0(s)$ voor alle $s \in S$. Men kan nu eenvoudig met volledige inductie bewijzen dat voor elke $s \in S$ geldt $u_n(s) \geq u_{n-1}(s)$, en $u_n(s) \leq M(1+\ldots+\beta^{n-1})$, $n \geq 1$. Derhalve bestaat

$$u^*(s) \overset{\text{def}}{=} \lim_{n\to\infty} u_n(s) \qquad \text{voor alle } s \in S.$$

Uit de eigenschappen van de funkties u_n volgt dat de funktie u^* begrensd is en dat $u^* \geq u_n$ voor alle $n \geq 0$. Wij zullen nu aantonen dat u^* aan (1) voldoet. Definieer daartoe de funktie v door

$$v(s) = \sup_{a \in A} \{r(s,a) + \beta p(s,a)u^*\} \qquad \text{voor } s \in S.$$

De funktie v is begrensd. Aangezien $u^* \geq u_n$ voor alle n, geldt voor alle $s \in S$ dat

$$v(s) \geq \sup_{a \in A} \{r(s,a) + \beta p(s,a)u_n\} = u_{n+1}(s) \qquad \text{voor } n \geq 0.$$

Uit deze ongelijkheid volgt $v \geq u^*$. De funkties u_1 en v zijn beiden begrensd, derhalve bestaat een constante c zó dat $u_1(s) \geq v(s) - c$ voor alle $s \in S$. Wij zullen nu met volledige induktie aantonen dat voor elke $s \in S$ geldt $u_n(s) \geq v(s) - \beta^{n-1}c$ voor alle $n \geq 1$. Stel dat voor alle s geldt $u_k(s) \geq v(s) - \beta^{k-1}c$, dan volgt uit (5) en de reeds bewezen ongelijkheid $v \geq u^*$ dat

$$\begin{aligned} u_{k+1}(s) &\geq \sup_{a \in A} \{r(s,a) + \beta p(s,a)v\} - \beta^k c \\ &\geq \sup_{a \in A} \{r(s,a) + \beta p(s,a)u^*\} - \beta^k c \\ &= v(s) - \beta^k c \qquad \text{voor alle } s \in S. \end{aligned}$$

Hieruit volgt dat $u^* \geq v$. Dus $u^* = v$ en u^* voldoet aan (1). Wij laten nu de aanname $r(s,a) \geq 0$ vallen. Aangezien $|r(s,a)| \leq M$ voor alle s,a, geldt dat $\hat{r}(s,a) = r(s,a) + M \geq 0$ voor alle s,a. Voorts is $\hat{r}(s,a)$ uniform begrensd. Derhalve bestaat een begrensde funktie $\hat{u}^*$ zó dat

$$\text{(6)} \qquad \hat{u}^*(s) = \sup_{a \in A} \{\hat{r}(s,a) + \beta p(s,a)\hat{u}^*\} \qquad \text{voor } s \in S.$$

Definieer nu u^* door $u^*(s) = \hat{u}^*(s) - M/(1-\beta)$, $s \in S$. Uiteraard is de funktie u^* begrensd. Door nu $\hat{u}^*(s) = u^*(s) + M/(1-\beta)$ en $\hat{r}(s,a) = r(s,a) + M$ te substitueren in (6) vinden wij dat u^* aan (1) voldoet

(b) Laat u^* een begrensde oplossing van (1) zijn. Uit $u^*(s) \geq r(s,a) + \beta p(s,a)u^*$ voor alle s,a volgt dat voor elke $\pi \in P(A|S)$ geldt

$$L(\pi)u^* \leq u^* \quad .$$

Uit een herhaalde toepassing van deze ongelijkheid en stelling 2(a) volgt dat voor elke $R = (\pi_0, \pi_1, \ldots) \in C_{rm}$ geldt dat $L(\pi_0)\ldots L(\pi_N)u^* \le u^*$, voor alle $N \ge 1$. Uit stelling 2(b) volgt nu dat $V_R \le u^*$ voor elke $R \in C_{rm}$. Toepassing van stelling 1 leert vervolgens dat $V_R \le u^*$ voor alle $R \in C$. Derhalve geldt

$$\sup_{R \in C} V_R \le u^*. \tag{7}$$

Zij $\varepsilon > 0$ een vast doch willekeurig gekozen getal. Aangezien u^* aan (1) voldoet, is bij elke s een actie a_s te vinden zó dat

$$r(s,a_s) + \beta p(s,a_s)u^* \ge u^*(s) - \varepsilon \qquad \text{voor alle } s \in S.$$

Definieer nu $f \in F$ door

$$f(s) = a_s \qquad \text{voor alle } s \in S,$$

dan geldt

$$L(f)u^*(s) \ge u^*(s) - \varepsilon \qquad \text{voor alle } s \in S. \tag{8}$$

Met volledige inductie tonen wij nu aan dat voor elke $s \in S$ geldt $L^n(f)\, u^*(s) \ge u^*(s) - (\varepsilon + \varepsilon\beta + \ldots + \varepsilon\beta^{n-1})$ voor $n \ge 1$. Voor $n = 1$ is de ongelijkheid juist. Stel dat de ongelijkheid juist is voor $n = k$, dan volgt met behulp van stelling 2(a) en (8) dat

$$L^{k+1}(f)\, u^*(s) = L(f)\, L^k(f)\, u^*(s) \ge L(f)\, u^*(s) - \beta\varepsilon \sum_{j=0}^{k-1} \beta^j \ge \tag{9}$$

$$\ge u^*(s) - \varepsilon \sum_{j=0}^{k} \beta^j \qquad \text{voor } s \in S.$$

Aangezien (9) geldt voor elke $k \ge 1$, volgt door toepassing van stelling 2(b) dat

$$V_{\beta, f^{(\infty)}}(s) \ge u^*(s) - \frac{\varepsilon}{1-\beta} \qquad \text{voor alle } s \in S.$$

Aangezien $\varepsilon > 0$ willekeurig gekozen is, geldt dus dat bij elke $\varepsilon > 0$ een $f \in F$ te vinden is met $V_{\beta,f^{(\infty)}}(s) \geq u^*(s) - \varepsilon$ voor alle $s \in S$.
Uit deze relatie en uit relatie (7) volgt direct dat u^* gegeven wordt door (2).
Hiermee is het bewijs van de stelling voltooid.

Een direct gevolg van stelling 3 is

Stelling 4.
Een strategie $R_0 \in C$ is optimaal dan en slechts dan als

$$(10) \qquad V_{\beta,R_0}(s) = \sup_{a \in A} \{r(s,a) + \beta p(s,a) V_{R_0}\} \qquad \text{voor alle } s \in S.$$

Stelling 5.
Stel dat de klasse C een optimale strategie bevat. Dan bestaat ook een optimale stationaire strategie. Verder geldt dan dat de funktionaalvergelijking

$$(11) \qquad u(s) = \max_{a \in A} \{r(s,a) + \beta p(s,a) u\} \qquad \text{voor } s \in S$$

als unieke begrensde oplossing heeft

$$(12) \qquad u^* = \max_{R \in C} V_R \quad ,$$

en voorts is elke stationaire strategie $f^{(\infty)}$, zó dat voor elke $s \in S$ de actie $f(s)$ het rechterlid van (11) maximaliseert, optimaal.

Bewijs.
Laat $R_0 \in C$ optimaal zijn. Uit stelling 4 volgt dat V_{R_0} aan de funktionaalvergelijking (10) voldoet. Wij zullen nu eerst aantonen dat wij in (10) het supremum wordt aangenomen voor alle s, d.w.z. wij mogen sup vervangen door max. Stel dat een toestand s_0 bestaat, zó dat

$$(13) \qquad V_{\beta,R_0}(s_0) > r(s_0,a) + \beta p(s_0,a) V_{R_0} \qquad \text{voor alle } a \in A.$$

Zij $R^{*} = (\pi_0^{*}, \pi_1^{*}, \ldots) \in C_{rm}$ zo gekozen dat $V_{\beta,R^{*}}(s_0) = V_{\beta,R_0}(s_0)$ (zie stelling 1). Laat $R^{**} = (\pi_1^{*}, \pi_2^{*}, \ldots)$. Uit (13) volgt dan

$$\begin{aligned} V_{\beta,R^{*}}(s_0) = V_{\beta,R_0}(s_0) &> \sum_a \pi_0^{*}(a|s_0)\ \{r(s_0,a) + \beta p(s_0,a) V_{R_0}\} \geq \\ &\geq \sum_a \pi_0^{*}(a|s_0)\ \{r(s_0,a) + \beta p(s_0,a) V_{R^{**}}\} = \\ &= V_{\beta,R^{*}}(s_0). \end{aligned}$$

Dit is een tegenspraak. Dus bij elke s behoort een $a \in A$, noem deze a_s, zó dat

$$(14) \qquad V_{\beta,R_0}(s) = r(s,a_s) + \beta p(s,a_s) V_{R_0} \qquad \text{voor alle } s \in S.$$

Zij $f \in F$ gedefinieerd door $f(s) = a_s$ voor alle $s \in S$. Uit (14) volgt dan $L(f)V_{R_0} = V_{R_0}$. Een toepassing van stelling 2c leert vervolgens dat $V_{f^{(\infty)}} = V_{R_0}$, oftewel $f^{(\infty)}$ is optimaal. Hiermee is bewezen dat een optimale stationaire strategie bestaat en dat elke $f^{(\infty)}$, zó dat voor elke s de actie f(s) het rechterlid van (11) maximaliseert, optimaal is. Relatie (12) volgt tenslotte uit stelling 3.

Wij zullen nu aan de hand van een tegenvoorbeeld laten zien dat een optimale strategie niet behoeft te bestaan als A aftelbaar is.

Tegenvoorbeeld.

$$\begin{aligned} &S = A = \{1,2,\ldots\}; \\ &q(1|1,a) = 1,\ r(1,a) = 0 \quad \text{voor alle } a \in A; \\ &q(1|s,a) = 1,\ r(s,a) = 1 - \frac{1}{a} \quad \text{voor alle } s \geq 2,\ a \in A. \end{aligned}$$

Het zal duidelijk zijn dat

$$(15) \qquad V_{\beta,R}(s) < 1 \qquad \text{voor alle } R \in C \text{ en } s \geq 2.$$

Voor strategie $f^{(\infty)}$ met $f(s) = a$ voor alle $s \in S$ geldt uiteraard

(16) $$V_{\beta,f^{(\infty)}}(s) = 1 - \frac{1}{a} \qquad \text{voor } s \geq 2.$$

Uit (15) en (16) volgt dat bij elke $R \in C$ en $s_0 \in S$, $s_0 \geq 2$, een $f \in F$ te vinden is, zó dat $V_{\beta,f^{(\infty)}}(s_0) > V_{\beta,R}(s_0)$. Derhalve bestaat in dit voorbeeld geen optimale strategie.

3.1.2. Een eindige actieruimte A.

Wij nemen nu aan dat A eindig is. Wij zullen aantonen dat nu altijd een optimale strategie bestaat.

Stelling 6.

Er bestaat een optimale stationaire strategie.

Bewijs.

Aangezien A eindig is; wordt in de funktionaalvergelijking (1) voor elke $s \in S$ het supremum aangenomen. Dus bij elke s bestaat een actie a_s zó dat

$$u^*(s) = r(s,a_s) + \beta p(s,a_s)u^* \qquad \text{voor alle } s \in S,$$

waarbij $u^* = \sup_{R\in C} V_R$. Zij nu $f_0 \in F$ gedefinieerd door $f_0(s) = a_s$, $s \in S$, dan geldt dus $L(f_0)u^* = u^*$. Uit stelling 2(c) volgt nu

$$V_{f_0^{(\infty)}} = u^* = \sup_{R\in C} V_R ,$$

oftewel strategie $f_0^{(\infty)}$ is optimaal.

Opmerking 3.2.1.

Men kan gemakkelijk nagaan, dat de resultaten van deze paragraaf 3.1 geldig blijven indien in plaats van een actieruimte A voor iedere toestand s een verzameling A(s) van toegelaten beslissingen gegeven is.

Voor sectie 3.1.2. moeten wij dan aannemen $A(s)$ eindig voor alle s. De bewijzen verlopen geheel analoog, alleen de notatie in de bewijsvoering wordt gecompliceerder.

Opmerking 3.2.2.

De stelling 6 blijft geldig als de aanname dat $r(s,a)$ uniform begrensd is, als volgt wordt verzwakt: Er bestaat een eindige constante M zó dat

$$(17) \qquad \inf_{s\in S} \max_{a\in A} r(s,a) \geq -M \quad \text{en} \quad r(s,a) \leq M \quad \text{voor alle } s \in S \text{ en } a \in A.$$

Bewijs: Eerst merken wij op dat $V_{\beta,R}(s)$ bestaat voor alle $s \in S$ en $R \in C$ (met $-\infty$ als mogelijke waarde), omdat $r(s,a)$ uniform van boven begrensd is. Voorts geldt dat $V_{\beta,R}(s) \leq M/(1-\beta)$ voor alle s,R. Definieer

$$(18) \qquad \bar{r}(s,a) = \max\{r(s,a); -\frac{(1+\beta)}{1-\beta} M-1\} \quad \text{voor } s \in S,\ a \in A\ .$$

De funktie $r(s,a)$ is uniform begrensd in s en a. Laat $\bar{V}_{\beta,R}(s)$ de totale verwachte verdisconteerde opbrengst zijn behorende bij begintoestand s en strategie $R \in C$ voor het beslissingsprobleem dat uit het oorspronkelijke volgt door $r(s,a)$ te vervangen door $\bar{r}(s,a)$. Uiteraard geldt

$$(19) \qquad \bar{V}_{\beta,R}(s) \geq V_{\beta,R}(s) \quad \text{voor alle } s \in S \text{ en } R \in C.$$

Aangezien $\bar{r}(s,a)$ uniform begrensd is, bestaat volgens de stelling 4 en 6 een strategie $f_0^{(\infty)}$ met $f_0 \in F$ zodat $\bar{V}_{f^{(\infty)}} \geq \bar{V}_R$ voor alle $R \in C$. Als wij nu kunnen aantonen dat $\bar{r}(s,f_0(s)) = r(s,f_0(s))$ voor alle $s \in S$ dan geldt $V_{f^{(\infty)}} = \bar{V}_{f^{(\infty)}}$ en met behulp van (19) volgt dan dan $V_{f^{(\infty)}} \geq V_R$ voor alle $R \in C$, oftewel $f_0^{(\infty)}$ is ook optimaal voor het oorspronkelijke beslissingprobleem. Om aan te tonen dat $\bar{r}(s,f_0(s)) = r(s,f_0(s))$ voor alle s merken wij op dat

(20) $$\overline{V}_{\beta,f^{(\infty)}}(s) = \max_{R\in C} \overline{V}_{\beta,R}(s) \geq \frac{-M}{1-\beta} \quad \text{voor alle } s \in S,$$

aangezien bij elke $s \in S$ een $a \in A$ te vinden is met $\overline{r}(s,a) \geq -M$ (zie de aanname (17)). Vanwege $\overline{r}(s,a) \leq M$ voor alle s,a, geldt dat

(21) $$\overline{V}_{\beta,R}(s) \leq M/(1-\beta) \qquad \text{voor alle } s \in S \text{ en } R \in C.$$

De funktie $\overline{V}_{f^{(\infty)}}$ voldoet aan

(22) $$\overline{V}_{\beta,f^{(\infty)}}(s) = \overline{r}(s,f(s)) + \beta p(s,f(s))\overline{V}_{f^{(\infty)}} \qquad \text{voor } s \in S.$$

Stel nu dat een toestand s_0 bestaat met $\overline{r}(s_0,f(s_0)) \neq r(s,f(s))$, dan volgt uit (18), (21) en (22) dat

(23) $$\overline{V}_{\beta,f^{(\infty)}}(s_0) \leq -\frac{(1+\beta)}{1-\beta} M-1 + \frac{\beta M}{1-\beta} < \frac{M}{1-\beta} \quad .$$

Echter de relatie (23) is in tegenspraak met (20). Derhalve geldt $\overline{r}(s,f(s)) = r(s,f(s))$ voor alle s. Waarmee het bewijs voltooid is.

Opmerking 3.2.3.

Stelling 6 behoeft niet geldig te zijn als van de funktie $r(s,a)$ alleen wordt geeist dat $r(s,a)$ eindig is voor alle s,a. Wij zullen dit laten zien aan de hand van een tegenvoorbeeld, dat afkomstig is van DERMAN.

Tegenvoorbeeld.

$S = \{0,1,-1,2,-2,\ldots\}$; $A = \{1,2\}$;

$q(0|0,1) = q(0|0,2) = 1$; $r(0,1) = r(0,2) = 0$;

$q(n+1|n,1) = p, q(0|n,1) = 1-p$, waarbij $0 < p < 1$,
$q(-n|n,2) = 1$, $r(n,1) = r(n,2) = 1/(2\beta p)^{n-1}$, $n = 1,2,\ldots$;

$q(0|-n,1) = q(0|-n,2) = 1$, $r(-n;1) = r(-n;2) = (n-1)/(\beta p)^{n-1}$, $n = 1,2,\ldots$

Als $f \in F$ en $f(n) = 1$ voor alle $n \geq 1$, dan geldt

$$V_{\beta,f^{(\infty)}}(1) = \sum_{k=0}^{\infty} (\beta p)^k \frac{1}{(2\beta p)^k} = \sum_{k=0}^{\infty} 2^{-k} < \infty .$$

Als $f(n) = 2$ voor tenminste één $n \geq 1$ en n_0 is de kleinste $n \geq 1$ met $f(n) = 2$, dan geldt

$$V_{\beta,f^{(\infty)}}(1) = \sum_{k=0}^{n_0-1} (\beta p)^k \frac{1}{(2\beta p)^k} + (\beta p)^{n_0-1} \beta \frac{(n_0-1)}{(\beta p)^{n_0-1}} < \infty.$$

Derhalve geldt voor elke $f \in F$ dat

(24) $$V_{\beta,f^{(\infty)}}(1) < \infty .$$

Laat de strategie $R(t)$ als volgt gedefinieerd zijn. Kies de actie 2 op het beslissingstijdstip t en kies de actie 1 op de overige beslissingstijdstippen. Voor de strategie $R(t)$ geldt

(25) $$V_{\beta,R(t)}(1) = \sum_{k=0}^{t} (\beta p)^k \frac{1}{(2\beta p)^k} + (\beta p)^t \frac{\beta t}{(\beta p)^t}$$

$$= \sum_{k=0}^{t} 2^{-k} + \beta t .$$

De strategie $R(t)$ behoort niet tot de klasse van de stationaire strategieen. Uit (24) en (25) volgt dat voor elke $f_0 \in F$ geldt dat

$$V_{\beta,R(t)}(1) > V_{\beta,f_0^{(\infty)}}(1) \qquad \text{voor } t \text{ voldoende groot.}$$

Opmerking 3.2.4 (tegenvoorbeeld voor S aftelbaar, A aftelbaar en $r(s,a)$ uniform begrensd).

In het tegenvoorbeeld gegeven op blz. 101 geldt voor elke $s \in S$ dat $\sup_a r(s,a)$ voor geen enkele a wordt aangenomen. Men kan zich nu afvragen of voor het geval S en A aftelbaar en $r(s,a)$ uniform begrensd een optimale strategie bestaat als $\sup_a r(s,a) = \max_a r(s,a)$ voor alle $s \in S$. Het volgende tegenvoorbeeld, dat afkomstig is van J. WIJNGAARD, toont aan dat dit niet het geval behoeft te zijn.

Tegenvoorbeeld

$$S = A = \{1,2,\ldots\}; \quad q(a|s,a) = 1 \text{ en } r(s,a) = 1-\frac{1}{s} \text{ voor alle } s \in S,\ a \in A.$$

Wij zullen aantonen dat bij elke $s_0 \in S$ en $f \in F$ een $g \in F$ te vinden is, zó dat

$$V_{\beta,g^{(\infty)}}(s_0) > V_{\beta,f^{(\infty)}}(s_0). \tag{26}$$

Laten $s_0 \in S$ en $f \in F$ vast gekozen zijn. Stel $a_0 = f(s_0)$, $a_1 = f(a_0)$ $a_2 = f(a_1),\ldots,a_{i+1} = f(a_i),\ldots$. Dan geldt

$$V_{\beta,f^{(\infty)}}(s_0) = \sum_{i=0}^{\infty} \beta^i(1-\frac{1}{a_i}) .$$

Laat $b_0 > a_0$ zijn. Kies nu $g \in F$ zó dat $g(s_0) = b_0$, $b_1 = g(b_0) > a_1$, $b_2 = g(b_1) > a_2,\ldots,b_{i+1} = g(b_i) > a_i,\ldots$. Voor strategie $g^{(\infty)}$ geldt de ongelijkheid (26). Uit (26) en stelling 5 volgt nu dat geen optimale strategie bestaat.

Opmerking 3.2.5 (vervolg op opmerking 3.2.3; tegenvoorbeeld voor S aftelbaar, A eindig en $r(s,a)$ eindig).

In het tegenvoorbeeld gegeven in opmerking 3.2.3 bestaat weliswaar geen optimale stationaire strategie, maar bestaat wel een optimale strategie. Wij zullen aantonen dat voor de strategie $R^* = (\pi_0^*,\pi_1^*,\ldots) \in C_{rm}$ zó dat

$$\pi_0^*(2|s) = \pi_1^*(2|s) = 0 \text{ en } \pi_n^*(2|s) = \frac{1}{n} \text{ voor alle } s \geq 1,\ n \geq 2,$$

geldt

$$V_{\beta,R^*}(s) = \infty \qquad \text{voor } s = 1,2,\dots \ .$$

Voor strategie R^* en begintoestand $s \geq 1$ geldt

$$(27) \qquad V_{\beta,R^*}(s) = \sum_{n=0}^{\infty} \beta^n \prod_{i=0}^{n-1} \pi_i^*(1|s+i)p^n \frac{1}{(2\beta p)^{n+s-1}} +$$

$$+ \sum_{n=0}^{\infty} \pi_n^*(2|n) \prod_{i=0}^{n-1} \pi_i^*(1|s+i)p^n \left\{ \frac{\beta^n}{(2\beta p)^{n+s-1}} + \frac{\beta^{n+1}(n+s-1)}{(\beta p)^{n+s-1}} \right\} .$$

Aangezien $\prod_{i=0}^{n-1} \pi_i^*(1|s+i) = \prod_{i=2}^{n-1} (1-\frac{1}{i}) = \frac{1}{n-1}$ en $\pi_n^*(2|n) = \frac{1}{n}$, $n \geq 3$ volgt uit (27) dat $V_{\beta,R^*}(s) = \infty$. Dus strategie R^* is optimaal.

Wij zullen nu voor het geval S aftelbaar, A eindig en $r(s,a)$ eindig, een tegenvoorbeeld geven waarin geen optimale strategie bestaat

$$S = \{1,1',2,2',3,3',\dots\} \ ; \quad A = \{1,2\};$$

$$q(i+1|i,1) = 1, \ q(i'|i,2) = 1, \ r(i,1) = 0, \ r(i,2) = (1-\tfrac{1}{i})\beta^{-i}, \qquad i = 1,2,\dots;$$

$$q(i'|i',1) = q(i'|i',2) = 1, \ r(i',1) = r(i',2) = 0, \qquad i' = 1',2',\dots \ .$$

Voor elke $R \in C$ en elke toestand $s \in \{1',2',\dots\}$ geldt uiteraard $V_{\beta,R}(s) = 0$. Wij zullen aantonen dat

$$(28) \qquad V_{\beta,R}(s) < \beta^{-s} \qquad \text{voor alle } R \in C, \ s = 1,2,\dots \ ,$$

en

$$(29) \qquad \sup_{R \in C} V_{\beta,R}(s) = \beta^{-s} \qquad \text{voor } s = 1,2,\dots \ .$$

Op grond van stelling 1 (blz. 95) kunnen wij volstaan met (28) aan te tonen voor elke $R \in C_{rm}$.
Laten $R = (\pi_0,\pi_1,\ldots) \in C_{rm}$ en begintoestand $s_0 (s_0 \geq 1)$ vast gekozen zijn. Als $\pi_n(1|s_0+n) = 1$ voor alle $n \geq 0$, dan geldt $V_{\beta,R}(s_0) = 0$ en is dus aan (28) voldaan. Neem derhalve aan dat $\pi_n(2|s_0+n)$ positief is voor tenminste één n.
Er geldt

$$(30) \qquad V_{\beta,R}(s_0) = \sum_{n=0}^{\infty} \beta^n \pi_n(2|s_0+n) \prod_{i=0}^{n-1} \pi_i(1|s_0+i)(1 - \frac{1}{s_0+n}) \beta^{-(s_0+n)}.$$

Voor elke rij van getallen $\{d_i\}$, $i \geq 0$, met $0 \leq d_i \leq 1$, $i \geq 0$, geldt

$$(31) \qquad \sum_{n=0}^{\infty} (1-d_n) \prod_{i=0}^{n-1} d_i = 1 - \prod_{n=0}^{\infty} d_n \leq 1$$

Uit (30), (31) en het feit dat $\pi_n(2|s_0+n) \prod_{i=0}^{n-1} \pi_i(1|s_0+i)$ positief is voor tenminste één n, volgt dat

$$(32) \qquad V_{\beta,R}(s_0) < \beta^{-s_0} \sum_{n=0}^{\infty} \pi_n(2|s_0+n) \prod_{i=0}^{n-1} \pi_i(1|s_0+i) \leq \beta^{-s_0}.$$

Hiermee is (28) bewezen.
Laat $f_n \in F$ zodanig zijn dat

$$f_n(s) = \begin{cases} 1 & \text{voor } s = 1,\ldots,n-1, \\ 2 & \text{voor } s = n, n+1,\ldots , \end{cases}$$

dan geldt voor elke $n > s_0$ dat

$$(33) \qquad V_{\beta,f_n^{(\infty)}}(s_0) = \beta^{n-s_0}(1 - \frac{1}{n})\beta^{-n} = \beta^{-s_0}(1 - \frac{1}{n}).$$

Uit (33) volgt nu (29). Uit (28) en (29) volgt dat geen optimale strategie bestaat.

3.2. De gemiddelde verwachte opbrengst per tijdseenheid.

In deze paragraaf zullen we eerst noodzakelijke voorwaarden opstellen waaronder een optimale strategie bestaat (secties 3.2.1 en 3.2.2). In sectie 3.2.3 geven wij een drietal tegenvoorbeelden voor het geval r(s,a) uniform begrensd is en A eindig is. In het eerste tegenvoorbeeld bestaat geen optimale strategie. In het tweede tegenvoorbeeld bestaat een optimale gerandomiseerde stationaire strategie, doch bestaat geen optimale stationaire strategie. In het derde tegenvoorbeeld bestaat een optimale Markov strategie, doch geen optimale gerandomiseerde stationaire strategie.

Tenslotte laten wij voor een tweetal specifieke beslissingsproblemen zien dat een optimale strategie altijd bestaat; in de secties 3.2.4 en 3.2.5 behandelen wij een vervangings- en een voorraadprobleem.

3.2.1 S aftelbaar en A hoogstens aftelbaar.

Wij nemen aan dat de opbrengstfunctie r(s,a) zodanig is dat $\sum_{s,a} P_R(x_t=s,y_t=a|x_0=s_0)\ r(s,a)$ bestaat voor alle $t \geq 1$, $s_0 \in S$ en $R \in C$. Een voldoende voorwaarde hiervoor is dat r(s,a) uniform van onder begrensd is.

Als optimaliteitskriterium nemen wij het sterkste kriterium

(1) $$G_R(s_0) = \limsup_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \sum_{s,a} P_R(x_t=s,y_t=a|x_0=s_0)\ r(s,a).$$

Wij merken op dat bij elke $R \in C$ en $s_0 \in S$ een $R^* \in C_{rm}$ bestaat zó dat

(2) $$G_{R^*}(s_0) = G_R(s_0),$$

aangezien de stelling van Derman en Strauch ook geldig is voor S aftelbaar en A hoogstens aftelbaar (zie stelling 1 in paragraaf 2.2).

Wij zullen nu de volgende belangrijke stelling bewijzen:

Stelling 1 (ROSS)

Stel er bestaat een verzameling van eindige getallen $\{g,v(s)\}$, $s \in S$, zó dat

(3) $$g+v(s) = \max_{a\in A}\{r(s,a) + p(s,a)v\} \qquad \text{voor alle } s \in S,$$

en

(4) $$\lim_{n\to\infty} \frac{1}{n} E_R(v(x_n)|x_0=s_0) = 0 \qquad \text{voor alle } s_0 \in S,\ R \in C.$$

Dan bestaat een optimale stationaire strategie $f_0^{(\infty)}$. Er geldt

(5) $$G_{f_0^{(\infty)}}(s) = g \qquad \text{voor alle } s \in S,$$

en verder geldt dat elke stationaire strategie $f^{(\infty)}$, zó dat voor elke $s \in S$ de actie $f(s)$ het rechterlid van (3) maximaliseert, optimaal is.

Bewijs

Zij $f_0 \in F$ zodanig dat voor elke $s \in S$ de actie $f_0(s)$ het rechterlid van (3) maximaliseert. Stel $x_0 = s_0$ en stel dat strategie $R \in C$ wordt gevolgd. Voor elke $n \geq 1$ geldt de identiteit

(6) $$E_R[\sum_{t=1}^{n} \{v(x_t) - E_R(v(x_t)|x_0,y_0,\ldots,x_{t-1},y_{t-1})\}].$$

Er geldt

(7) $$E_R(v(x_t)|x_0=s_0,y_0=a_0,\ldots,x_{t-1}=s_{t-1},y_{t-1}=a_{t-1}) = p(s_{t-1},a_{t-1})v =$$

$$= r(s_{t-1},a_{t-1}) + p(s_{t-1},a_{t-1})v - r(s_{t-1},a_{t-1}) \leq$$

$$\leq \max_{a\in A}\{r(s_{t-1},a) + p(s_{t-1},a)v\} - r(s_{t-1},a_{t-1}) =$$

$$= g + v(s_{t-1}) - r(s_{t-1},a_{t-1}).$$

Als $R = f_0^{(\infty)}$, dan geldt in (7) het gelijkheidsteken. Uit (6) en (7) volgt dat voor alle $n \geq 1$ geldt

$$(8) \qquad E_R[\sum_{t=1}^{N} \{v(x_t)-g-v(x_{t-1})+r(x_{t-1},y_{t-1})\}] \leq 0.$$

Als $R = f_0^{(\infty)}$, dan geldt in (8) het gelijkheidsteken. Uit (8) volgt

$$(9) \qquad g \geq \frac{1}{n} E_R\, v(x_n) - \frac{1}{n} E_R\, v(x_0) + \frac{1}{n} E_R\{\sum_{t=0}^{n-1} r(x_t,y_t)\} \quad \text{voor } n \geq 1.$$

Als $R = f_0^{(\infty)}$, dan geldt in (9) het gelijkheidsteken. Uit (9) volgt dat $g \geq G_R(s_0)$, waarbij het gelijkheidsteken geldt als $R = f_0^{(\infty)}$. Aangezien s_0 en R willekeurig gekozen zijn, is hiermee de stelling bewezen.

Opmerking

Als $\{g,v(s)\}$, $s \in S$ aan (3) voldoet en c is een constante, dan voldoet ook $\{g,v(s)+c\}$, $s \in S$, aan (3).

Indien de rij $\{v(s)\}$, $s \in S$, begrensd is, dan is automatisch aan de voorwaarde (4) voldaan.

3.2.2. S aftelbaar en A eindig.

Wij zullen in deze sectie aantonen dat een optimale stationaire strategie bestaat onder de volgende voorwaarden.

(A) Er bestaat een eindige constante M zó dat $|r(s,a)| \leq M$ voor alle $s \in S$, $a \in A$.

(B) Voor elke $f \in F$ geldt dat de Markov-keten $\{x_n\}$, $n \geq 0$ behorende bij strategie $f^{(\infty)}$ geen twee of meer disjuncte fuiken bevat.

(C) Er bestaat één of ander toestand, stel toestand 1, en een constante $N < \infty$, zó dat voor elke $f \in F$ toestand 1 een positieve terugkeer-toestand is onder strategie $f^{(\infty)}$ en $\mu_{s1}(f) \leq N$ voor alle $s \in S$. Hierbij is $\mu_{s1}(f)$ de verwachte tijd nodig om uitgaande van toestand s de toestand 1 te bereiken, indien strategie $f^{(\infty)}$ wordt gevolgd.

Daartoe eerst het volgend lemma.

Lemma 1

Stel dat voor elke $i = 1, 2, \ldots$ de rij van reële getallen $\{u_n(i)\}$, $n \geq 1$, begrensd is. Dan bestaan eindige getallen $\{u(i)\}$, $i \geq 1$, en een deelrij $\{m_k\}$, $k \geq 1$, zó dat

$$\lim_{k\to\infty} u_{m_k}(i) = u(i) \qquad \text{voor alle } i = 1, 2, \ldots .$$

Bewijs

Uit de analyse is het welbekend dat elke begrensde rij tenminste één convergente deelrij bevat en voorts dat elke deelrij van een convergente rij met limiet b ook limiet b heeft. Laat $\{u_{m_k^{(1)}}(1)\}$, $k \geq 1$, een convergente deelrij met limiet $u(1)$ zijn van de begrensde rij $\{u_n(1)\}$, $n \geq 1$. De begrensde rij $\{u_{m_k^{(1)}}(2)\}$, $k \geq 1$, bevat een convergente deelrij $\{u_{m_k^{(2)}}(2)\}$, $k \geq 1$. Laat $u(2)$ de limiet van deze convergente deelrij zijn. Zo voortgaand kunnen wij voor elke i een convergente deelrij $\{u_{m_k^{(i)}}(i)\}$, $k \geq 1$, van de rij $\{u_{m_k^{(i-1)}}(i)\}$, $k \geq 1$, construeren. Laat $u(i)$ de limiet van de convergente rij $\{u_{m_k^{(i)}}(i)\}$, $k \geq 1$, zijn. Stel nu $m_k = m_k^{(k)}$, $k \geq 1$. Voor elke vaste i geldt dat de rij $\{u_{m_k}(i)\}$, $k \geq i$, een deelrij is van de convergente rij $\{u_{m_k^{(i)}}(i)\}$, $k \geq 1$. Dus voor elke i convergeert de rij $\{u_{m_k}(i)\}$, $k \geq 1$, naar $u(i)$.

Stelling 2

Onder de voorwaarden (A), (B) en (C) bestaat een optimale stationaire strategie.

Bewijs

Het bewijs valt in twee delen uiteen. In deel (a) tonen wij aan dat voor elke $\beta \in (0,1)$ en elke $f \in F$ geldt

(10) $$|V_{\beta,f^{(\infty)}}(s) - V_{\beta,f^{(\infty)}}(1)| \leq 2MN \qquad \text{voor alle } s \in S.$$

In deel (b) van het bewijs zullen wij laten zien dat uit (10), de resultaten van paragraaf 3.1 en stelling 1 op eenvoudige wijze het bestaan van een optimale stationaire strategie volgt.

(a) Stel $\hat{r}(s,a) = r(s,a) + M$ voor alle $s \in S$, $a \in A$. Op grond van aanname (A) geldt $0 \leq \hat{r}(s,a) \leq 2M$ voor alle s, a. Laat $\hat{V}_{\beta,R}(s)$ behoren bij de opbrengstfunctie $\hat{r}(s,a)$. Voor elke $R \in C$ en $\beta \in (0,1)$ geldt dat $\hat{V}_{\beta,R}(s) = V_{\beta,R}(s) + M/(1-\beta)$ voor alle $s \in S$, dus voor alle $f \in F$ en $\beta \in (0,1)$ geldt

(11) $$\hat{V}_{\beta,f^{(\infty)}}(s) - \hat{V}_{\beta,f^{(\infty)}}(1) = V_{\beta,f^{(\infty)}}(s) - V_{\beta,f^{(\infty)}}(1), \quad s \in S.$$

Definieer

$$T = \min\{t \mid t \geq 1, x_t = 1\}.$$

Zij $f \in F$ willekeurig gekozen. Als de strategie $f^{(\infty)}$ wordt toegepast, dan is de stochastische variable T welgedefinieerd en heeft dan als verwachting

$$E_{f^{(\infty)}}(T \mid x_0 = s) = \mu_{s1}(f), \qquad s \in S.$$

Voor de totale verdisconteerde opbrengst $\hat{V}_{\beta,f^{(\infty)}}(s)$ geldt

(12) $$\hat{V}_{\beta,f^{(\infty)}}(s) = E_{f^{(\infty)}}[\{\sum_{t=0}^{T-1} \beta^t \hat{r}(x_t, f(x_t))\} \mid x_0 = s] +$$
$$+ E_{f^{(\infty)}}[\{\sum_{t=T}^{\infty} \beta^t \hat{r}(x_t, f(x_t))\} \mid x_0 = s] \text{ voor } s \in S.$$

Aangezien $\hat{r}(s,a) \leq 2M$, $\hat{V}_{\beta,R}(s) \geq 0$ en $\mu_{s1}(f) \leq N$, volgt uit (12) dat

(13) $$\hat{V}_{\beta,f^{(\infty)}}(s) \leq 2M\, \mu_{s1}(f) + E_{f^{(\infty)}}(\beta^T \mid x_0 = s)\, \hat{V}_{\beta,f^{(\infty)}}(1) \leq$$

$$\leq 2MN + \hat{V}_{\beta,f^{(\infty)}}(1) \qquad \text{voor alle } s \in S.$$

Aangezien $\hat{r}(s,a) \geq 0$ is, volgt uit (12) tevens dat

$$\hat{V}_{\beta,f^{(\infty)}}(s) \geq E_{f^{(\infty)}}(\beta^T | x_0=s)\ \hat{V}_{\beta,f^{(\infty)}}(1) \qquad \text{voor alle } s \in S.$$

Deze ongelijkheid schrijven wij in de equivalente vorm

$$(14) \qquad \hat{V}_{\beta,f^{(\infty)}}(1) \leq \hat{V}_{\beta,f^{(\infty)}}(s) + (1-E_{f^{(\infty)}}(\beta^T|x_0=s))\ \hat{V}_{\beta,f^{(\infty)}}(1), \quad s \in S.$$

De funktie β^x, $x \geq 0$, is een convexe funktie van x. Derhalve leert de ongelijkheid van Jense *) ons dat

$$E_{f^{(\infty)}}(\beta^T|x_0=s) \geq \beta^{\mu_{s1}(f)} \geq \beta^N \qquad \text{voor alle } s \in S.$$

Voorts geldt $\hat{V}_{\beta,f^{(\infty)}}(s) \leq 2M/(1-\beta)$ voor alle $s \in S$. Derhalve volgt uit (14) en de relatie $(1-\beta^N)/(1-\beta) = 1+\beta+\ldots+\beta^{N-1}$ dat

$$(15) \qquad \hat{V}_{\beta,f^{(\infty)}}(1) \leq \hat{V}_{\beta,f^{(\infty)}}(s) + (1-\beta^N)\ 2M/(1-\beta) \leq$$

$$\leq \hat{V}_{\beta,f^{(\infty)}}(s) + 2MN \qquad \text{voor alle } s \in S.$$

Uit (13), (14) en (11) volgt nu de te bewijzen relatie (10).

(b) De stellingen 4, 5 en 6 uit paragraaf 3.1 leren ons dat voor elke vaste $\beta \in (0,1)$ een optimale stationaire strategie $f_\beta^{(\infty)}$ bestaat en dat $V_{\beta,f}{}^{(\infty)}(s)$ de unieke oplossing is van de funktionaalvergelijking

$$(16) \qquad V_{\beta,f^{(\infty)}}(s) = \max_{a\in A}\{r(s,a)+\beta p(s,a)V_{f_\beta^{(\infty)}}\}, \qquad s \in S.$$

Definieer nu voor elke $\beta \in (0,1)$

$$(17) \qquad v_\beta(s) = V_{\beta,f^{(\infty)}}(s) - V_{\beta,f^{(\infty)}}(1), \qquad s \in S$$

---

*) Zie b.v. W. Feller, Introduction to the Probability Theory and its Applications, Volume II, p. 151.

en

(18) $\qquad g_\beta = (1-\beta)\, V_{\beta,f^{(\infty)}}(1).$

Uit (10) volgt dat $|v_\beta(s)| \leq 2MN$ voor alle $s \in S$ en $\beta \in (0,1)$. Voorts geldt $|g_\beta| \leq M$ voor alle $\beta \in (0,1)$, omdat $|V_{\beta,R}(s)| \leq M/(1-\beta)$ voor alle $s \in S$ en $R \in C$.

Door van beide leden van (16) de grootheid $V_{\beta,f^{(\infty)}}(1)$ af te trekken, vindt men

(19) $\qquad g_\beta + v_\beta(s) = \max_{a\in A}\{r(s,a) + \beta p(s,a)v\} \qquad$ voor $s \in S$.

Aangezien de funkties g_β en $v_\beta(s)$ uniform begrensd zijn in β en s, bestaan op grond van lemma 1 een begrensde verzameling van getallen $\{g,v(s)\}$, $s \in S$ en een deelrij $\{\beta_r\}$, $r \geq 1$, met $\beta_r \uparrow 1$ voor $r \to \infty$, zó dat

(20) $\qquad \lim_{r\to\infty} g_{\beta_r} = g$ en $\lim_{r\to\infty} v_{\beta_r}(s) = v(s) \qquad$ voor alle $s \in S$.

Uit (19), (20), de eindigheid van A en stelling 1.2.1c (hoofdstuk 1) volgt dat

$$g + v(s) = \max_{a\in A}\{r(s,a) + p(s,a)v\} \qquad \text{voor alle } s \in S.$$

Aangezien de verzameling van getallen $\{g,v(s)\}$, $s \in S$, begrensd is, is voorwaarde (4) uit stelling 1 ook vervuld. Derhalve bestaat een optimale stationaire strategie.

<u>Opmerking 3.2.2.1</u>

Uit het bewijs van een stelling 2 blijkt dat de noodzakelijke voorwaarden (A), (B) en (C) als volgt een weinig verzwakt kunnen worden. "Er bestaat één of andere toestand, stel toestand 1, een constante $N < \infty$ en een rij $\{\beta_r\}$, $r \geq 1$ met $\beta_r \uparrow 1$ voor $r \to \infty$, zó dat

$$|v_{\beta_r}(s)| \leq N \qquad \text{voor alle } r \geq 1 \text{ en } s \in S,$$

waarbij $v_\beta(s)$ gedefinieerd is door $v_\beta(s) = \max_{R\in C} V_{\beta,R}(s) - \max_{R\in C} V_{\beta,R}(1)$".

Opmerking 3.2.2.2

De resultaten van de secties 3.2.1 en 3.2.2 blijven geldig indien in plaats van eenzelfde actieruimte A voor iedere toestand s een verzameling A(s) van toegelaten beslissingen gegeven wordt voor elke toestnad s. In sectie 3.2.2 moeten wij dan eisen dat A(s) eindig is voor elke $s \in S$.

3.2.3. Tegenvoorbeelden

Voorbeeld 1 (MAITRA) Er bestaat geen optimale strategie.

$S = (0,0',1,1',2,2',\dots)$; $A = (1,2)$;

$q(i+1|i,1) = 1$, $q(i'|i,2) = 1$, $r(i,1) = r(i,2) = -1$

voor $i = 0, 1, 2, \dots$;

$q(i'|i',1) = q(i'|i',2) = 1$, $r(i',1) = r(i',2) = w(i')$

voor $i' = 0',1',2',\dots,$

waarbij $\{w(i')\}$ een monotoon stijgende rij van negatieve getallen is met 0 als limiet.

Stel $x_0 = 0$. Wij zullen eerst aantonen dat

$$G_R(0) < 0 \qquad \text{voor elke } R \in C. \tag{21}$$

Op grond van (2) kunnen wij volstaan met (21) aan te tonen voor elke $R \in C_{rm}$. Laat $R = (\pi_0,\pi_1,\dots) \in C_{rm}$. Stel er bestaat een i met $\pi_i(2|i) > 0$. Laat i_0 de kleinste $i \geq 0$ zijn met $\pi_i(2|i) > 0$. Dan geldt $G_R(0) \leq \pi_{i_0}(2|i_0)w(i_0') < 0$. Als $\pi_i(2|i) = 0$ voor alle $i \geq 0$, dan geldt $G_R(0) = -1$.

Hiermee is (21) aangetoond. Zij nu R_k een strategie die actie 1 voorschrijft in de toestanden $i = 0, \dots, k-1$ en actie 2 in de toestand $i = k$ voorschrijft. Dan geldt

(22) $\quad G_{R_k}(0) = w(k')$.

Uit (21), (22) en het feit dat $w(k') \uparrow 0$ als $k' \to \infty$ volgt dat $\sup_{R \in C} G_R(0) = 0$. Dus er bestaat geen optimale strategie, omdat het supremum voor geen enkele $R \in C$ wordt aangenomen.

Voorbeeld 2 (DERMAN) Er bestaat een optimale gerandomiseerde stationaire strategie en er bestaat geen optimale stationaire strategie.

$S = \{0,1,2,\ldots\}$; $A = \{1,2\}$

$q(0|0,1) = q(0|0,2) = 0$, $q(s|0,1) = q(s|0,2) = q_s > 0$ voor $s \geq 1$;

$q(s|s,1) = 1$, $q(0|s,2) = 1$ voor $s \geq 1$; $r(s,1) = r(s,2) = w(s)$ voor $s \geq 0$,

waarbij $\{w(s)\}$ een rij van negatieve getallen is met $\sum |w(s)| < \infty$.

Uit $r(s,a) < 0$ voor alle s, a, volgt dat

(23) $\quad G_R(s) \leq 0 \qquad$ voor alle $s \in S$ en $R \in C$.

Wij zullen nu aantonen dat voor elke $f \in F$ geldt

(24) $\quad G_{f^{(\infty)}}(0) < 0$

en dat de klasse C_{rs} een strategie $\pi^{(\infty)}$ bevat met

(25) $\quad G_{\pi^{(\infty)}}(s) = 0 \qquad$ voor alle $s \in S$.

Uit (23) t/m (25) volgt dan dat de klasse C_{rs} een optimale strategie bevat en dat geen optimale stationaire bestaat.

Stel $x_0 = 0$. Laat $f \in F$ willekeurig gekozen zijn. Definieer

(26) $\quad S_f = \{s | s \geq 1, f(s)=1\}$.

Onderscheid (a) S_f is leeg en (b) S_f is niet-leeg.

(a) S_f is leeg. Dan is het systeem op de tijdstippen $t = 0, 2, 4, \ldots$ in toestand 0 en op de tijdstippen $t = 1, 3, 5, \ldots$ in één van de toestanden $1, 2, \ldots$. De kans dat het systeem op $t = 2n+1$ in toestand s is wordt gegeven door q_s. Dus $r(x_{2n}, f(x_{2n})) = w(0)$ en
$E_{f^{(\infty)}}\, r(x_{2n+1}, f(x_{2n+1})) = \sum_{s \geq 1} w(s) q_s$, $n \geq 0$.
Derhalve

$$G_{f^{(\infty)}}(0) = (w_0 + \sum_{s=1}^{\infty} w(s) q_s)/2 < 0.$$

(b) S_f is niet-leeg. Het stochastische proces $\{x_n\}$, $n \geq 0$, behorende bij strategie $f^{(\infty)}$ is een aftelbare Markov-keten met $P(f)$ als matrix van overgangskansen. Laten $p_{ss'}^{(k)}(f)$ de k-stapsovergangskansen van deze Markov-keten zijn, en laat $p_{ss'}^{(0)}(f) = 1$ voor $s = s'$ en laat $p_{ss'}^{(0)}(f) = 0$ voor $s \neq s'$. Aangezien de getallen $w(s)$ negatief zijn en $\sum |w(s)| < \infty$, volgt met behulp van stelling 1.2.1(b), 1.2.1(c) en stelling 1.2.12(a) dat

$$G_{f^{(\infty)}}(0) = \limsup_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \sum_{s\in S} p_{0s}^{(t)}(f) w(s) =$$

$$= \limsup_{n\to\infty} \sum_{s\in S} w(s) \frac{1}{n} \sum_{t=0}^{n-1} p_{0s}^{(t)}(f) =$$

$$= \sum_{s\in S} w(s)\, p_{0s}^{*}(f).$$

Na toestand $s \notin S_f$ wordt toestand 0 als volgende toestand aangenomen. Vanuit toestand 0 kan iedere toestand in S_f bereikt worden en elke toestand in S_f is absorberend. Derhalve is elke toestand $s \notin S_f$ inessentieel. Uit stelling 1.2.14 en stelling 1.2.12(c) volgt dan dat $p_{0s}^{*}(f) = 0$ voor $s \notin S_f$. Dus

$$(28) \qquad G_{f^{(\infty)}}(0) = \sum_{s\in S_f} w(s) p_{0s}^{*}(f).$$

Definieer voor elke $s \in S$,

$$g_{0s}^{(n)}(f) = P_{f^{(\infty)}}(x_n=s,\ x_k\neq s \text{ voor } 1\leq k<n \mid x_0=0), \qquad n \geq 1$$

en

$$g_{0s}(f) = \sum_{n=1}^{\infty} g_{0s}^{(n)}(f).$$

Stelling 1.2.12(b) leert ons dat $p_{0s}^{*}(f) = g_{0s}(f)p_{ss}^{*}(f)$ voor alle $s \in S$. Elke toestand $s \in S_f$ is absorberend, dus $p_{ss}^{(k)}(f) = 1$ voor alle $k \geq 1$ als $s \in S_f$. Hieruit volgt dat $p_{ss}^{*}(f) = 1$ voor $s \in S_f$. Derhalve geldt

$$(29) \qquad p_{0s}^{*}(f) = g_{0s}(f) \qquad \text{voor } s \in S_f.$$

Het zal duidelijk zijn dat voor elke $s \in S_f$ geldt

$$(30) \qquad g_{0s}^{(n)}(f) = \begin{cases} 0 & \text{voor } n \text{ even}, \\ \left(\sum_{j\notin S_f} q_j\right)^k q_s & \text{voor } n = 2k+1,\ k = 0, 1, \ldots . \end{cases}$$

Derhalve geldt

$$(31) \qquad g_{0s}(f) = q_s / \left(\sum_{j\in S_f} q_j\right) \qquad \text{voor } s \in S_f.$$

Uit (28), (29) en (31) volgt nu

$$(32) \qquad G_{f^{(\infty)}}(0) = \left\{\sum_{s\in S_f} w(s)q_s\right\} / \sum_{s\in S_f} q_s < 0.$$

Hiermee is het bewijs van relatie (24) voltooid.

Zij $\pi \in P(A|S)$ zo gekosen dat

$$(33) \qquad \pi(2|s) = q_s/s \qquad \text{voor } s \geq 1.$$

Het stochastische proces $\{x_n\}$, $n \geq 0$ geassocieerd met strategie $\pi^{(\infty)}$ is een aftelbare Markov-keten. Aangezien voor deze Markov-keten alle toestanden onderling bereikbaar zijn, is de verzameling van alle toestanden een kernfuik.

Wij zullen nu aantonen dat toestand 0 een nultoestand is onder strategie $\pi^{(\infty)}$. Uit

$$P_{\pi^{(\infty)}}(x_t=0 \text{ voor één of andere } t\geq 1 | x_0=0\} =$$

$$= 1 - P_{\pi^{(\infty)}}(x_t\neq 0 \text{ voor alle } t\geq 1 | x_0=0\} =$$

$$= 1 - \lim_{t\to\infty} \sum_{s=1}^{\infty} q_s(1-\pi(2|s))^t = 1$$

volgt dat toestand 0 een terugkeertoestand is. In bovenstaande relatie mogen limiet en sommatie verwisseld worden op grond van stelling 1.2.1 (c). Onder strategie $\pi^{(\infty)}$ is de verwachte tijd nodig om uitgaande van toestand $s \neq 0$ de toestand 0 te bereiken gelijk aan $1/\pi(2|s) = s/q_s$. Derhalve is de verwachte tijd nodig om uitgaande van toestand 0 in toestnad 0 terug te keren gelijk aan

$$1 + \sum_{s=1}^{\infty} q_s \frac{s}{q_s} = \infty.$$

Dus toestand 0 is een nultoestand onder strategie $\pi^{(\infty)}$. Aangezien de verzameling van alle toestanden een kernfuik is, volgt uit stelling 1.2.15 dat alle toestanden nultoestanden zijn onder $\pi^{(\infty)}$. Stelling 1.2.12c leert ons vervolgens dat elk element van de matrix $P^*(\pi)$ gelijk aan nul is. Laten $p_{ss'}^{(k)}(\pi)$ de elementen zijn van het k-voudig matrixprodukt $P^k(\pi)$, en stel $p_{ss'}^{(0)}(\pi) = 0$ voor $s = s'$ en $p_{ss'}^{(0)}(\pi) = 0$ voor $s \neq s'$. Aangezien de getallen $w(s)$ negatief zijn en $\sum |w(s)| < \infty$, volgt met behulp van stelling 1.2.1(a), 1.2.12(a) en 1.2.1(c) dat

$$G_{\pi^{(\infty)}}(s) = \limsup_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \sum_{s'\in S} p_{ss'}^{(t)}(\pi)w(s') =$$

$$= \limsup_{n\to\infty} \sum_{s'\in S} w(s') \frac{1}{n} \sum_{t=0}^{n-1} p_{ss'}^{(t)}(\pi) =$$

$$= \sum_{s'\in S} w(s')\, p_{ss'}^{*}(\pi) = 0 \qquad \text{voor } s \in S.$$

Hiermee is het bewijs voltooid.

Voorbeeld 3 (FISHER en ROSS) Er bestaat een optimale Markov-strategie en er bestaat geen optimale gerandomiseerde stationaire strategie.

$$S = \{0,1,1',2,2',\dots\};\ A = \{1,2\}$$

$$q(i|0,1) = q(i'|0,2) = (\tfrac{3}{2})(\tfrac{1}{4})^i,\ i > 0,\ r(0,1) = r(0,2) = -1;$$

$$q(0|i,1) = q(i'|i,1) = \tfrac{1}{2},\ q(0|i,2) = q(i+1|i,2) = \tfrac{1}{2},\ i > 0;$$

$$q(0|i',1) = q(0|i',2) = (\tfrac{1}{2})^i,\ q(i'|i',1) = q(i'|i',2) = 1 - (\tfrac{1}{2})^i,\ i > 0;$$

$$r(i,1) = r(i,2) = r(i',1) = r(i',2) = 0,\ i > 0.$$

Wij zullen aantonen dat

(34) $\quad G_R(s) \le -\frac{1}{5}$ $\quad$ voor alle $s \in S$, $R \in C$,

en

(35) $\quad G_R(s) < -\frac{1}{5}$ $\quad$ voor alle $s \in S$, $R \in C_{rs}$.

Voorts zullen wij een strategie $R^* \in C_m$ construeren, waarvoor geldt

(36) $\quad G_{R^*}(s) = -\frac{1}{5}$ $\quad$ voor alle $s \in S$.

Uit (34) t/m (36) volgt dan dat een optimale Markov-strategie bestaat en dat geen optimale gerandomiseerde stationaire strategie bestaat.

Wij merken op dat op grond van (2) het voldoende is om (34) aan te tonen voor elke $R \in C_{rm}$.

Wij zullen eerst enige notatie invoeren. Stel dat $x_0 = 0$ en dat strategie $R \in C_{rm}$ wordt gevolgd. Wij definieren $I_0 = 0$ en I_n als de lengte van het tijdsinterval tussen het tijdstip waarop toestand 0 voor de $(n-1)^{ste}$ keer wordt aangenomen *) en het tijdstip waarop toestand 0 voor de n^{de} keer wordt aangenomen. Wij spreken hierbij af dat de toestand 0 op $t = 0$ voor de 0^{de} keer wordt aangenomen. Wij definieren

$$T_n = I_0 + \dots + I_n \qquad n \ge 1.$$

*) Wij zullen hierna zien dat deze stochastische variabelen met kans 1 eindig zijn.

Dus op tijdstip T_n treedt toestand 0 voor de n^{de} keer op. Aangezien strategie $R \in C_{rm}$ "geheugenloos" is, is de verdeling van I_n volledig bepaald door T_{n-1}. Laat N_n, $n \geq 0$ gedefinieerd zijn door

$$T_{N_n} \leq n < T_{N_n+1}.$$

Dus $N_0 = 0$ en voor $n \geq 1$ is N_n gelijk aan het aantal keren dat toestand 0 in het tijdsinterval $[1,n]$ wordt aangenomen.

Uit de definitie van N_n en de structuur van de opbrengstfunktie $r(s,a)$ volgt direct dat

$$\text{(37)} \qquad G_R(0) = \limsup_{n\to\infty} -\frac{1}{n} E_R(N_n+1) =$$

$$= -\liminf_{n\to\infty} \frac{1}{n} E_R(N_n+1) \qquad \text{voor alle } R \in C_{rm}.$$

Wij zullen nu eerst aantonen dat voor elke $R \in C_{rm}$ en elke mogelijke waarde k van T_{n-1} geldt

$$\text{(38)} \qquad E_R(I_n | T_{n-1}=k) < 5 \qquad \text{voor alle } n \geq 1.$$

Daartoe definieren wij voor elke $s \in S$ en elke $R \in C_{rm}$ de grootheid $\mu_{s0}(R)$ als het verwachte tijdstip waarop voor het eerst een overgang naar toestand 0 plaatsvindt, gegeven dat $x_0 = s$ en dat strategie R wordt gevolgd. Stel ter afkorting $\mu_{s0}(\pi^{(\infty)}) = \mu_{s0}(\pi)$ voor $\pi \in P(A|S)$.

Laat $f_n \in F$ als volgt gedefinieerd zijn

$$f_n(s) = \begin{cases} 2 & \text{voor } s = 1, \ldots, n-1, \\ 1 & \text{voor } s \neq 1, \ldots, n-1. \end{cases}$$

Het zal duidelijk zijn dat voor vaste $n \geq 2$ geldt

$$\text{(39)} \qquad \mu_{00}(f_n) = 1 + \sum_{j=1}^{\infty} (\tfrac{3}{2})(\tfrac{1}{4})^j \mu_{j0}(f_n) + \sum_{j=1}^{\infty} (\tfrac{3}{2})(\tfrac{1}{4})^j 2^j.$$

Voor $\mu_{j0}(f_n)$ geldt

$$(40)\qquad \mu_{j0}(f_n) = \begin{cases} 2^j & j \geq n, \\ \frac{1}{2} + 2(\frac{1}{2})^2 + \ldots + (n-j)(\frac{1}{2})^{n-j} + (\frac{1}{2})^{n-j}\{n-j+2^n\} = & \\ = 2 + 2^j - (\frac{1}{2})^{n-j-1}, & 1 \leq j < n. \end{cases}$$

Uit (39) en (40) volgt na enig rekenwerk

$$(41)\qquad \mu_{00}(f_n) = 5 - (\tfrac{1}{2})^n \sum_{j=1}^{n-1} (\tfrac{3}{2})(\tfrac{1}{2})^{j-1} - 3 \sum_{j=n}^{\infty} (\tfrac{1}{4})^j.$$

Dus

$$(42)\qquad \mu_{00}(f_n) < 5 \qquad \text{voor alle } n \geq 2,$$

en

$$(43)\qquad \lim_{n\to\infty} \mu_{00}(f_n) = 5.$$

Wij merken op dat relatie (43) intuitief niet overeenstemt met

$$\mu_{00}(f) = 1 + \sum_{j=1}^{\infty} (\tfrac{3}{2})(\tfrac{1}{4})^j 2^j + \sum_{j=1}^{\infty} (\tfrac{3}{2})(\tfrac{1}{4})^j 2 = 3\tfrac{1}{2},$$

waarin f gegeven wordt door $f(s) = 2$ voor alle $s \in S$.

Stel $R = (\pi_0, \pi_1, \ldots) \in C_{rm}$. Laat $R(k) = (\pi_k, \pi_{k+1}, \ldots)$, $k \geq 1$, dan behoort $R(k)$ ook tot C_{rm}. Aangezien $E_R(I_n | T_{n-1}=k) = E_{R(k)}I_1$, is het om (38) te bewijzen voldoende om aan te tonen dat voor elke $R \in C_{rm}$ geldt $E_R I_1 < 5$.

Zij $R = (\pi_0, \pi_1, \ldots) \in C_{rm}$ en stel $R(1) = (\pi_1, \pi_2, \ldots)$. Dan geldt

$$(44)\qquad E_R I_1 = 1 + \sum_{j=1}^{\infty} (\tfrac{3}{2})(\tfrac{1}{4})^j \mu_{j0}(R(1)) + \sum_{j=1}^{\infty} (\tfrac{3}{2})(\tfrac{1}{4})^j 2^j.$$

Voor $\mu_{j0}(R(1))$ geldt

$$(45) \qquad \mu_{j0}(R(1)) = \sum_{n=j}^{\infty} \{(1-\pi_{n-j+1}(2|n)) \prod_{k=j}^{n-1} \pi_{k-j+1}(2|k)\ \mu_{j0}(f_n)\} +$$

$$+ 2 \prod_{k=j}^{\infty} \pi_{k-j+1}(2|k) <$$

$$< (2+2^j)\ [\sum_{n=j}^{\infty} \{(1-\pi_{n-j+1}(2|n)) \prod_{k=j}^{n-1} \pi_{k-j+1}(2|k)\} +$$

$$+ \sum_{k=j}^{\infty} \pi_{k-j+1}(2|k)], \qquad j \geq 1.$$

Zij $\{a_i\}$, $i \geq 1$ een rij getallen met $0 \leq a_i \leq 1$ voor alle i. Dan geldt voor elke vaste $H \geq 1$ dat

$$(46) \qquad \sum_{n=H}^{N} (1-a_n) \sum_{k=H}^{n-1} a_k = 1 - (a_H \dots a_N) \qquad \text{voor alle } N \geq H+1.$$

Deze relatie kan eenvoudig met volledige induktie naar N bewezen worden.

Uit (45) en (46) volgt

$$(47) \qquad \mu_{j0}(R(1)) < 2+2^j \qquad \text{voor } j \geq 1.$$

De relaties (44) en (47) leiden tot

$$(48) \qquad E_R I_1 < 5 \qquad \text{voor alle } R \in C_{rm}.$$

Hiermee is nu (38) bewezen. Als bijproduct van het bewijs vinden wij dat $\mu_{s0}(R) < \infty$ voor alle $s \in S$ en $R \in C_{rm}$.

Uit $q(s|0,a) > 0$ en $q(0|s,a) > 0$ vooralle s, a, volgt dat voor elke gerandomiseerde stationaire strategie het stochastische proces $\{x_n\}$, $n \geq 0$, een Markov-keten is, waarin alle toestanden onderling bereikbaar zijn. Dus onder elke strategie $\pi^{(\infty)}$ is de verzameling van alle toestanden een kernfuik. Voorts zijn alle toestanden positieve terugkeertoestanden, omdat $\mu_{00}(\pi) < \infty$ is (zie ook stelling 1.2.15). Voor elke strategie $\pi^{(\infty)}$ zijn I_1, I_2, ... onderling onafhankelijk en identiek verdeeld met verwachting $\mu_{00}(\pi)$. Laten $p_{ss'}^{(k)}(\pi)$ de elementen zijn van het k-voudig matrixproduct $P^k(\pi)$, en stel $p^{(0)}_{ss'}(\pi) = 1$ voor $s = s'$ en $p^{(0)}_{ss'}(\pi) = 0$ voor $s \neq s'$. Met behulp van stelling 1.2.12(a), 1.2.20 en 1.2.12(c) vinden wij dat voor elke gerandomiseerde stationaire strategie geldt

$$(49) \qquad G_{\pi^{(\infty)}}(s) = \limsup_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \sum_{s'\in S} p^{(t)}_{ss'}(\pi)\, r(\pi)(s') =$$

$$= \limsup_{n\to\infty} -\frac{1}{n} \sum_{t=0}^{n-1} p^{(t)}_{s0}(\pi) =$$

$$= -\lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} p^{(t)}_{s0}(\ )$$

$$= -p^{*}_{s0}(\pi) = -\frac{1}{\mu_{00}(\pi)} \qquad \text{voor alle } s \in S.$$

Uit (49) en de ongelijkheid $\mu_{00}(\pi) < 5$ volgt relatie (35).

Wij zullen nu (34) verifieren. Stel strategie $R \in C_{rm}$ wordt gevolgd. Definieer

$$(50) \qquad K_n = T_n - 5n, \qquad n \geq 0.$$

Dus $K_0 = 0$. Voor elke $n \geq 0$ geldt

$$E_R(K_{n+1}|K_0,\ldots,K_n) = E_R(K_n+I_{n+1}-5|K_0,\ldots,K_n) =$$

$$= K_n + E_R(I_{n+1}|T_n) - 5 \leq K_n.$$

Derhalve is de rij van stochastische variabelen $\{K_n\}$, $n \geq 0$, een supermartingaal. De stochastische variabele $N_n + 1$ is een stochastische stoptijd van de supermartingaal $\{K_n\}$. Aangezien $E_R|K_n| \leq E_R T_n + 5n < \infty$, $n \geq 0$, en voor n vast de stochastische variabele $N_n + 1$ begrensd is, kunnen wij een standaardstelling uit de martingalentheorie toepassen (zie J. Doob, Stochastic Processes, blz. 302). Deze stelling leert ons

(51) $$E_R(K_{N_n+1}) \leq E_R K_0 = 0 \qquad \text{voor alle } n \geq 1.$$

Uit (50) en (51) volgt

(52) $$E_R(T_{N_n+1}) \leq 5\; E_R(N_n+1).$$

Aangezien per definitie geldt $T_{N_n+1} > n$, volgt uit (52) dat

$$\frac{1}{n}\, E_R(N_n+1) > 1/5.$$

Derhalve geldt

$$\liminf_{n\to\infty} \frac{1}{n}\, E_R(N_n+1) \geq 1/5.$$

Hiermee is relatie (34) bewezen. Rest ons nog een strategie R^* te construeren die aan (36) voldoet.
Daartoe merken wij eerst het volgende op. Zij $f \in F$ en laat

$$C_n(s;f) \overset{\text{def}}{=} \sum_{t=0}^{n-1} p_{s0}^{(t)}(f) \qquad \text{voor } n \geq 1,\ s \in S.$$

Uit (49) volgt dat voor elke $f \in F$ geldt

(53) $$-\,C_n(s;f) = \sum_{t=0}^{n-1} E_{f^{(\infty)}}\{r(x_t,f(x_t))\,|\,x_0 = s\} \quad \text{voor } n \geq 1,\ s \in S$$

en

(54) $$\lim_{n\to\infty} -\frac{1}{n}\, C_n(s;f) = G_{f^{(\infty)}}(s) = G_{f^{(\infty)}}(0) \qquad \text{voor alle } s \in S.$$

Met volledige inductie toont men eenvoudig aan dat voor alle $s \in 0$ geldt $C_n(s;f) \leq C_{n-1}(0;f)$, $n \geq 2$. Hieruit volgt

(55) $$-\,C_n(s;f) \geq -\,C_n(0;f) \qquad \text{voor alle } s \neq 0,\ n \geq 1.$$

Op grond van (42), (43) en (49) kunnen wij een rij $\{h_k\}$, $k \geq 1$ met $h_k \in F$, construeren, zó dat

(56) $\quad G_{h_k(\infty)}(0) \geq -\frac{1}{5} - \frac{1}{2k}$ $\quad$ voor $k \geq 1$.

Stel ter afkorting

$$b(h_k) = G_{h_k(\infty)}(0).$$

Uit (54) volgt dat wij als volgt een rij indices L_1, L_2, ... kunnen construeren. Kies L_1 zó dat

(57) $\quad -\frac{1}{n}\, C_n(0;h_1) \geq b(h_1) - \frac{1}{2}$ $\quad$ voor $n \geq L_1$.

Gegeven $L_1, \ldots, L_{k-1}$, $k \geq 2$, wordt L_k zó gekozen dat

(58) $\quad \frac{1}{n}\{-(L_1+\ldots+L_{k-1}) - C_n(0;h_k)\} \geq b(h_k) - \frac{1}{2k}$ $\quad$ voor $n \geq L_k$.

Definieer nu strategie $R^* \in C_m$ als volgt: Zolang toestand 0 nog geen enkele keer aangenomen is, wordt actie 1 gevolgd. Vanaf het tijdstip dat toestand 0 voor het eerst wordt aangenomen, wordt als volgt te werk gegaan. De eerste L_1 tijdstippen wordt strategie $g_1^{(\infty)}$ gevolgd, de volgende L_2 tijdstippen wordt strategie $g_2^{(\infty)}$ gevolgd, de daarop volgende L_3 tijdstippen wordt strategie $g_3^{(\infty)}$ gevolgd, etc.

Aangezien $\mu_{s0}(f) < \infty$ voor alle $s \in S$ en $f \in F$, geldt voor elke begintoestand $s \neq 0$ dat het systeem met kans 1 na een eindige tijd toestand 0 bereikt, als strategie R^* wordt gevolgd. De opbrengst in deze tijd is gelijk aan nul. Het zal nu duidelijk zijn dat

(59) $\quad G_{R^*}(s) = G_{R^*}(0)$ $\quad$ voor alle $s \quad S$.

Zij $n_k = L_1 + \ldots + L_k$. Uit de structuur van de opbrengstfunktie $r(s,a)$, de definitie van R^*, (53), (55), (58) en (56) volgt dat

(60) $$\frac{1}{n_k}\sum_{t=0}^{n_k-1} E_{R^*}\{r(x_t,y_t)\,|\,x_0=0\} \geq \frac{1}{n_k}\{-(L_1+\ldots+L_{k-1})-C_n(0;h_k)\} \geq$$

$$\geq b(h_k) - \frac{1}{2k} \geq -\frac{1}{5} - \frac{1}{k} \qquad \text{voor } k \geq 1.$$

Op grond van (34) geldt

$$(61) \qquad G_R(0) = \limsup_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} E_{R^*}\{r(x_t,y_t)|x_0=0\} \le -\frac{1}{5}.$$

Stel $\{d_n\}$, $n \geq 1$, is een rij met getallen, zó dat $\limsup_{n\to\infty} d_n \leq d$ en voor elke $\varepsilon > 0$ geldt dat $d_n > d-\varepsilon$ voor oneindig veel indices n. Uit $\limsup_{n\to\infty} d_n \leq d$ volgt dat bij elke $\varepsilon > 0$ een $n_0(\varepsilon)$ te vinden is met $d_n \leq d+\varepsilon$ voor $n \geq n_0(\varepsilon)$. Dus d is een limietpunt van de rij $\{d_n\}$, hetgeen betekent dat $\limsup_{n\to\infty} d_n = d$.

Derhalve volgt uit (60) en (61) dat $G_R(0) = -\frac{1}{5}$. Hiermee is het bewijs van (36) voltooid.

3.2.4. Een vervangingsprobleem

Een vervangingsprobleem is een Markovbeslissingsprobleem waarin de toestandsruimte S een speciale toestand bevat, noem deze toestand 0, en de actieruimte A een speciale actie bevat, noem deze actie a^0, zó dat

$$q(0|s,a^0) = 1 \qquad \text{voor alle } s \in S. \tag{62}$$

Als in toestand s de actie a genomen wordt, dan worden directe kosten c(s,a) verkregen. Wij nemen aan dat de actieruimte A eindig is. Voorts nemen wij aan dat een constante M bestaat, zó dat

$$0 \leq c(s,a) \leq M \qquad \text{voor alle } s \in S,\ a \in A. \tag{63}$$

Definieer voor elke $R \in C$

$$\phi(s;R) = \liminf_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} E_R\{c(x_t,y_t)|x_0=s\}, \qquad s \in S, \tag{64}$$

en

$$\Phi(s;R) = \limsup_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} E_R\{c(x_t,y_t)|x_0=s\}, \qquad s \in S. \tag{65}$$

Een strategie $R_0 \in C$ heet optimaal voor het vervangingsprobleem als

$$\phi(0;R) = \min_{R\in C} \phi(0;R). \tag{66}$$

Wij eisen dus alleen dat een strategie optimaal is voor de begintoestand 0.

Wij zullen nu bewijzen dat voor het vervangingsprobleem altijd een optimale Markovstrategie bestaat. Daartoe eerst enige lemmas.

<u>Lemma 2</u>.

$$\inf_{R\in C} \phi(0;R) = \inf_{R\in C} \Phi(0;R). \tag{67}$$

Bewijs

Stel

$$g = \inf_{R \in C} \phi(0;R) \quad \text{en} \quad G = \inf_{R \in C} \Phi(0;R).$$

Uiteraard geldt $g \leq G$. Zij $\varepsilon > 0$ vast en willekeurig gekozen. Op grond van de stelling van Derman en Strauch (zie ook (2)) is $g = \inf\{\phi(0;R) | R \in C_{rm}\}$. Kies $R^* = (\pi_0^*, \pi_1^*, \ldots) \in C_{rm}$ zó dat

$$\phi(0;R^*) \leq g + \varepsilon/2, \tag{68}$$

en kies $N \geq 1$ zó dat

$$\frac{1}{N+1}\{\sum_{t=0}^{N-1} E_{R^*}(c(x_t,y_t)|x_0=0) + M\} \leq \phi(0;R^*) + \varepsilon/2. \tag{69}$$

Definieer nu strategie R^{**} als $R^{**} = (\pi_0^*, \ldots, \pi_{N-1}^*, f^0, \pi_0^*, \ldots, \pi_{N-1}^*, f^0, \ldots, f^0, \pi_0^*, \ldots, \pi_N^*, f^0, \ldots)$, waarbij $f^0(s) = a^0$ voor alle $s \in S$. Dus op de eerste N tijdstippen gaat R^{**} op precies dezelfde wijze te werk als R^*, daarna wordt actie a^0 genomen, op de volgende N tijdstippen gaat R^{**} weer op precies dezelfde wijze te werk als R^* op $t = 0, \ldots, N-1$, daarna wordt actie a^0 genomen, etc. Men gaat eenvoudig na dat

$$\Phi(0;R^{**}) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} E_{R^{**}}(c(x_t,y_t)|x_0=0) = \tag{70}$$

$$= \frac{1}{N+1}\{\sum_{t=0}^{N-1} E_{R^*}(c(x_t,y_t)|x_0=0) + E_{R^*}(c(x_N,a^0)|x_0=0)\}.$$

Uit (70), (63), (69) en (68) volgt dat

$$\Phi(0;R^{**}) \leq \phi(0;R^*) + \varepsilon/2 \leq g+\varepsilon. \tag{71}$$

Aangezien $G \leq \Phi(0;R^*)$ en ε willekeurig gekozen is, volgt uit (71) dat $G \leq g$ is. Per definitie geldt $G \geq g$. Dus $G = g$. Hiermee is het bewijs voltooid.

Opmerking

Uit lemma 2 volgt direct dat een strategie optimaal met betrekking tot $\phi(0;R)$ ook optimaal is met betrekking tot $\Phi(0;R)$, en omgekeerd.

Op grond van de resultaten van paragraaf 3.1 (de stellingen 3, 4 en 6) bestaat voor elke $\beta \in (0,1)$

(72) $$V_\beta(s) = \min_{R\in C} V_{\beta,R}(s) \qquad \text{voor alle } s \in S$$

en is V_β de unieke oplossing van de functionaalvergelijking

(73) $$V_\beta(s) = \min_{a\in A}\{c(s,a) + \beta p(s,a)V_\beta\}, \qquad s \in S.$$

Voorts geldt dat elke strategie $f^{(\infty)}$, zó dat voor elke $s \in S$ de actie $f(s)$ het rechterlid van (73) minimaliseert, optimaal is m.b.t. $V_{\beta,R}(s)$.

Lemma 3.

Voor elke $\beta \in (0,1)$ geldt

(74) $$V_\beta(s) \leq M + V_\beta(0) \qquad \text{voor alle } s \in S.$$

Bewijs

Uit (73), (62), (63) en het feit dat $V_\beta(0) \geq 0$ is, volgt direct dat $V_\beta(s) \leq c(s,a^0) + \beta V_\beta(0) \leq M + V_\beta(0)$ voor alle $s \in S$.

Lemma 4.

Zij $\sum_{n=0}^\infty a_n x^n$ een machtreeks met reële coëfficienten die absoluut convergeert voor $0 < x < 1$. Dan geldt

$$\limsup_{x\uparrow 1} (1-x) \sum_{n=0}^\infty a_n x^n \leq \limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^n a_k.$$

Bewijs

Stel ter afkorting

$$s_n = \sum_{k=0}^{n-1} a_k, \quad s = \limsup_{n\to\infty} \frac{s_n}{n} \text{ en } A(x) = \sum_{n=0}^\infty a_n x^n.$$

Aangezien $\sum_{n=0}^{\infty} a_n x^n$ en $\sum_{n=0}^{\infty} x^n$ beiden absoluut convergeren voor $0 < x < 1$, geldt

$$\frac{A(x)}{1-x} = \sum_{n=0}^{\infty} a_n x^n \sum_{m=0}^{\infty} x^m =$$

$$= \sum_{n=0}^{\infty} s_n x^n \qquad \text{voor } 0 < x < 1.$$

Zij $\varepsilon > 0$ vast en willekeurig gekozen. Er bestaat een natuurlijk getal N zó dat $s_n/(n+1) \leq s + \varepsilon/2$ voor alle $n \geq N$. Stel $K = \max\{s_n-(n+1)s \mid n=0,1,\ldots,N\}$. Aangezien $1/(1-x)^2 = \sum_{n=0}^{\infty} (n+1)x^n$, geldt

$$\frac{A(x)}{1-x} - \frac{s}{(1-x)^2} = \sum_{n=0}^{N} \{s_n-(n+1)s\}x^n + \sum_{n=N+1}^{\infty} \{s_n-(n+1)s\}x^n$$

$$\leq K\,\frac{1-x^{N+1}}{1-x} + \frac{\varepsilon}{2}\,\frac{1}{(1-x)^2} \qquad \text{voor } 0 < x < 1.$$

Hieruit volgt

$$(1-x)A(x) - s \leq K(1-x)(1-x^{N+1}) + \varepsilon/2.$$

Dus $(1-x)A(x) \leq s + \varepsilon$ voor x dicht genoeg bij 1. Hiermee is het lemma bewezen, aangezien ε willekeurig gekozen is.

Lemma 5.

Laat voor elke $\beta \in (0,1)$ de stationaire strategie $f_\beta^{(\infty)}$ zodanig zijn, dat voor elke $s \in S$ de actie $f_\beta(s)$ het rechterlid van (73) minimaliseert. Dan geldt

$$\lim_{\beta \uparrow 1} \phi(0;f_\beta^{(\infty)}) = \inf_{R \in C} \phi(0;R).$$

Bewijs

Uit de definitie van f_β volgt dat

$$(75) \qquad V_\beta(s) = c(s,f_\beta(s)) + \beta p(s,f_\beta(s))V_\beta \qquad \text{voor alle } s \in S.$$

Voor alle $n \geq 1$ geldt de identiteit

(76) $$E_{f_\beta^{(\infty)}}\Big[\sum_{t=1}^{n}\{V_\beta(x_t)-E_{f_\beta^{(\infty)}}(V_\beta(x_t)\,|\,x_0,y_0,\ldots,x_{t-1},y_{t-1})\}\Big] = 0.$$

Voorts geldt (gebruik (75))

(77) $$E_{f_\beta^{(\infty)}}\{V_\beta(x_t)\,|\,x_0=s_0,y_0=f_\beta(s_0),\ldots,x_{t-1}=s_{t-1},y_{t-1}=f_\beta(s_{t-1})\} =$$

$$= p(s_{t-1},f_\beta(s_{t-1}))V_\beta =$$

$$= \beta p(s_{t-1},f_\beta(s_{t-1}))V_\beta + c(s_{t-1},f_\beta(s_{t-1})) +$$

$$- c(s_{t-1},f_\beta(s_{t-1})) + (1-\beta)p(s_{t-1},f_\beta(s_{t-1}))V_\beta =$$

$$= V_\beta(s_{t-1}) - c(s_{t-1},f_\beta(s_{t-1})) + (1-\beta)p(s_{t-1},f_\beta(s_{t-1}))V_\beta.$$

Uit (76) en (77) volgt dat voor alle $n \geq 1$ geldt

(78) $$E_{f_\beta^{(\infty)}}\{V_\beta(x_0)-V_\beta(x_n)\} + (1-\beta)\sum_{t=1}^{n} E_{f_\beta^{(\infty)}}\{p(x_{t-1},f_\beta(x_{t-1}))V_\beta\} =$$

$$= \sum_{t=1}^{n} E_{f_\beta^{(\infty)}}\, c(x_{t-1},f_\beta(x_{t-1})).$$

Uit (78) en lemma 3 volgt dat voor alle $n \geq 1$ geldt

(79) $$\frac{1}{n}\sum_{t=1}^{n} E_{f_\beta^{(\infty)}}\, c(x_{t-1},f_\beta(x_{t-1})) \leq (1-\beta)(V_\beta(0)+M) +$$

$$- \frac{1}{n} E_{f_\beta^{(\infty)}}\{V_\beta(x_n)-V_\beta(x_0)\}.$$

Aangezien $|V_\beta(s)| \leq M/(1-\beta)$ voor alle $s \in S$, vinden wij door $n \to \infty$ te laten gaan in (79), dat

(80) $$\phi(s;f_\beta^{(\infty)}) \leq (1-\beta)V_\beta(0) + (1-\beta)M \quad \text{voor alle } s \in S.$$

Aangezien $V_\beta(s) \leq V_{\beta,R}(s)$ voor alle $R \in C$, $s \in S$, volgt door toepassing van lemma 4 dat

$$\text{(81)} \qquad \limsup_{\beta \uparrow 1} (1-\beta)V_\beta(s) \leq \limsup_{\beta \uparrow 1} (1-\beta)V_{\beta,R}(s) \leq$$

$$\leq \Phi(s;R) \qquad \text{voor alle } R \in C,\ s \in S.$$

Uit (80) en (81) volgt dat

$$\text{(83)} \qquad \limsup_{\beta \uparrow 1} \phi(0;f_\beta^{(\infty)}) \leq \Phi(0;R) \qquad \text{voor alle } R \in C.$$

Uit (83) en lemma 2 volgt dat

$$\text{(84)} \qquad \limsup_{\beta \uparrow 1} \phi(0;f_\beta^{(\infty)}) \leq \inf_{R \in C} \Phi(0;R) = \inf_{R \in C} \phi(0;R).$$

Aangezien $\phi(0;f_\beta^{(\infty)}) \geq \inf_{R \in C} \phi(0;R)$ voor alle $\beta \in (0,1)$, volgt uit (84) de bewering van de stelling.

Stelling 3.

Er bestaat een optimale Markovstrategie.

Bewijs

Op grond van lemma 5 kunnen wij bij elk natuurlijk getal n een stationaire strategie $f_n^{(\infty)}$ kiezen, zó dat

$$\text{(85)} \qquad \phi(0;f_n^{(\infty)}) \leq g + \frac{1}{2n},$$

waarbij

$$g = \inf_{R \in C} \phi(0;R).$$

Kies N_1 zó dat

$$\frac{1}{N_1+1}\left[\sum_{t=0}^{N_1-1} E_{f_1^{(\infty)}}(c(x_t,y_t)\,|\,x_0=0)+M\right] \leq \phi(0;f_1^{(\infty)}) + \frac{1}{2}.$$

Wij definieren de natuurlijke getallen N_1, N_2, ..., op de volgende recursieve wijze. Gegeven N_1, ..., N_{i-1}, kies N_i zó dat

$$(86)\qquad \frac{1}{N_1+\ldots+N_i+i}\,[\sum_{t=0}^{N_i-1} E_{f_i^{(\infty)}}(c(x_t,y_t)\,|\,x_0=0) + M(N_1+\ldots+N_{i-1}+i)] \le$$

$$\le \phi(0;f_i^{(\infty)}) + \frac{1}{2i} \qquad\qquad i \ge 1.$$

Laat de Markovstrategie R^* als volgt gedefinieerd zijn:

volg $f_1^{(\infty)}$ op $t = 0,\ldots,N_1-1$, neem op $t = N_1$ actie a^0,

volg $f_2^{(\infty)}$ op de volgende N_2 tijdstippen, neem op $t=N_1+N_2+1$ actie a^0,

.
.
.

volg $f_i^{(\infty)}$ op de volgende N_i tijdstippen, neem daarna actie a^0,

etc.

Stel

$$m_i = N_1 + \ldots + N_i + i, \qquad\qquad i \ge 1,$$

dan volgt met behulp van (62), (63), (85) en (86) dat

$$\frac{1}{m_i}\sum_{t=0}^{m_i-1} E_{R^*}(c(x_t,y_t)\,|\,x_0=0) \le \frac{1}{m_i}\,[M(N_1+\ldots+N_{i-1}+i) +$$

$$+ \sum_{t=0}^{N_i-1} E_{f_i^{(\infty)}}\,(c(x_t,y_t)\,|\,x_0=0) \le \phi(0;f_i^{(\infty)}) + \frac{1}{2i} \le$$

$$\le g + \frac{1}{i} \quad \text{voor } i \ge 1.$$

Uit deze ongelijkheid volgt $\phi(0;R^*) \le g$. Hiermee is de stelling bewezen, aangezien per definitie $g \le \phi(0;R^*)$ is.

3.2.5. Een (s,S) strategie voor een voorraadmodel met nalevering.

(I) Model formulering.

Wij beschouwen een voorraadprobleem waarin de behoeften ξ_1, ξ_2,... aan een bepaald produkt in de perioden 1,2,... onafhankelijke, niet-negatieve stochastische variabelen zijn met eenzelfde kansverdeling

$$p_j = P(\xi_t = j), \qquad j \geq 0;\ t \geq 1.$$

Veronderstel dat

$$\mu = \sum_{j=1}^{\infty} j\, p_j < \infty \quad \text{en} \quad p_0 < 1. \tag{87}$$

De voorraad kan alleen worden aangevuld aan het begin van elke periode. De levertijd wordt nul verondersteld. De vraag wordt verondersteld plaats te vinden aan het eind van elke periode. Als de voorraad ontoereikend is voor de vraag, dan wordt nageleverd. De voorraad kan dus ook negatief zijn (achterstand!)

De volgende kosten treden in dit model op. De bestelkosten van i eenheden bedragen $K\delta(i) + ci$, $i = 0,1,\ldots$, waarin $K \geq 0$, $\delta(0) = 0$, en $\delta(i) = 1$ voor $i \geq 1$. De voorraad- en naleveringskosten in een periode bedragen $L(k)$, als k de voorraad is vlak na een eventuele aanvulling in die periode.

De volgende aannames worden voor de funktie $L(k)$ gemaakt.

(i) Er bestaat een geheel getal S_0 zodat $L(i) \leq L(j)$ voor $j \leq i \leq S_0$ en $L(i) \geq L(j)$ voor $i \geq j \geq S_0$

(ii) $\lim\limits_{|k|\to\infty} L(k) > L(S_0) + K$.

Op grond van (ii) mogen wij aannemen dat S_0 het kleinste gehele getal is waarvoor (i) geldt.

Definieer s_1 als het kleinste gehele getal, waarvoor

$$L(s_1) \leq L(S_0) + (1-p_0)K \tag{88}$$

en definieer S_1 als het grootste gehele getal, waarvoor

$$L(S_1) \leq L(S_0) + K \tag{89}$$

Aan de keuze van de bestelgrootte is de volgende zwakke voorwaarde opgelegd. Er zijn eindige getallen $u \leq s_1$ en $U \geq S_1$ gegeven, zó dat geen order wordt geplaatst als de voorraad $i \geq U$ is, de orderomvang tenhoogste $U - i$ is als de voorraad $i < U$ is, en de orderomvang tenminste $u - i$ is als de voorraad $i < u$ is.

Definieer de toestand i van het systeem als de voorraad vlak voor beslissen. Als toestandsruimte I nemen wij de verzameling van alle gehele getallen. Elke actie k identificeren wij met de voorraad vlak na beslissen. Laat $A(i)$ de verzameling zijn van toegelaten acties in toestand i, dan geldt

$$A(i) = \begin{cases} \{i\} & \text{voor } i \geq U \\ \{k \mid \max(i,u) \leq k \leq U\} & \text{voor } i < U. \end{cases} \tag{90}$$

Laten i_t en k_t de toestand en de actie in de periode $t(t \geq 1)$ voorstellen. Wij kunnen stellen dat direkte kosten $K\delta(k-i) + (k-i)c + L(k)$ gemaakt worden, als in toestand i de actie k genomen wordt.

Onder de voorwaarde $i_1 = i$ geldt $u \leq k_t \leq \max(i,U)$, $t \geq 1$. Uit deze betrekking, de relatie $i_{t+1} = k_t - \xi_t$, $t \geq 1$, en $E\xi_t = \mu < \infty$, $t \geq 1$, volgt dat $E_R\{K\delta(k_t-i_t) + (k_t-i_t)c + L(k_t)\}$ bestaat voor alle $R \in C$, $t \geq 1$.

Het optimaliteitskriterium is

$$\phi(i;R) = \liminf_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} E_R\{K\delta(k_t-i_t) + (k_t-i_t)c + L(k_t) \mid i_1=i\}.$$

Door gebruik te maken van $i_{t+1} = k_t - \xi_t$, en de begrensdheid van de rij $\{E_R(k_n \mid i_1=i\}$, $n \geq 1$, vinden wij dat uit de relatie

$$\sum_{t=1}^{n} E_R\{(k_t-i_t)|i_1=i\} =$$

$$= -i + E_R(k_1|i_1=i) + \sum_{t=2}^{n} E_R\{(k_t-k_{t-1}+\xi_{t-1})|i_1=i\} =$$

$$= -i + (n-1)\mu + E_R(k_1|i_1=i) + E_R(k_n|i_1=i),\ n \geq 2$$

volgt dat

$$\phi(i;R) = \liminf_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E_R\{K\delta(k_t-i_t) + L(k_t)|i_1=i\} + \mu c\ .$$

Aangezien de term μc niet van de strategie R afhangt, mogen wij zonder bezwaar aannemen dat $c = 0$. Dus wij mogen $\phi(i;R)$ herdefinieren als

$$(91)\qquad \phi(i;R) = \liminf_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E_R\{K\delta(k_t-i_t) + L(k_t)|i_1=i\},\quad i \in I;\ R \in C.$$

Een strategie $R^* \in C$ heet optimaal als $\phi(i;R^*) \leq \phi(i;R)$ voor alle $i \in I$, $R \in C$.

Wij zullen in deel (IV) van de sectie 3.2.5 bewijzen dat een optimale stationaire strategie bestaat en dat tenminste één optimale stationaire strategie een (s,S) voorraadstrategie is. Het bewijs berust op stelling 1 (blz. 107) en de resultaten die wij in deel (II) en deel (III) van deze sectie zullen afleiden voor de (s,S) voorraadstrategieën.

(II) De verwachte gemiddelde kosten per tijdseenheid voor een (s,S) strategie.

Een (s,S) strategie ($s \leq S$) is een stationaire strategie van de volgende vorm: "als de voorraad aan het begin van een periode kleiner dan s is, vul de voorraad dan aan tot S; is de voorraad groter dan of gelijk aan s, plaats dan geen order".

Alvorens wij de kriteriumfunktie (91) voor de (s,S) strategie kunnen berekenen, moeten wij eerst enig voorbereidend werk verrichten. Eerst zullen wij enige resultaten geven, die welbekend zijn uit de "renewal theory", en daarna zullen wij de Markovketen $\{i_t\}$, $t \geq 1$, behorende bij een (s,S) strategie analyseren.

De behoeften in de perioden 1,2,... zijn onderling onafhankelijke, niet-negatieve en identiek verdeelde stochastische variabelen $\xi_1,\xi_2,\ldots$ met $p(j) = P(\xi_t=j)$, $j \geq 0$. Hierbij is $p(0) < 1$. Definieer

$$(92)\qquad p^{(n)}(j) = P(\xi_1+\ldots+\xi_n=j) \text{ en } F^{(n)}(j) = P(\xi_1+\ldots+\xi_n\leq j),$$

$$j \geq 0;\ n \geq 1.$$

Uit deze definitie volgt direct dat

$$(93)\qquad p^{(n)}(j) = \sum_{k=0}^{j} p(j-k)\, p^{(n-1)}(k), \qquad j \geq 0;\ n \geq 1,$$

waarin $p^{(0)}(0) = 1$ en $p^{(0)}(j) = 0$ voor $j \geq 1$.
Voor elke $j \geq 0$ geldt

$$(94)\qquad \lim_{n\to\infty} F^{(n)}(j) = 0.$$

Dit kan als volgt bewezen worden. Aangezien $p(0) < 1$ is, is bij elke $j \geq 0$ een natuurlijk getal r te vinden met $F^{(r)}(j) < 1$. De gebeurtenis $\{\xi_1+\ldots+\xi_n \leq j\}$ impliceert het simultane optreden van de gebeurtenissen $\{\xi_k \leq j\}$, $k = 1,\ldots,n$. Dus $F^{(rm)}(j) \leq [F^{(r)}(j)]^m$, $m \geq 1$. Hieruit volgt dat de monotoon niet-stijgende rij $\{F^{(n)}(j)\}$, $n \geq 1$, als limiet 0 heeft.

Definieer

$$(95)\qquad m(j) = \sum_{n=1}^{\infty} p^{(n)}(j) \quad \text{en} \quad M(j) = \sum_{n=1}^{\infty} F^{(n)}(j), \quad j \geq 0.$$

Uit deze definitie volgt direct dat

$$(96)\qquad M(j) = \sum_{k=0}^{j} m(k), \qquad j \geq 0.$$

Zonder bewijs vermelden wij dat $M(k)$ eindig is voor alle $k \geq 0$ en dat *)

---

*) Zie bijv. N. Prabhu, Stochastic Processes.

$$(97) \qquad \lim_{k\to\infty} \frac{M(k)}{k} = \frac{1}{\mu} .$$

Definieer de stochastische variabelen

$$T_0 = 0 \text{ en } T_n = \xi_1+\ldots+\xi_n, \; n \geq 1$$

en definieer voor elke $k \geq 0$ de stochastische variabele

$$(98) \qquad N(k) = \max\{n | n \geq 0, \; T_n \leq k\}.$$

Dus $1 + EN(k)$ is het verwachte aantal perioden nodig voor een totale vraag groter dan k. Er geldt

$$(99) \qquad EN(k) = M(k) \qquad \text{voor } k \geq 0.$$

Dit kan als volgt bewezen worden. Definieer bij k vast, $Z_n = 1$ als $T_n \leq k$ en $Z_n = 0$ als $T_n > k$. Dan geldt $EN(k) = E(\sum_{n=1}^{\infty} Z_n) = \sum_{n=1}^{\infty} F^{(n)}(k) =$
$= M(k)$, $k \geq 0$.
Definieer voor elke $k \geq 0$ de stochastische variabele

$$\gamma(k) = T_{N(k)+1} - k .$$

Er geldt voor elke $k \geq 0$ dat

$$(100) \qquad P(\gamma(k) \leq j) = P(k < T_1 \leq k+j) + \sum_{n=1}^{\infty} \sum_{h=0}^{k} P(T_n=h, \; k < T_{n+1} \leq k+j)$$
$$= F(k+j) - F(k) + \sum_{h=0}^{k} \{F(k+j-h) - F(k-h)\} \, m(h), \; j \geq 1$$

waarin $F(j) = F^{(1)}(j) = p(0) + \ldots + p(j)$, $j \geq 0$.

Stel dat een (s,S) strategie wordt toegepast ($s \leq S$, s en S vast). Het stochastische proces $\{i_t\}$, $t \geq 1$, is dan een aftelbare Markov-keten, waarvan de een-stapsovergangskansen p_{ij}, $i, j \in I$, gegeven worden door

$$(101) \qquad p_{ij} = \begin{cases} p(i-j) & \text{voor } i \geq s, \; j \leq i, \\ p(S-j) & \text{voor } i < s, \; j \leq S, \\ 0 & \text{anders.} \end{cases}$$

De Markov-keten $\{i_t\}$, $t \geq 1$, bevat geen twee of meer disjuncte fuiken, omdat vanuit elke toestand de toestanden $j \leq S$ met $p(S-j) > 0$ bereikbaar zijn. Elke toestand j met $p(S-j) > 0$ is essentieel. Dus de Markov-keten $\{i_t\}$, $t \geq 1$, bevat precies één kernfuik K (stelling 1.2.6). Laat m het grootste getal $j \leq S$ zijn met $p(S-j) > 0$. Wij zullen nu aantonen dat m een positieve terugkeertoestand is. Laat $f_{ij} = P(i_t = j$ voor één of andere $t \geq 2 | i_1 = i)$. Het zal direct duidelijk zijn dat voor het geval $m \leq s$ is, geldt

$$(102) \qquad 1 - f_{mm} = \lim_{t\to\infty} \{1-p(S-m)\}^t = 0.$$

Door gebruik te maken van het feit dat de stochastische variabele $N(S-s)$ met kans 1 eindig is (zie (98) en (99), ziet men eenvoudig in dat relatie (102) ook van toepassing is voor het geval $m \geq s$ is. Dus toestand m is een terugkeertoestand. Laten N_1, N_2,... onafhankelijke stochastische variabelen zijn die eenzelfde verdeling bezitten als $1 + N(S-s)$, en laat de stochastische variabele V de geometrische verdeling

$$P(V=k) = p(S-j)\ \{1 - p(S-j)\}^k, \qquad k \geq 0$$

bezitten. Het zal na enig nadenken duidelijk zijn dat voor de verwachte terugkeertijd μ_{mm} van toestand m geldt

$$(103) \qquad \mu_{mm} = \begin{cases} 1/p(S-m) & \text{als } m < s \\ E\left(\sum_{k=1}^{V} N_k\right) + 1 & \text{als } m \geq s. \end{cases}$$

Een toepassing van de identiteit van Wald *) leert ons dat

$$E(N_1+\ldots+N_V) = EV\{1 + EN(S-s)\} =$$
$$= \{1 + M(S-s)\}/p(S-m) < \infty.$$

*) Zie b.v. W. Feller, Introduction to the Probability Theory and its Applications, Volume II, p. 567.

Uit (103) en op de vorige blz. staande relatie volgt nu dat $\mu_{mm} < \infty$. Dus m is een positieve terugkeertoestand. Stelling 1.2.15 leert ons vervolgens dat alle toestanden van de kernfuik K positieve terugkeertoestanden zijn.

Laten $p_{ij}^{(k)}$ de k-stapsovergangskansen van de Markov-keten $\{i_t\}$, $t \geq 1$, zijn en laat (zie ook stelling 1.2.12),

$$p_{ij}^{*} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)} \qquad i,j \in I.$$

Voor de getallen p_{ij}^{*} geldt

$$p_{ij}^{*} = p_{jj}^{*} \qquad \text{voor alle } i,j \in I. \tag{104}$$

Dit volgt eenvoudig uit de resultaten van paragraaf 1.2. Aangezien elke toestand $j \notin K$ inessentieel is, volgt direct uit stelling 1.2.14 en stelling 1.2.12(c) dat (104) juist is voor $j \notin K$. Uit stelling 1.2.12(b) en stelling 1.2.18 volgt dat (104) juist is voor $i,j \in K$. Rest nog om (104) aan te tonen voor $i \notin K$, $j \in K$. Daartoe merken wij op dat uit het bewijs van stelling 1.2.21 volgt dat $f_{ij} = f_{iK}$, $i \notin K$, $j \in K$, waarbij f_{iK} de kans is om uitgaande van toestand i ooit in de kernfuik K terecht te komen. Als $i_1 = i < s$ is, dan geldt $f_{iK} = P(i_2 \in K | i_1=i) = 1$, en als $i \geq s$ dan volgt met behulp van (94) dat

$$1 - f_{iK} = P\{\xi_1+\ldots+\xi_n \leq i-s \text{ voor alle } n \geq 1\} =$$

$$= \lim_{n\to\infty} F^{(n)}(i-s) = 0.$$

Hiermee is de relatie (104) geverifieerd voor alle i en j.

Aangezien de Markov-keten $\{i_t\}$, $t \geq 1$, precies één kernfuik bezit en deze kernfuik uit positieve terugkeertoestanden bestaat, is $(p_{jj}^{*}, j \in I)$ de unieke invariante kansverdeling van de Markov-keten (zie stelling 1.2.22 en de bijbehorende opmerking op blz. 37). De kansverdeling $(p_{jj}^{*}, j \in I)$ is de unieke oplossing van het stelsel

$$(105)\qquad \sum_{j\in I} p^*_{jj} = 1 \quad ; \quad p^*_{jj} = \sum_{k\in I} p_{kj}\, p^*_{kk}, \quad j \in I.$$

Stel ter afkorting

$$p^*_j = p^*_{jj}, \qquad\qquad j \in I.$$

Elke toestand $j > S$ is inessentieel, dus (zie stelling 1.2.14 en stelling 1.2.12(c))

$$(106)\qquad p^*_j = 0 \qquad\qquad j > S.$$

Uit (101), (105) en (106) volgt dat

$$(107)\qquad p^*_j = \sum_{k=-\infty}^{s-1} p(S-j)p^*_k + \sum_{k=j}^{S} p(k-j)p^*_k \quad \text{voor } s \leq j \leq S$$

$$(108)\qquad p^*_j = \sum_{k=-\infty}^{s-1} p(S-j)p^*_k + \sum_{k=s}^{S} p(k-j)p^*_k \quad \text{voor } j < s \quad .$$

De vergelijking (107) kunnen wij eenvoudig tot de standaardvorm van de "renewal equation" herleiden. Voeren wij de variabelen $S-j = h$ en $m = S-k$ in, en stellen wij

$$(109)\qquad p^*_{S-i} = p^{**}_i \qquad\qquad i \leq S,$$

dan kunnen wij (107) in de equivalente vorm

$$(110)\qquad p^{**}_h = Q^*_s\, p(h) + \sum_{m=0}^{h} p(h-m)p^{**}_m \qquad \text{voor } 0 \leq h \leq S-s$$

schrijven, waarin

$$(111)\qquad Q^*_s = \sum_{k=-\infty}^{s-1} p^*_k.$$

Passen wij de gelijkheid (110) herhaald toe, dan vinden wij met behulp van (93) dat voor elke $n \geq 1$ geldt

$$(112)\qquad p^{**}_h = Q^*_s \sum_{t=1}^{n} p^{(t)}(h) + \sum_{m=0}^{h} p^{(n)}(h-m)p^{**}_m \quad \text{voor } 0 \leq h \leq S-s.$$

Laten wij nu $n \to \infty$ gaan in (112), dan vinden wij met behulp van (94) dat

(113) $$p_h^{**} = Q_s^* \, m(h) \qquad \text{voor } 0 \le h \le S-s.$$

Uit (109) en (113) volgt dat

(114) $$p_j^* = Q_s^* \, m(S-j) \qquad \text{voor } s \le j \le S.$$

Uit (105) en (106) volgt dat

(115) $$\sum_{j=s}^{S} p_j^* = 1-Q_s^*.$$

Uit (114) en (115) volgt nu dat

(116) $$Q_s^* = 1/\{1 + M(S-s)\}.$$

Uit (106), (108), (114) en (116) volgt nu dat de invariante kansverdeling $(p_j^*, \, j \in I)$ gegeven wordt door

(117) $$p_j^* = \begin{cases} 0 \, , \\ m(S-j)/\{1 + M(S-s)\}, & s \le j \le S, \\ \{m(S-j) + \sum_{k=0}^{S-s} p(S-j-k) \, m(k)\}/\{1 + M(S-s)\}, & j < s. \end{cases}$$

Definieer

(118) $$w_i = \begin{cases} L(i) & \text{voor } i \ge s \\ K+L(S) & \text{voor } i < s. \end{cases}$$

Als de begintoestand i is en de (s,S) strategie wordt gevolgd, dan kunnen alleen toestanden j met $j \le \max(i,S)$ worden aangenomen. Voor de strategie $R = (s,S)$ geldt derhalve

(119) $$\frac{1}{n} \sum_{t=1}^{n} E_R\{K|(k_t-i_t) + L(k_t)|i_1=i\} = \frac{1}{n} \sum_{t=1}^{n} \sum_{j=-\infty}^{\max(i,S)} p_{ij}^{(t)} w_j =$$

$$= w_{s-1} \sum_{j=-\infty}^{s-1} \frac{1}{n} \sum_{t=1}^{n} p_{ij}^{(t)} + \sum_{j=s}^{\max(i,S)} w_j \frac{1}{n} \sum_{t=1}^{n} p_{ij}^{(t)}, \qquad i \in I.$$

Uit $\sum_{j=-\infty}^{\max(i,S)} \frac{1}{n} \sum_{t=1}^{n} p_{ij}^{(t)} = 1$ voor alle $n \geq 1$, $i \in I$, en (105) en (106) volgt

$$(120) \qquad \lim_{n\to\infty} \sum_{j=-\infty}^{\max(i,S)} \frac{1}{n} \sum_{t=1}^{n} p_{ij}^{(t)} = 1 = \sum_{j=-\infty}^{S} p_j^*, \qquad i \in I.$$

Uit (104) en (106) volgt dat

$$(121) \qquad \lim_{n\to\infty} \sum_{j=s}^{\max(i,S)} \frac{1}{n} \sum_{t=1}^{n} p_{ij}^{(t)} = \sum_{j=s}^{S} p_j^*, \qquad i \in I.$$

Uit (120) en (121) volgt nu dat

$$(122) \qquad \lim_{n\to\infty} \sum_{j=-\infty}^{s-1} \frac{1}{n} \sum_{t=1}^{n} p_{ij}^{(t)} = \sum_{j=-\infty}^{s-1} p_j^*, \qquad i \in I.$$

Uit (119), (121) en (122) volgt nu dat

$$(123) \qquad \phi(i;(s,S)) = w_{s-1} \sum_{j=-\infty}^{s-1} p_j^* + \sum_{j=s}^{S} w_j p_j^* \quad \text{voor alle } i \in I.$$

Uit (111) en (116) t/m (118) volgt dat het rechterlid van (123) gelijk is aan

$$(124) \qquad g(s,S) \stackrel{\text{def}}{=} \{L(S) + \sum_{k=0}^{S-s} L(S-k)\, m(k) + K\}/\{1 + M(S-s)\}.$$

Wij merken op dat uit het voorgaande volgt dat wij voor een (s,S) strategie lim inf in (91) mogen vervangen door lim. Voorts vinden wij dat voor een (s,S) strategie de verwachte gemiddelde kosten per tijdseenheid niet van de begintoestand afhangen.

III <u>Enige eigenschappen van de funktie g(s,S)</u>.

Beschouw de funktie $g(s,S)$, waarbij de variabelen s en S beiden geheel zijn en $s \leq S$ is.

Wij merken eerst op dat uit (124) en de relatie $M(0) = p(0)/\{1-p(0)\}$. (zie (95)) volgt dat

(125) $g(S,S) = L(S) + \dfrac{K}{1+M(0)} = L(S) + (1-p(0))K$ voor alle $S \in I$.

Lemma 6.

Er bestaan gehele getallen s^* en S^*, $s^* \leq S^*$, zodat $g(s^*,S^*) \leq g(s,S)$ voor alle $s,S \in I$, $s \leq S$.

Bewijs.

Wij zullen aantonen dat een geheel getal B bestaat, zó dat

(126) $g(s,S) > L(S_0) + K$ voor alle $S > B$, $s \leq S$.

Geheel analoog aan het bewijs van het bestaan van zo'n getal B, kan bewezen worden dat een getal C bestaat, zó dat

(127) $g(s,S) > L(S_0) + K$ voor alle $s < C$, $S \geq s$.

Uit (126), (127) en de ongelijkheid $g(S_0,S_0) \leq L(S_0) + K$ (zie (125)) volgt dan het Lemma.

De eigenschappen (i) en (ii) van de funktie $L(k)$ (zie blz. 133) volgt dat een getal $\delta > 0$ bestaat en gehele getallen u_1 en u_2 met $u_1 < u_2$ bestaan, zó dat

(128) $L(j) \geq L(S_0) + K + \delta$ voor zowel $j \geq u_2$ als $j \leq u_1$.

Onderscheid de volgende gevallen

(1) $S \geq s > u_2$. Aangezien $K \geq 0$ is, volgt direct uit (124) en (128) dat $g(s,S) \geq L(S_0) + K + \delta > L(S_0) + K$.

(2) $S > u_2$, $u_1 \leq s \leq u_2$. Splits de sommatie van $\sum_{k=0}^{S-s} L(S-k)\,m(k)$ in $[0,S-u_2]$ en $[S-u_2+1, S-s]$. Uit (128) volgt dat $L(S-k) \geq L(S_0) + K + \delta$ voor $k \in [0,S-u_2]$ en voor $k \in [S-u_2+1,S-s]$ geldt $L(k) \geq L(S_0)$, omdat $L(k)$ het absolute minimum aanneemt in S_0. Door van deze ongelijkheden gebruik te maken, vinden wij dat

(129) $g(s,S) \geq L(S_0) + K + \delta + \{K+\delta\}\{M(S-u_2)-M(S-s)\}/\{1+M(S-s)\}$.

Aangezien de niet-negatieve funktie $M(k)$, $k \geq 0$, monotoon niet-dalend is en $u_1 \leq s \leq u_2$ is, geldt

$$|M(S-u_2) - M(S-s)|/\{1 + M(S-s)\} \leq \{M(S-u_1) - M(S-u_2)\}/\{1 + M(S-u_2)\}.$$

Uit (97) volgt dat een geheel getal $B_1 > u_2$ bestaat, zó dat het rechterlid van bovenstaande ongelijkheid $\leq \delta/2(K+\delta)$ is voor $S > B_1$. Uit (129) volgt nu dat $g(s,S) > L(S_0) + K$ voor alle $S > B_1$, $u_1 \leq s \leq u_2$.

(3) $S > u_2$, $s < u_1$. Splits nu de sommatie van $\sum_{k=0}^{S-s} L(S-k)\, m(k)$ in $[0,S-u_2]$, $[S-u_2+1,S-u_1]$, $[S-u_1+1,S-s]$. Op analoge wijze als in (2) kan het bestaan van een geheel getal B_2 worden aangetoond, zó dat $g(s,S) > L(S_0) + K$ voor $S > B_2$, $s < u_1$.

De relatie (126) geldt nu voor $B = \max(B_1,B_2)$. Hiermee is het bewijs voltooid.

Definieer

(130) $$g^* = \min\{g(s,S) | s,S \in I,\ s \leq S\}.$$

Wij merken op dat uit (124) en het feit dat $L(k)$ het absolute minimum aanneemt in S_0 direct volgt dat

(131) $$g^* = g(S_0,S_0) = L(S_0) \qquad \text{als } K = 0.$$

<u>Lemma 7</u>.

Laten s^* en S^* , $s^* \leq S^*$, gehele getallen zijn, zó dat $g(s^*,S^*) = g^*$

(a) als $m(S^*-s^*+1) > 0$, dan geldt $L(s^*-1) \geq g^*$

(b) als $s^* = S^*$, dan geldt $L(s^*) \leq g^*$

(c) als $s^* < S^*$ en als $m(S^*-s^*) > 0$, dan geldt $L(s^*) \leq g^*$

(d) als $p_1 > 0$, dan geldt $L(s^*-1) \geq g^* \geq L(s^*)$

(e) als $L(s^*-1) \geq g^* \geq L(s^*)$ en als $K > 0$ is, dan geldt $s_1 \leq s^* \leq S_0$.

<u>Bewijs</u>.

(a) Beschouw $g(s,S)$ als funktie van $\Delta = S-s$ en S. Stel $h(\Delta,S) = g(S-\Delta,S)$. De funktie $h(\Delta,S)$ is minimaal voor $\Delta = \Delta^* = S^*-s^*$ en $S = S^*$.

Dus $h(\Delta^*+1,S^*) - h(\Delta^*,S^*) \geq 0$. Deze ongelijkheid leidt na enig rekenwerk tot

$$m(\Delta^*+1)[L(s^*-1)(1+M(\Delta^*)) - \{L(S^*) + \sum_{k=0}^{\Delta^*} L(S^*-k)\, m(k) + K\}] \geq 0.$$

Omdat $m(\Delta^*+1) > 0$ is, volgt uit bovenstaande ongelijkheid en uit (124) dat $L(s^*-1) \geq g(s^*,S^*) = g^*$.

(b) Aangezien $s^* = S^*$, geldt $g^* = g(s^*,s^*) = L(s^*) + K/\{1+M(0)\} \geq L(s^*)$.

(c) Punt (c) volgt na enig rekenwerk uit $h(\Delta^*-1,S^*) - h(\Delta^*,S^*) \geq 0$ en $m(\Delta^*) > 0$.

(d) Als $p_1 > 0$ is, dan is $m(k) > 0$ voor alle $k \geq 1$. Punt (d) volgt nu uit (a), (b) en (c)

(e) Uit (125) en de aanname van punt (e) volgt dat $L(s^*) \leq g^* \leq g(S_0,S_0) = L(S_0) + (1-p(0))K$. Uit de definitie (88) van s_1 volgt nu dat $s_1 \leq s^*$ is. Stel $s^* > S_0$. Aangezien $L(k)$ monotoon niet-dalend is op $[S_0,\infty)$, volgt uit (124) en uit de aanname $K > 0$ dat $g^* = g(s^*,S^*) \geq L(s^*-1) + K/\{1+M(S^*-s^*)\} > L(s^*-1)$. Dit is in tegenspraak met $g^* \leq L(s^*-1)$. Dus $s^* \leq S_0$.

Lemma 8.

Er bestaan gehele getallen s^* en S^* zó dat $g(s^*,S^*) = g^*$ en $L(s^*-1) \geq g^* \geq L(s^*)$. Als $K = 0$, dan voldoen $s^* = S_0$ en $S^* = S_0$ aan deze voorwaarden.

Bewijs.

Uit Lemma 6 volgt het bestaan van gehele getallen s' en S' zó dat $g(s',S') = g^*$. Het bewijs van lemma 6 (zie (127)) leert ons dat $g(s,S') > g^*$ voor s voldoende klein. Derhalve bestaan gehele getallen s en S zó dat $g(s,S) = g^*$ en $m(S-s+1) > 0$. Uit lemma 7(a) volgt nu dat de verzameling $V = \{(s,S) | g(s,S) = g^* \leq L(s-1)\}$ niet-leeg is. Laat $(s^*,S^*) \in V$ zo gekozen zijn dat $S^*-s^* \leq S-s$ voor alle $(s,S) \in V$.

Wij zullen nu aantonen dat $g^* \geq L(s^*)$. Als $s^* = S^*$, dan volgt dit direct uit lemma 7(b). Beschouw nu het geval $s^* < S^*$. Stel dat $L(s^*) > g^*$. Uit lemma 7(c) volgt dan dat $m(S^*-s^*) = 0$. Vervolgens impliceert definitie (124) dat $g(s^*+1,S^*) = g^*$. Hieruit volgt dat $(s^*+1,S^*) \in V$, omdat $L(s^*) > g^*$. Dit is in tegenspraak met de keuze van (s^*,S^*). Dus $g^* \geq L(s^*)$.

Aangezien S_0 het kleinste gehele getal is, waarvoor $L(k)$ het absolute minimum aanneemt, volgt direct uit (131) dat voor het geval $K = 0$ geldt $g^* = L(S_0) < L(S_0-1)$. Hiermee is het bewijs voltooid.

IV <u>Het bestaan van een optimale (s,S) strategie</u>.

Laten s^* en S^* twee vast gekozen gehele getallen zijn, zó dat

(132) $$g(s^*,S^*) = g^* \qquad \text{en} \qquad L(s^*-1) \geq g^* \geq L(s^*).$$

Voor het geval $K = 0$, kiezen wij $s^* = S^* = S_0$. Bij deze keuze is voldaan aan (132) (zie lemma 8).

Uit lemma 7(e) volgt dat s^* voldoet aan

(133) $$s_1 \leq s^* \leq S_0 \quad .$$

Definieer de funktie $v^*(i)$, $i \in I$, als volgt

(134) $$v^*(i) = \begin{cases} 0 & \text{voor } i < s^*, \\ L(i) - a^* + \sum\limits_{j=0}^{i-s^*} v^*(i-j)\ p(j) & \text{voor } i \geq s^*. \end{cases}$$

<u>Opmerking</u>. In deze opmerking zullen wij de definitie (134) nader toelichten. Neem eens aan dat $\{g,v(i)\}$, $i \in I$, een verzameling van eindige getallen is, zó dat voldaan is aan de funktionaalvergelijking (zie stelling 1 op blz. 107)

(135) $$g+v(i) = \min_{k \in A(i)} \{K\delta(k-i) + L(k) + \sum_{j=-\infty}^{k} v(j)\ p(k-j)\}, \quad i \in I.$$

Neem verder aan dat het rechterlid van (135) voor $i < s^*$ minimaal is voor $k = S^*$ en voor $i \geq s^*$ minimaal is voor $k = i$. Dan geldt

$$v(i) = \begin{cases} L(i) - g + \sum_{j=0}^{\infty} v(i-j)\, p(j) & \text{voor } i \geq s^* \\ K + v(S^*) & \text{voor } i < s^*. \end{cases}$$

Als c een constante is, dan voldoet ook $\{g, v(i)+c\}$, $i \in I$, aan (135). Stellen wij nu $v(s^*-1) = 0$, dan is definitie (134) verklaard.

De funktie $v^*(i)$, $i \in I$, is eindig en is ondubbelzinnig vastgelegd door de "renewal equation" (134). Immers de getallen $v^*(i)$ kunnen successievelijk uit (134) berekend worden.
Door de vergelijking (134) te itereren en gebruik te maken van (93) t/m 96) vinden wij (vgl. de afleiding van (113)):

$$(136) \qquad v^*(i) = \begin{cases} 0, & i < s^*, \\ L(i) + \sum_{j=0}^{i-s^*} L(i-j)\, m(j) - g^*\{1+M(i-s^*)\}, & i \geq s^*. \end{cases}$$

Definieer de funktie

$$(137) \qquad J(k) = L(k) - g^* + \sum_{j=0}^{\infty} v^*(k-j)\, p(j), \qquad k \in I.$$

Uit (134) en (137) volgt

$$(138) \qquad J(k) = \begin{cases} L(k) - g^* & \text{voor } k < s^*, \\ v^*(k) & \text{voor } k \geq s^*. \end{cases}$$

<u>Lemma 9</u>.

(a) $J(k)$ is monotoon niet-stijgend op $(-\infty, s^*-1]$;

(b) $K+J(S^*) = 0$, $J(s^*-1) \geq 0$;

(c) $J(k) \geq J(S^*)$ voor alle $k \in I$;

(d) $J(k) \leq 0$ voor $s^* \leq k \leq S_0$;

(e) $J(k)$ is monotoon niet-stijgend op $[s^*, S_0]$;

(f) $J(k) - J(i) \geq L(k) - L(i) - K$ voor $k \geq i \geq S_0$.

Bewijs

(a) Aangezien $L(k)$ monotoon niet-stijgend op $(-\infty, S_0]$ is en $s^* \leq S_0$ is (zie (133)), volgt (a) direct uit (138).

(b) Uit (138), (136), (132) en (124) volgt

$$J(S^*) = L(S^*) + \sum_{j=0}^{S^*-s^*} L(S^*-j)\, m(j) - g(s^*,S^*)\{1+M(S^*-s^*)\} = -K.$$

Uit (138) en (132) volgt direct $J(s^*-1) = L(s^*-1) - g^* \geq 0$.

(c) Aangezien $K \geq 0$ is, volgt uit (a) en (b) dat $J(k) \geq J(s^*-1) \geq J(S^*)$ voor $k < s^*$. Om aan te tonen dat $J(k) \geq J(S^*)$ voor $k \geq s^*$, nemen wij aan dat een geheel getal $r \geq s^*$ bestaat, zó dat $J(r) < J(S^*) = -K$. Uit $J(r) < -K$ en uit (138) en (136) volgt dan

$$g^* > \{L(r) + \sum_{j=0}^{r-s^*} L(r-j)\, m(j) + K\}/\{1+M(r-s^*)\} .$$

Aangezien het rechterlid van deze ongelijkheid gelijk aan $g(r,s^*)$ is, hebben wij een tegenspraak verkregen. Hiermee is (c) bewezen.

(d) Aangezien $L(k)$ monotoon niet-stijgend op $[s^*, S_0]$ is, volgt uit (138), (134) en (132) dat $J(k) \leq \{L(s^*) - g^*\}\{1+M(k-s^*)\} \leq 0$ voor $s^* \leq k \leq S_0$.

(e) Uit (134) en (138) volgt dat

$$(139) \qquad J(k) = L(k) - g^* + \sum_{j=0}^{k-s^*} J(k-j)\, p(j) \qquad \text{voor } k \geq s^*.$$

Uit (d) en (139) volgt dat voor $s^* \leq i \leq k \leq S_0$ geldt

$$J(i) - J(k) \geq L(i) - L(k) + \sum_{j=0}^{i-s^*} \{J(i-j) - J(k-j)\}\, p(j).$$

Door deze ongelijkheid herhaald te iteren en gebruik te maken van (93) t/m (95), vinden wij dat voor $s^* \leq i \leq k \leq S_0$ geldt

$$J(i) - J(k) \geq L(i) - L(k) + \sum_{j=0}^{i-s^*} \{L(i-j) - L(k-j)\}\, m(j).$$

Punt (c) volgt nu uit deze ongelijkheid en het feit dat $L(k)$ monotoon niet-stijgend op $[s^*, S_0]$ is.

(f) Uit (139) volgt dat voor $k \geq i \geq S_0$ geldt

$$J(k) - J(i) = L(k) - L(i) + \sum_{j=0}^{i-S_0} \{J(k-j) - J(i-j)\}\, p(j) +$$

$$+ \sum_{j=i-S_0+1}^{k-s^*} J(k-j)\, p(j) - \sum_{j=i-S_0+1}^{i-s^*} J(i-j)\, p(j).$$

Uit (b) en (c) volgt dat $J(k) \geq -k$ voor alle $k \in I$, en uit (d) volgt dat $J(k) \leq 0$ voor $s^* \leq k \leq S_0$. Derhalve geldt voor $k \geq i \geq S_0$

$$J(k) - J(i) \geq L(k) - L(i) + \sum_{j=0}^{i-S_0} \{J(k-j) - J(i-j)\}\, p(j) +$$

$$-K\{F(k-s^*) - F(i-S_0)\}.$$

Door deze ongelijkheid herhaald te itereren, vinden wij dat voor $k \geq i \geq S_0$ geldt

$$J(k) - J(i) \geq L(k) - L(i) + \sum_{j=0}^{i-S_0} \{L(k-j) - L(i-j)\}\, m(j) +$$

$$-K[F(k-s^*) - F(i-S_0) + \sum_{j=0}^{i-S_0} \{F(k-s^*-j) - F(i-S_0-j)\} m(j)].$$

Punt (f) volgt nu uit het feit dat de coefficient van $-K$ in bovenstaande ongelijkheid kleiner dan of gelijk aan 1 is (zie (100)) en het feit dat $L(k)$ monotoon niet-dalend is op $[S_0, \infty]$.

Stelling 4.

(a) De verzameling van eindige getallen $\{g^*, v^*(i)\}$, $i \in I$, voldoet aan

$$(140) \quad g^* + v^*(i) = \min_{k \geq i}\{K\delta(k-i) + L(k) + \sum_{j=0}^{\infty} v^*(k-j)\, p(j)\ , \quad i \in I,$$

waarbij het rechterlid van (140) wordt geminimaliseerd door $k = S^*$ voor $i < s^*$ en door $k = i$ voor $i \geq s^*$;

(b) $$s_1 \leq s^* \leq S^* \leq S_0.$$

Bewijs.

Uit (137) volgt dat voor elke $i \in I$ geldt

$$K\delta(k-i) + L(k) - g^* + \sum_{j=0}^{\infty} v^*(k-j)\, p(j) = K\delta(k-i) + J(k), \quad k \geq i.$$

Beschouw nu voor elke vaste $i \in I$ de funktie $K\delta(k-i) + J(k)$, $k \geq i$, en onderscheid de volgende drie gevallen.

(1) $i < s^*$. Uit lemma 9(a), 9(b) en 9(c) volgt dat

$$J(i) \geq J(s^*-1) \geq K+J(S^*) = \min_{k>i}\{K+J(k)\}.$$

Dus het rechterlid van (140) wordt geminimaliseerd door $k = S^*$ voor $i < s^*$. Uit lemma 9(b) en uit (134) volgt dat $K+J(S^*) = 0 = v^*(i)$, $i < s^*$. Hiermee is (a) geverifieerd voor $i < s^*$.

(2) $s^* \leq i \leq S_0$. Uit lemma 9(b), 9(c), 9(d) en uit (138) volgt dat $K+J(k) \geq K+J(S^*) = 0 \geq J(i) = v^*(i)$ voor $k > i$. Hiermee is (a) geverifieerd voor $s^* \leq i \leq S_0$.

(3) $i > S_0$. Aangezien $L(k)$ monotoon niet-dalend is op $[S_0,\infty)$, volgt uit lemma 9(f) en uit (138) dat $K+J(k) \geq J(i) = v^*(i)$ voor $k > i$. Hiermee is (a) geverifieerd voor $i > S_0$.

(b) Wij hebben reeds aangetoond dat $s_1 \leq s^* \leq S_0$ (zie (133)). Stel eens $S^* < S_0$. Dan geldt $L(S_0) < L(S^*)$, omdat S_0 het kleinste gehele getal is waarvoor $L(k)$ het absolute minimum aanneemt. Uit (139), lemma 9(d) en 9(e) volgt nu

$$J(S_0) - J(S^*) = L(S_0) - L(S^*) + \sum_{j=0}^{S^*-s^*} \{J(S_0-j) - J(S^*-j)\}\, p(j) +$$

$$+ \sum_{j=S^*-s^*+1}^{S_0-s^*} J(S_0-j)\, p(j) \leq L(S_0) - L(S^*) < 0 .$$

Deze ongelijkheid is in tegenspraak met lemma 9(c). Derhalve geldt $S^* \geq S_0$.

Uit lemma 9(f) en 9(c) volgt dat $L(S^*) - L(S_0) - K \leq 0$. De definitie (89) impliceert nu $S^* \leq S_1$. Hiermee is het bewijs voor de stelling voltooid.

Aangezien de bestelgrenzen u en U (zie (90) op blz. 134) voldoen aan $u \leq s_1$ en $U \geq S_1$, volgt uit stelling 4 dat de verzameling van eindige getallen $\{g^*, v^*(i)\}$, $i \in I$, ook voldoet aan de funktionaalvergelijking (135), en dat tevens geldt dat het rechterlid van (135) wordt geminimaliseerd door $k = S^*$ voor $i < s^*$ en door $k = i$ voor $i \geq s^*$.

Aangezien $v^*(j) = 0$ voor $j < s^*$, en onder de voorwaarde $i_1 = i$ geldt $i_t \leq \max(i,U)$ voor alle $t \geq 1$, is voor elke $R \in C$ en $i \in I$ de rij $\{E_R(v^*(i_n)|i_1=i)\}$, $n \geq 1$, begrensd. Derhalve geldt

$$\lim_{n\to\infty} \frac{1}{n} E_R(v^*(i_n)|i_1=i) = 0 \qquad \text{voor alle } R \in C,\ i \in I.$$

Uit stelling 1 (blz. 107) volgt nu dat de (s^*,S^*) strategie optimaal is.

Samenvattend kunnen wij de volgende stelling formuleren (zie ook de lemma's 7 en 8):

<u>Stelling 5</u>.

Er bestaat een optimale (s,S) strategie en

$$g^* = \min_{R\in C} \phi(i;R) \qquad \text{voor alle } i \in I,\ R \in C,$$

waarbij $g^* = \min\{g(s,S)|s,s \in I,\ s \leq S\}$.

Als $K = 0$, dan is de (S_0,S_0) strategie optimaal. Als $K > 0$, dan voldoet elke optimale (s,S) strategie met $L(s-1) \geq g^* \geq L(s)$ aan $s_1 \leq s^* \leq S_0 \leq S^* \leq S_1$. Als $p(1) > 0$, dan geldt voor elke optimale (s,S) strategie dat $L(s-1) \geq g^* \geq L(s)$.