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.

MC SYLLABUS 1.7 b

G. DE LEVE
H.C. TIJMS

LEERGANG BESLISKUNDE

DEEL 7 b

DYNAMISCHE PROGRAMMERING 2

MATHEMATISCH CENTRUM AMSTERDAM 1974

AMS (MOS) subject classification scheme(1970): 90C40 49D99

ISBN 90 6196 055 X

1e druk 1970

2e (bijgewerkte) druk 1974

INHOUD

blz.

Bij de tweede druk

In de tweede druk zijn de bladzijden 13, 87-90 en 128-132 herzien en zijn een aantal drukfouten verwijderd. Voor het overige is de tekst ongewijzigd gebleven.

Amsterdam, 1974.

Voorwoord

In deel 7a van de syllabusserie Leergang Besliskunde zijn beslissingsproblemen met een eindig aantal beslissingstijdstippen behandeld. Daarentegen worden in het voor U liggende deel beslissingssituaties beschouwd, waarin op een onbegrensd aantal equidistant gelegen tijdstippen een beslissing genomen moet worden. Van deze beslissingsproblemen vormen voorraadproblemen een belangrijk onderdeel in deze syllabus.

Hoewel het de voorkeur verdient, is het niet noodzakelijk eerst deel 7a te bestuderen voordat men de hierna volgende tekst kan begrijpen. Eveneens is het mogelijk de drie paragrafen uit dit deel onafhankelijk van elkaar te lezen.

In deze paragrafen wordt een tweetal mogelijke optimaliteitskriteria beschouwd en wel de totale verwachte verdisconteerde opbrengst (§6) en de verwachte gemiddelde opbrengst per tijdseenheid (§7 en §8). Overal in het boekje wordt de behandelde theorie aan de hand van een aantal voorbeelden uitgebreid toegelicht. Omdat in het boekje geen index is opgenomen, is de inhoudsopgave zodanig uitgebreid, dat de lezer hieruit direct de samenstelling van het boekje kan terugvinden.

Op deze uitgave zal nog een deel 7c volgen, waarin beslissingsproblemen behandeld zullen worden, waarin de beslissingstijdstippen niet van te voren gegeven zijn en de beslisser op elk gewenst ogenblik kan ingrijpen.

De auteurs danken de heren J. HILLEBRAND, J.K. LENSTRA, Drs. A.W. SCHURINGA en Ir. P.J.WEEDA, die ieder hun bijdrage hebben geleverd aan de totstandkoming van dit deeltje. Ook zijn zij Mej. A. FASEN erkentelijk voor de voortreffelijke typografische verzorging van de tekst.

Amsterdam, 1970 GdL, HCT

6. Het ∞-stapsbeslissingsprobleem met als kriterium de totale verwachte verdisconteerde opbrengst.

Een situatie wordt beschouwd waarin een beslisser gedurende een periode van onbegrensde lengte een systeem moet beheren. Dit beheer houdt in dat de beslisser op equidistant gelegen tijdstippen een beslissing moet nemen. Het systeem doorloopt toestanden S uit een toestandsruimte $\mathcal{S}$. Bevindt het systeem zich op een beslissingstijdstip in een toestand S, dan moet de beslisser een keuze maken uit een gegeven verzameling $\mathcal{X}(S)$ van toegelaten beslissingen. Kiest de beslisser een toegelaten beslissing X in een toestand S, dan verkrijgt hij een directe opbrengst h(S,X). De opbrengstfunktie h(S,X) is voor alle $X \in \mathcal{X}(S)$ en alle $S \in \mathcal{S}$ gegeven en wordt begrensd verondersteld.

In het deterministische ∞-stapsbeslissingsprobleem wordt na een beslissing X in een toestand S, de toestand S' op het volgende beslissingstijdstip gegeven door de transformatieformule $S' = T(S,X)$. De funktie T(S,X) is voor elke $X \in \mathcal{X}(S)$ en elke $S \in \mathcal{S}$ gegeven.

In het stochastische ∞-stapsbeslissingsprobleem wordt na een beslissing X in een toestand S, de kans op [de kansdichtheid van] de toestand S' op het volgende beslissingstijdstip gegeven door de bekend veronderstelde voorwaardelijke kans $p(S'|S,X)$ [kansdichtheid $g(S'|S,X)$]. Ter wille van een uniforme notatie geven wij de verdelingsfunktie van deze voorwaardelijke kansverdeling aan met $P(S'|S,X)$.

Wij merken op dat zowel de funkties h(S,X), T(S,X) en $P(S'|S,X)$ als de verzameling $\mathcal{X}(S)$ voor elk beslissingstijdstip dezelfde zijn. Verder merken wij op dat het deterministische ∞-stapsbeslissingsprobleem als een speciaal geval te beschouwen is van het stochastische ∞-stapsbeslissingsprobleem (namelijk als $P(S'|S,X)$ voor alle $X \in \mathcal{X}(S)$ en $S \in \mathcal{S}$ de verdelingsfunktie is van een kansverdeling die geconcentreerd is in $S' = T(S,X)$). Uit een oogpunt van duidelijkheid zullen wij het deterministische en het stochastische ∞-stapsbeslissingsprobleem herhaaldelijk apart behandelen.

Voor het ∞-stapsbeslissingsprobleem moeten wij voorzichtig zijn met de keuze van een optimaliteitskriterium. Wij kunnen niet zonder meer de totale (verwachte) opbrengst over alle beslissingstijdstippen als kriterium nemen. In de meeste beslissingsproblemen zal deze opbrengst oneindig groot zijn, ongeacht de begintoestand van het systeem en ongeacht welke strategie wordt toegepast; een voorbeeld vormt een beslissingsprobleem met $h(S,X) \geq 1$ voor alle $X \in \mathcal{X}(S)$ en alle $S \in \mathcal{S}$. Het is voor het ∞-stapsbeslissingsprobleem zelfs mogelijk dat de totale (verwachte) opbrengst niet gedefinieerd is. Om dit laatste toe te lichten beschouwen wij een deterministisch ∞-stapsbeslissingsprobleem waarin $\mathcal{S}$ uit twee toestanden S_1 en S_2 bestaat en voor de beslissingen $X_1 \in \mathcal{X}(S_1)$ en $X_2 \in \mathcal{X}(S_2)$ geldt dat $S_2 = T(S_1,X_1)$, $S_1 = T(S_2,X_2)$, $h(S_1,X_1) = 1$ en $h(S_2,X_2) = -1$. Als de begintoestand S_1 is en in de toestand S_1 resp. S_2 steeds de beslissing X_1 resp. X_2 wordt genomen, verkrijgen wij de reeks van opbrengsten $1-1+1-1+\ldots$; deze som is niet gedefinieerd.

Moeilijkheden van deze aard worden voorkomen door toekomstige opbrengsten niet gelijkelijk te waarderen.*) Een opbrengst $h(S_j,X_j)$ te verkrijgen op het j^{de} beslissingstijdstip, kan in de beschouwingen op het eerste beslissingstijdstip gereduceerd worden tot een bedrag $\alpha^{j-1}h(S_j,X_j)$, waarin α een vast, gegeven getal voorstelt met $0 \leq \alpha < 1$. Bij een goede keuze van α stelt $\alpha^{j-1}h(S_j,X_j)$ het bedrag voor dat, uitgezet op het eerste beslissingstijdstip, met samengestelde interest op het j^{de} beslissingstijdstip de waarde $h(S_j,X_j)$ heeft. De <u>verdisconteringsfactor</u> α is dus gelijk aan $1/(1+r)$, waarin r de van toepassing zijnde rentefactor is. De opbrengst $\alpha^{j-1}h(S_j,X_j)$ wordt dikwijls aangegeven met de naam <u>verdisconteerde</u> opbrengst.

Veronderstel dat alle te verkrijgen opbrengsten naar het eerste beslissingstijdstip verdisconteerd worden. Laat α de vaste, gegeven verdisconteringsfactor zijn. Tenzij anders vermeld, voldoet α aan $0 \leq \alpha < 1$. De opbrengstfunktie $h(S,X)$ is begrensd en de verdisconteringsfactor is kleiner dan 1. Bijgevolg geldt voor het deterministische resp. stochastische ∞-stapsbeslissingsprobleem dat de totale verdisconteerde opbrengst resp. de totale verwachte verdisconteerde opbrengst gedefinieerd en begrensd is, ongeacht de begintoestand en ongeacht welke strategie wordt toegepast.

---

*) In §7 en §8 ontkomen wij aan deze moeilijkheden door gebruik te maken van een ander kriterium: de gemiddelde opbrengst per tijdseenheid in de "long run".

Neem eens aan dat voor het deterministische resp. het stochastische ∞-stapsbeslissingsprobleem voor elke begintoestand S gesproken kan worden van een maximaal te verkrijgen totale verdisconteerde opbrengst f(S) resp. van een maximaal te verkrijgen totale verwachte verdisconteerde opbrengst f(S); d.w.z. er bestaat een strategie waarvan de totale (verwachte) verdisconteerde opbrengst groter is dan of gelijk aan de totale (verwachte) verdisconteerde opbrengst van welke andere strategie dan ook en wel voor elke begintoestand. Onder deze veronderstelling kan bewezen worden dat voor het deterministische resp. stochastische ∞-stapsbeslissingsprobleem geldt*)

$$(6.1) \qquad f(S) = \max_{X \in \mathfrak{X}(S)} \{h(S,X) + \alpha f(T(S,X))\} \qquad \text{voor alle } S \in \mathfrak{S}$$

resp.

$$(6.2) \qquad f(S) = \max_{X \in \mathfrak{X}(S)} \{h(S,X) + \alpha \int_{S' \in \mathfrak{S}} f(S')dP(S'|S,X)\} \text{ voor alle } S \in \mathfrak{S}^{**)}$$

Deze relaties zijn intuïtief duidelijk als men bedenkt dat de uitdrukking tussen accolades in (6.1) resp. (6.2) gelijk is aan de totale (verwachte) verdisconteerde opbrengst, die verkregen wordt als op het eerste beslissingstijdstip in de begintoestand S de beslissing X wordt genomen en vanaf het tweede beslissingstijdstip een strategie wordt toegepast waarvan de totale (verwachte) verdisconteerde opbrengst maximaal is. De funktionaalvergelijking (6.1) resp. (6.2) zullen wij de optimaliteitsvergelijking noemen.

Omgekeerd kan men zich, uitgaande van de funktionaalvergelijking (6.1) resp. (6.2), afvragen of er een funktie f(S) bestaat die aan (6.1) resp. (6.2) voldoet, en zo ja, of deze oplossing uniek is. Verder kan men zich afvragen of een oplossing van (6.1) resp. (6.2) een strategie bepaalt,

---

*) D. BLACKWELL, Discounted dynamic programming, The Annals of Mathematical Statistics, 36, (1965), p. 226-235.

**) Indien $\mathfrak{S}$ een continuum is en $P(S'|S,X)$ een kansdichtheid $g(S'|S,X)$ heeft resp. indien $\mathfrak{S}$ discreet is en $P(S'|S,X)$ de verdelingsfunktie is van een discrete kansverdeling $\{p(S'|S,X)\}$, leze men $\int_{\mathfrak{S}} f(S')g(S'|S,X)dS'$ resp. $\sum_{S' \in \mathfrak{S}} f(S')p(S'|S,X)$ voor $\int_{\mathfrak{S}} f(S')dP(S'|S,X)$.

waarvan de totale (verwachte) verdisconteerde opbrengst maximaal is voor elke begintoestand.*) Wij zullen op deze vragen nader ingaan in de stellingen 6.1 t/m 6.5.

Daartoe voeren wij eerst een speciale en belangrijke klasse $\mathcal{Z}$ van strategieën in. Iedere strategie $z \in \mathcal{Z}$ heeft de volgende vorm: kies, ongeacht het beschouwde beslissingstijdstip, in elke toestand S altijd dezelfde beslissing $z(S) \in \mathfrak{X}(S)$.

Uiteraard kunnen algemenere klassen van strategieën beschouwd worden; bijv. strategieën waarvan het beslissingsvoorschrift afhangt van het beschouwde beslissingstijdstip of, nog algemener, strategieën waarvan het beslissingsvoorschrift afhangt van de in het verleden door het systeem doorlopen toestanden en de daarin genomen beslissingen. Voor de beschouwde ∞-stapsbeslissingsproblemen is het plausibel zich te beperken tot de klasse $\mathcal{Z}$ van strategieën. In dit verband merken wij op dat in het op blz. 3 genoemde artikel het volgende bewezen is. Als voor elke begintoestand gesproken kan worden van een maximaal te verkrijgen totale (verwachte) verdisconteerde opbrengst, dan bevat de klasse $\mathcal{Z}$ een strategie waarvoor deze maximale opbrengst wordt bereikt voor elke begintoestand. Een reden te meer om zich voor het ∞-stapsbeslissingsprobleem te beperken tot de klasse $\mathcal{Z}$ van strategieën, hetgeen wij dan ook zullen doen.

Voor het deterministische resp. stochastische ∞-stapsbeslissingsprobleem definiëren wij de funktie $f(S;z)$ als de totale verdisconteerde opbrengst resp. de totale verwachte verdisconteerde opbrengst ingeval de begintoestand S is en strategie $z \in \mathcal{Z}$ wordt toegepast. Aangezien $h(S,X)$ begrensd en $0 \leq \alpha < 1$ is, geldt voor elke strategie $z \in \mathcal{Z}$ dat de funktie $f(S;z)$ bestaat en begrensd is. Op eenvoudige wijze kan nagegaan worden dat voor het deterministische resp. stochastische ∞-stapsbeslissingsprobleem de funktie $f(S;z)$ voldoet aan

---

*) BLACKWELL (zie de voetnoot op blz. 3) heeft de volgende uitspraak bewezen. Als de totale (verwachte) verdisconteerde opbrengst van een strategie aan de optimaliteitsvergelijking (6.1) resp. (6.2) voldoet, is deze opbrengst voor elke begintoestand de maximaal te verkrijgen totale (verwachte) verdisconteerde opbrengst.

$$f(S;z) = h(S,z(S)) + \alpha f(T(S,z(S));z) \qquad \text{voor alle } S \in \mathcal{S} \tag{6.3}$$

resp.

$$f(S;z) = h(S,z(S)) + \alpha \int_{S' \in \mathcal{S}} f(S';z)dP(S'|S,z(S)) \qquad \text{voor alle } S \in \mathcal{S}.^{*)} \tag{6.4}$$

In het hiernavolgende zullen wij gebruik maken van een speciaal type gemengde strategieën. Volgens de gemengde strategie $(z_1)^k z_2$ met $z_1, z_2 \in \mathcal{Z}$ worden de eerste k beslissingen genomen in overeenstemming met strategie z_1 en daarna wordt uitsluitend strategie z_2 toegepast. Laat $f(S;(z_1)^k z_2)$ gelijk zijn aan de totale (verwachte) verdisconteerde opbrengst, die verkregen wordt als S de begintoestand is en de gemengde strategie $(z_1)^k z_2$ wordt toegepast. Voor het deterministische resp. stochastische ∞-stapsbeslissingsprobleem geldt voor alle $k \geqq 1$ en alle $S \in \mathcal{S}$

$$f(S;(z_1)^k z_2) = h(S,z_1(S)) + \alpha f(T(S,z_1(S));(z_1)^{k-1} z_2) \tag{6.5}$$

resp.

$$f(S;(z_1)^k z_2) = h(S,z_1(S)) + \alpha \int_{S' \in \mathcal{S}} f(S';(z_1)^{k-1} z_2)dP(S'|S,z_1(S)), \tag{6.6}$$

waarin $(z_1)^0 z_2 = z_2$.

Wij zullen nu een aantal stellingen bewijzen, die nader ingaan op de op blz. 3 en 4 opgeworpen problemen.

Stelling 6.1

Als voor een tweetal strategieën $z_1, z_2 \in \mathcal{Z}$ geldt

$$f(S;(z_1)z_2) \geqq f(S;z_2) \qquad \text{voor alle } S \in \mathcal{S}, \tag{6.7}$$

dan is

$$f(S;z_1) \geqq f(S;(z_1)z_2) \geqq f(S;z_2) \qquad \text{voor alle } S \in \mathcal{S}. \tag{6.8}$$

De stelling blijft geldig als in (6.7) en (6.8) de ongelijkheidstekens worden omgedraaid. Het bewijs is dan precies hetzelfde.

---

*) Door (6.3) resp. (6.4) herhaald in zichzelf te substitueren kan m.b.v. de begrensdheid van $f(S;z)$ en de relatie $\alpha^k \to 0$ voor $k \to \infty$, eenvoudig bewezen worden dat (6.3) resp. (6.4) een unieke oplossing heeft.

Bewijs:

Wij bewijzen de stelling alleen voor het deterministische ∞-stapsbeslissingsprobleem (het bewijs voor het stochastische beslissingsprobleem verloopt geheel analoog). Aangezien bij een zelfde begintoestand de eerste $(k-1)$ beslissingen van de gemengde strategieën $(z_1)^k z_2$ en $(z_1)^{k-1} z_2$ tot dezelfde opbrengsten leiden, volgt uit (6.7) dat voor elke $k \geqq 1$ geldt

$$(6.9) \qquad f(S;(z_1)^k z_2) \geqq f(S;(z_1)^{k-1} z_2) \qquad \text{voor alle } S \in \mathcal{S}.$$

Uit deze relatie volgt dat voor elke $k \geqq 1$ geldt

$$(6.10) \qquad f(S;(z_1)^k z_2) \geqq f(S;(z_1) z_2) \geqq f(S;z_2) \qquad \text{voor alle } S \in \mathcal{S}.$$

Aangezien de opbrengstfunktie $h(S,X)$ begrensd is en de verdisconteringsfactor α kleiner dan 1 is, bestaat er een constante c zodat $|f(S;z)| \leqq c$ is voor alle $S \in \mathcal{S}$ en alle $z \in \mathcal{Z}$. Laat S_k de toestand zijn na k maal beslissen volgens strategie z_1 als S de begintoestand is. Met behulp van (6.5) vinden wij dat voor alle $S \in \mathcal{S}$ en alle $k \geqq 1$ geldt

$$(6.11) \qquad |f(S;(z_1)^k z_2 - f(S;z_1)| = \alpha^k |f(S_k;z_2) - f(S_k;z_1)| \leqq 2\alpha^k c.$$

Omdat $\alpha^k \to 0$ voor $k \to \infty$, volgt (6.8) uit (6.10) en (6.11).

Stelling 6.2

Als in het deterministische resp. stochastische ∞-stapsbeslissingsprobleem voor een strategie $z^* \in \mathcal{Z}$ geldt

$$(6.12) \qquad f(S;z^*) = \max_{X \in \mathcal{X}(S)} \{h(S,X) + \alpha f(T(S,X);z^*)\} \qquad \text{voor alle } S \in \mathcal{S}$$

resp.

$$(6.13) \quad f(S;z^*) = \max_{X \in \mathcal{X}(S)} \{h(S,X) + \alpha \int_{S' \in \mathcal{S}} f(S';z^*) dP(S'|S,X)\} \qquad \text{voor alle } S \in \mathcal{S}$$

(m.a.w. als $f(S;z^*)$ voldoet aan de optimaliteitsvergelijking (6.1) resp.

(6.2)), dan is

(6.14) $$f(S;z^*) \geq f(S;z) \qquad \text{voor alle } S \in \mathcal{S} \text{ en alle } z \in \mathcal{Z}\,.$$

<u>Bewijs:</u>

Uit (6.5) en (6.12) resp. (6.6) en (6.13) volgt dat

(6.15) $$f(S;(z)z^*) \leq f(S;z^*) \qquad \text{voor alle } S \in \mathcal{S} \text{ en alle } z \in \mathcal{Z}\,.$$

Uit stelling 6.1 volgt nu dat (6.14) juist is.

Wij voeren de volgende definitie in. Een strategie $z^* \in \mathcal{Z}$ heet <u>optimaal</u> als

(6.16) $$f(S;z^*) \geq f(S;z) \qquad \text{voor alle } S \in \mathcal{S} \text{ en alle } z \in \mathcal{Z}\,.$$

Wij merken op dat (6.16) equivalent is met

(6.17) $$f(S;z^*) = \max_{z \in \mathcal{Z}} f(S;z) \qquad \text{voor alle } S \in \mathcal{S}.$$

<u>Stelling 6.3</u>

In het deterministische resp. stochastische ∞-stapsbeslissingsprobleem geldt voor een optimale strategie $z^* \in \mathcal{Z}$

(6.18) $$f(S;z^*) = \max_{X \in \mathcal{X}(S)} \{h(S,X) + \alpha f(T(S,X);z^*)\} \qquad \text{voor alle } S \in \mathcal{S}$$

resp.

(6.19) $$f(S;z^*) = \max_{X \in \mathcal{X}(S)} \{h(S,X) + \alpha \int_{S' \in \mathcal{S}} f(S';z^*)dP(S'|S,X)\} \text{ voor alle } S \in \mathcal{S};$$

m.a.w. $f(S;z^*)$ voldoet aan de optimaliteitsvergelijking (6.1) resp. (6.2).

<u>Bewijs:</u>

Wij beperken ons tot het deterministische beslissingsprobleem (het bewijs voor het stochastische beslissingsprobleem verloopt geheel analoog). Stel eens dat (6.18) niet juist is; d.w.z. er bestaat een toestand $S' \in \mathcal{S}$ waarvoor een beslissing $X' \in \mathcal{X}(S')$ gevonden kan worden zodat

$$h(S',X') + \alpha f(T(S',X');z^*) > f(S';z^*) \tag{6.20}$$

Definieer op $\mathcal{S}$ de strategie $z' \in \mathcal{Z}$ als volgt:

$$z'(S) = \begin{cases} z^*(S) & \text{als } S \neq S', \\ X' & \text{als } S = S'. \end{cases} \tag{6.21}$$

Dan geldt (vgl. (6.3) en (6.5))

$$f(S;(z')z^*) = f(S;z^*) \qquad \text{voor } S \in \mathcal{S} \text{ en } S \neq S' \tag{6.22}$$

en

$$f(S';(z')z^*) > f(S';z^*). \tag{6.23}$$

Uit (6.22), (6.23) en stelling 6.1 volgt dat

$$f(S;z') \geqq f(S;(z')z^*) \qquad \text{voor alle } S \in \mathcal{S}. \tag{6.24}$$

Uit (6.23) en (6.24) volgt dat

$$f(S';z') > f(S';z^*). \tag{6.25}$$

hetgeen in tegenspraak is met (6.16), omdat strategie z^* optimaal is. De veronderstelling (6.20) is dus onjuist, waarmee de stelling bewezen is.

De stellingen 6.2 en 6.3 leren ons dat een strategie $z^* \in \mathcal{Z}$ dan en slechts dan optimaal is als $f(S;z^*)$ voldoet aan de optimaliteitsvergelijking (6.1) resp. (6.2).

Wij zullen nu aantonen, dat de optimaliteitsvergelijking (6.1) resp. (6.2) een unieke oplossing bezit, die tevens een optimale strategie uit $\mathcal{Z}$ bepaalt, aangenomen dat zo'n strategie bestaat. In het bewijs wordt voor het eerst gebruikt dat voor elke $z \in \mathcal{Z}$ de funktie $f(S;z)$ ondubbelzinnig is vastgelegd door (6.3) resp. (6.4).

Stelling 6.4

Stel dat de klasse $\mathcal{Z}$ een optimale strategie bevat. Dan heeft in het deterministische resp. stochastische ∞-stapsbeslissingsprobleem de optimaliteitsvergelijking

$$(6.26)\qquad f(S) = \max_{X\in\mathcal{X}(S)} \{h(S,X) + \alpha f(T(S,X))\} \qquad \text{voor alle } S\in\mathcal{S}$$

resp.

$$(6.27)\qquad f(S) = \max_{X\in\mathcal{X}(S)} \{h(S,X) + \alpha \int_{S'\in\mathcal{S}} f(S')dP(S'|S,X)\} \qquad \text{voor alle } S\in\mathcal{S},$$

als unieke oplossing

$$(6.28)\qquad f(S) = \max_{z\in\mathcal{Z}} f(S;z) \qquad \text{voor alle } S\in\mathcal{S}.$$

Verder geldt dat een strategie $z^*\in\mathcal{Z}$ dan en slechts dan optimaal is als

$$(6.29)\qquad f(S) = h(S,z^*(S)) + \alpha f(T(S,z^*(S))) \qquad \text{voor alle } S\in\mathcal{S}$$

resp.

$$(6.30)\qquad f(S) = h(S,z^*(S)) + \alpha \int_{S'\in\mathcal{S}} f(S')dP(S'|S,z^*(S)) \qquad \text{voor alle } S\in\mathcal{S}.$$

Bewijs:

Wij geven alleen het bewijs voor het deterministische beslissingsprobleem (het bewijs voor het stochastische beslissingsprobleem verloopt geheel analoog).

Aangezien wij verondersteld hebben dat de klasse $\mathcal{Z}$ een optimale strategie bevat, volgt uit (6.17) en stelling 6.3 dat (6.28) voldoet aan (6.26). Om te bewijzen dat deze oplossing uniek is, gaan wij uit van een oplossing f(S) van (6.26). Omdat voor elke toestand S het maximum in het rechterlid van (6.26) wordt aangenomen, kunnen wij bij elke toestand S een beslissing $X_S \in \mathcal{X}(S)$ vinden, zodat geldt

$$f(S) = h(S,X_S) + \alpha f(T(S,X_S)) =$$

(6.31)

$$= \max_{X \in \mathcal{X}(S)} \{h(S,X) + \alpha f(T(S,X))\} \qquad \text{voor alle } S \in \mathcal{S}.$$

Laten wij strategie $z^* \in \mathcal{Z}$ definiëren als $z^*(S) = X_S$ voor alle $S \in \mathcal{S}$. Uit de eigenschap, dat (6.3) voor iedere $z \in \mathcal{Z}$ een unieke oplossing heeft en de eerste gelijkheid in (6.31) volgt dat $f(S) = f(S;z^*)$ voor alle $S \in \mathcal{S}$. Uit stelling 6.2, de gelijkheid $f(S) = f(S;z^*)$ voor alle $S \in \mathcal{S}$ en (6.31) volgt dat strategie z^* optimaal is. Dus $f(S)$ wordt gegeven door (6.28) (vgl. (6.17)), waarmee aangetoond is dat (6.28) de unieke oplossing is van (6.26). In bovenstaande bewijsvoering hebben wij als nevenresultaat gevonden dat strategie $z^* \in \mathcal{Z}$ optimaal is, als aan (6.29) voldaan is. Als omgekeerd strategie $z^* \in \mathcal{Z}$ optimaal is, dan volgt uit (6.28) en relatie (6.3) dat aan (6.29) voldaan is. Hiermee is het bewijs van de stelling voltooid.

Opmerking 6.1

Stel dat de klasse $\mathcal{Z}$ een optimale strategie bevat. Dan volgt uit de op blz. 4 vermelde voetnoot en stelling 6.4, dat geen "betere" strategie dan een optimale strategie uit $\mathcal{Z}$ bestaat; d.w.z. welke strategie ook wordt toegepast, voor elke begintoestand S geldt dat de te verkrijgen totale (verwachte) verdisconteerde opbrengst niet groter kan zijn dan $\max_{z \in \mathcal{Z}} f(S;z)$.

Wij hebben tot nu toe alleen uitspraken gedaan en stellingen bewezen voor ∞-stapsbeslissingsproblemen waarin de opbrengsten verdisconteerd worden met een verdisconteringsfactor die kleiner dan 1 is. Laten wij nu eens de opbrengsten niet verdisconteren; d.w.z. stel $\alpha = 1$. Wij beschouwen de speciale klasse van beslissingsproblemen met de eigenschap dat voor elke begintoestand S en elke strategie $z \in \mathcal{Z}$ de totale verwachte opbrengst, aan te geven met $f(S;z)$, bestaat en begrensd is. Voor deze beslissingsproblemen gelden de stellingen 6.1 t/m 6.3 niet zonder meer. Door de structuur van de bewijzen van deze drie stellingen voor het geval $0 \leq \alpha < 1$ te beschouwen, kan men op eenvoudige wijze nagaan dat deze stellingen ook geldig zijn voor het geval $\alpha = 1$, als voor elk tweetal strategieën $z_1, z_2 \in \mathcal{Z}$ geldt

$$(6.32) \qquad \lim_{k\to\infty} f(S;(z_1)^k z_2) = f(S;z_1) \qquad \text{voor alle } S \in \mathcal{S}.$$

De funktie $f(S;(z_1)^k z_2)$ is gedefinieerd als de totale (verwachte) opbrengst wanneer de gemengde strategie $(z_1)^k z_2$ wordt toegepast en S de begintoestand is. Het zal duidelijk zijn dat de relaties (6.7) en (6.8) ook voor $\alpha = 1$ geldig zijn. Wij zullen nu voorwaarden afleiden, waaronder (6.32) geldt.

In het deterministische ∞-stapsbeslissingsprobleem geldt voor alle $S \in \mathcal{S}$

$$(6.33) \qquad |f(S;(z_1)^k z_2) - f(S;z_1)| = |f(S_k^{z_1};z_2) - f(S_k^{z_1};z_1)|,$$

waarin $S_k^{z_1}$ de toestand aangeeft na k maal beslissen volgens strategie z_1, als de begintoestand S is. Indien, naarmate k toeneemt, de toestanden $S_k^{z_1}$ (ongeacht de toegepaste strategie z_1) steeds "slechtere" uitgangstoestanden worden voor welke strategie $z \in \mathcal{Z}$ dan ook, d.w.z. als voor alle $S \in \mathcal{S}$ en elke $z_1 \in \mathcal{Z}$ geldt

$$(6.34) \qquad \lim_{k\to\infty} f(S_k^{z_1};z) = 0 \qquad \text{voor alle } z \in \mathcal{Z},$$

dan volgt (6.32) uit (6.33) en (6.34). Bijgevolg blijven de stellingen 6.1 t/m 6.3 gelden als (6.34) van toepassing is.

Bij stochastische ∞-stapsbeslissingsproblemen kunnen soortgelijke situaties zich eveneens voordoen. Veelal treft men daarin een toestand S^* aan, waarvoor geldt

$$(6.35) \qquad f(S^*;z) = 0 \qquad \text{voor alle } z \in \mathcal{Z}.$$

In het stochastische ∞-stapsbeslissingsprobleem geldt voor $\alpha = 1$

$$(6.36) \qquad |f(S;(z_1)^k z_2 - f(S;z_1)| = |\mathcal{E} f(\underline{S}_k^{z_1};z_2) - \mathcal{E} f(\underline{S}_k^{z_1};z_1)|,$$

waarin $\underline{S}_k^{z_1}$ de stochastische toestand na k maal beslissen volgens strategie z_1 voorstelt, als de begintoestand S is.

Veronderstel dat de volgende relatie geldt:

$$\lim_{k\to\infty} P(\underline{S}_k^z = S^*) = 1 \qquad \text{voor alle } z \in \mathcal{Z}. \tag{6.37}$$

Dan volgt (6.32) uit (6.36) en (6.37).

Wij zullen hieronder een dergelijk probleem ontmoeten. In verband met de eigenschappen (6.34) en (6.37) worden problemen van dit type ook wel "uitputtingsproblemen" genoemd.

Stelling 6.4 is niet geldig voor het geval $\alpha = 1$; de optimaliteitsvergelijking (6.1) resp. (6.2), waarin $\alpha = 1$, heeft geen unieke oplossing. Als immers $f(S)$ voldoet, dan voldoet ook $f(S) + c$, waarin c een willekeurige constante is. Dezelfde opmerking is van toepassing op de oplossing $f(S;z)$ van (6.3) resp. (6.4), waarin $\alpha = 1$ wordt gesteld.

De in deze paragraaf beschreven ∞-stapsbeslissingsproblemen kunnen op verschillende manieren worden opgelost. Voor een vijftal methoden wordt hierna het oplossingsprincipe geschetst. Een aantal basisveronderstellingen wordt daarbij genoemd.

a) Methode van de successieve approximatie.

Laat $f_0(S)$, $S \in \mathcal{S}$, een willekeurige funktie zijn en laat α een vaste verdisconteringsfactor zijn met $0 \leq \alpha \leq 1$. Voor het deterministische resp. stochastische ∞-stapsbeslissingsprobleem definiëren wij de funkties $f_1(S)$, $f_2(S)$, ... door

$$f_N(S) = \max_{X \in \mathcal{X}(S)} \{h(S,X) + \alpha f_{N-1}(T(S,X))\} \qquad \text{voor alle } S \in \mathcal{S} \tag{6.38}$$

resp.

$$f_N(S) = \max_{X \in \mathcal{X}(S)} \{h(S,X) + \alpha \int_{S' \in \mathcal{S}} f_{N-1}(S')dP(S'|S,X)\} \text{ voor alle } S \in \mathcal{S}\,. \tag{6.39}$$

Wij nemen stilzwijgend aan dat deze funkties bestaan. De grootheid $f_N(S)$ kan geïnterpreteerd worden als de maximaal te verkrijgen totale (verwachte) verdisconteerde opbrengst over de eerste N beslissingstijdstippen als S de begintoestand is.

Wij beschouwen eerst het geval $0 \leq \alpha < 1$. Voor elke $S \in \mathfrak{S}$ is dan de rij van opbrengstfunkties $\{f_N(S), N \geq 1\}$ begrensd. Onder algemene voorwaarden convergeert $f_N(S)$ voor $N \to \infty$ naar de oplossing $f(S)$ van de optimaliteitsvergelijking (vgl. (6.1) en (6.2)). De funkties $f_N(S)$ kunnen wij successievelijk uit (6.38) of (6.39) bepalen. De relatie (6.38) resp. (6.39) geeft voor iedere begintoestand S een optimale beslissing $z_N(S)$ voor het eerste beslissingstijdstip van een N-stapsbeslissingsprobleem, m.a.w. voor elke S is $z_N(S)$ gedefinieerd als een beslissing waarvoor het rechterlid van (6.38) resp. (6.39) maximaal is. Wij krijgen op deze wijze een rij van strategieën $\{z_N\}$, die mogelijkerwijs voor $N \to \infty$ convergeert naar een optimale strategie van het ∞-stapsbeslissingsprobleem. Indien dit laatste niet het geval is, maar wel de limietfunktie $f(S)$ van de rij $\{f_N(S)\}$ bepaald kan worden, dan kan een optimale strategie via (6.1) resp. (6.2) verkregen worden (zie stelling 6.4).*)

Vervolgens beschouwen wij het geval $\alpha = 1$. Nu geldt veelal dat $f_N(S)$ onbegrensd groot wordt als N toeneemt. Dit behoeft echter niet te impliceren dat de achtereenvolgens verkregen beslissingsvoorschriften divergeren. Rekentechnisch kan het wel een bezwaar zijn dat bij toenemende N de funkties $f_N(S)$ elke eindige waarde overschrijden. Dit bezwaar kan in vele gevallen ondervangen worden. Wij zullen dit nader toelichten aan de hand van het deterministische ∞-stapsbeslissingsprobleem. Laat S^* een vaste, doch willekeurig gekozen toestand zijn. Voor elke toestand $S \in \mathfrak{S}$ geldt dat de volgende uitdrukkingen voor dezelfde $X \in \mathfrak{X}(S)$ hun maximale waarde bereiken:

(6.41) $\qquad h(S,X) + f_N(T(S,X))$

---

*) Beschouw het geval waarin $\mathfrak{S}$ eindig is en $\mathfrak{X}(S)$ eindig is voor alle S. Definieer $L_N(S) = f_N(S) + (1-\alpha)^{-1} \min_{S' \in \mathfrak{S}} \{f_N(S') - f_{N-1}(S')\}$ en definieer $U_N(S)$ overeenkomstig $L_N(S)$ waarbij min door max vervangen wordt. Dan geldt voor elke $N \geq 1$ dat $L_N(S) \leq f(S;z_N) \leq f(S) \leq U_N(S)$ voor alle S, waarbij $L_N(S)$ resp. $U_N(S)$ niet-dalend resp. niet-stijgend in N is naar $f(S)$. Bovendien bestaat een eindige N_0 zodat voor elke $N \geq N_0$ geldt $f(S;z_N) = f(S)$ voor alle S, m.a.w. voor elke $N \geq N_0$ is de strategie z_N optimaal. Voor een bewijs van deze beweringen verwijzen wij naar J.B. MACQUEEN, A modified dynamic programming method for Markovian decision problems, Journal of Mathematical Analysis and Applications, 14, (1966), p. 38-43 en J.F. SHAPIRO, Turnpike planning horizons for a Markovian decision model, Management Science, 14, (1968), p.292-300.

(6.42)
$$h(S,X) + f_N(T(S,X)) - f_N(S^*).$$

Als de funktie $v_k(S)$ voor alle $S \in \mathcal{S}$ gedefinieerd wordt door

(6.43)
$$v_k(S) = \begin{cases} f_k(S) - f_k(S^*) & \text{voor } k \geqq 1, \\ 0 & \text{voor } k = 0, \end{cases}$$

dan gaat (6.42) voor alle $S \in \mathcal{S}$ over in

(6.44)
$$h(S,X) + v_N(T(S,X)).$$

Uit (6.38) en (6.43) kan op eenvoudige wijze afgeleid worden dat voor elke $N \geqq 1$ geldt

(6.45)
$$\begin{aligned} v_N(S) = & \max_{X \in \mathcal{X}(S)} \{h(S,X) + v_{N-1}(T(S,X))\} + \\ & - \max_{X \in \mathcal{X}(S^*)} \{h(S^*,X) + v_{N-1}(T(S^*,X))\} \qquad \text{voor alle } S \in \mathcal{S}. \end{aligned}$$

De funkties $v_N(S)$ kunnen successievelijk uit (6.46) berekend worden. Voor de bepaling van het optimale beslissingsvoorschrift op het eerste beslissingstijdstip van het N-stapsbeslissingsprobleem kan men in plaats van (6.41) ook (6.44) gebruiken.

In een groot aantal beslissingsproblemen geldt voor elke $S \in \mathcal{S}$ dat de rij van funkties $\{v_N(S), N \geqq 1\}$ begrensd is en convergeert naar een limietfunktie $v(S)$, terwijl $\lim_{N\to\infty} f_N(S) = \infty$ voor alle $S \in \mathcal{S}$. In paragraaf 7 zullen wij op blz. 93 en 131 dergelijke v-funkties ontmoeten.

Voor het stochastische ∞-stapsbeslissingsprobleem kunnen wij op overeenkomstige wijze te werk gaan.

<u>Voorbeeld 6.0</u> <u>Een leverantieprobleem (methode a).</u>

Als toelichting op de methode van de successieve approximatie beschouwen wij nogmaals het voorbeeld 4.1 (vgl. blz. 53, deel 7a).

Neem nu aan dat het contract voor een onbegrensd aantal jaren afgesloten is. De opbrengsten worden niet verdisconteerd, zodat $\alpha = 1$ is. De numerieke gegevens zijn de volgende. Bij levering van kolensoort 1, 2 resp. 3 is de winst gelijk aan $a_1 = 435$, $a_2 = 790$ resp. $a_3 = 1050$ en is de kans dat het contract wordt opgezegd gelijk aan $p_1 = 0{,}2$, $p_2 = 0{,}4$ resp. $p_3 = 0{,}6$. Het systeem kan alleen de toestand 0 (het contract is opgezegd) en de toestand 1 (het contract loopt nog) aannemen. In toestand 0 is alleen de beslissing X = 0 (geen levering van kolen) toegelaten en in toestand 1 zijn de beslissingen X = 1,2 en 3 toegelaten (X = i betekent levering van kolensoort i). De opbrengstfunktie h(S,X) en de kansen p(S'|S,X) worden gegeven door

$$(6.46)\quad \begin{cases} h(0,0) = 0\ ;\quad h(1,X) = a_X & \text{voor } X = 1,2,3, \\ p(0|0,0) = 1\ ;\quad p(1|0,0) = 0\ ; & \\ p(0|1,X) = p_X\ ;\quad p(1|1,X) = 1 - p_X & \text{voor } X = 1,2,3. \end{cases}$$

Wanneer het systeem eenmaal in toestand 0 is, dan blijft het voor altijd in toestand 0. Voor elke strategie $z \in \mathcal{Z}$ geldt derhalve

$$(6.47)\qquad f(0;z) = 0.$$

Omdat de kansen $p_i > 0$ zijn, zal het systeem na verloop van een eindige tijd in toestand $S^* = 0$ belanden, ongeacht de toegepaste strategie. Aan de zgn. "uitputtingsvoorwaarde" (6.37) is dus voldaan.

Uiteraard geldt

$$(6.48)\qquad f_N(0) = f(0) = 0 \qquad \text{voor alle } N \geqq 1.$$

In deel 7a (blz. 55) hebben wij reeds gevonden dat

$$(6.49)\quad \begin{cases} f_1(1) = 1050\ ;\ f_2(1) = 1470\ ;\ f_3(1) = 1672\ ; \\ f_4(1) = 1793{,}20\ ;\ f_5(1) = 1869{,}56\ ;\ f_6(1) = 1930{,}65. \end{cases}$$

Voor een contract van 10, 15, 25, 35 en 50 jaar vinden wij na enig rekenwerk

$$(6.50)\quad \begin{cases} f_{10}(1) = 2074{,}91 \ ; \ f_{15}(1) = 2142{,}20 \ ; \ f_{25}(1) = 2171{,}80 \ ; \\ \\ f_{35}(1) = 2174{,}64 \ ; \ f_{50}(1) = 2174{,}99. \end{cases}$$

Een nadere bestudering van de rij $\{f_N(1), N \geq 1\}$ leidt tot de conclusie

$$(6.51)\qquad \lim_{N\to\infty} f_N(1) = 2175.$$

Kiezen wij $f(1) = 2175$, dan volgt voor X uit de optimaliteitsvergelijking

$$(6.52)\qquad 2175 = \max_{X=1,2,3} \{a_X + p_X.0 + (1-p_X)\, 2175\}$$

de oplossing $X = 1$. Laten wij de strategie $z^* \in \mathcal{Z}$ definiëren door

$$(6.53)\qquad z^*(0) = 0 \quad \text{en} \quad z^*(1) = 1.$$

Voor strategie z^* geldt dat

$$(6.54)\qquad f(0;z^*) = 0 \quad \text{en} \quad f(1;z^*) = \sum_{n=0}^{\infty} (1-p_1)^n\, a_1 = 2175$$

voldoen aan (vgl. (6.52))

$$(6.55)\quad \begin{cases} f(0;z^*) = f(0;z^*), \\ \\ f(1;z^*) = \max\limits_{X=1,2,3} \{a_X + p_X\, f(0;z^*) + (1-p_X)\, f(1;z^*)\}. \end{cases}$$

De "uitputtingsvoorwaarde" (6.37) is vervuld; bijgevolg is stelling 6.2 ook voor het geval $\alpha = 1$ geldig. Uit deze stelling en (6.55) volgt dat strategie z^* optimaal is.

b) Oplossing met behulp van de optimaliteitsvergelijking.

In de eerste plaats kan men proberen de optimaliteitsvergelijking (6.1) resp. (6.2) op te lossen. Zodra een oplossing bepaald is, levert voor vele beslissingsproblemen het vaststellen van een optimale strategie geen moeilijkheden meer op (vgl. de stellingen 6.2 en 6.4).

Oplossen van funktionaalvergelijkingen is een vak apart. Wij zullen laten zien dat voor het leverantieprobleem uit voorbeeld 6.0 de oplossing snel verkregen kan worden. Voor dit stochastische ∞-stapsbeslissingsprobleem luidt de optimaliteitsvergelijking (6.2) voor het geval $\alpha = 1$ als volgt:

$$(6.56) \quad \begin{cases} f(0) = f(0) \\ \\ f(1) = \max\limits_{X=1,2,3} \{a_X + p_X\, f(0) + (1-p_X)\, f(1)\}. \end{cases}$$

De maximaal te verkrijgen opbrengst is nul als de begintoestand 0 is. Stellen wij $f(0) = 0$ in (6.56), dan vinden we

$$(6.57) \qquad 0 = \max_{X=1,2,3} \{a_X - p_X\, f(1)\}.$$

De relatie (6.57) is equivalent met

$$(6.58) \qquad a_X - p_X\, f(1) \leqq 0 \qquad \text{voor } X = 1,2,3,$$

waarbij het gelijkheidsteken geldt voor de maximaliserende beslissing X. Aangezien $p_X > 0$ is voor $X = 1,2,3$, volgt uit (6.58) dat

$$(6.59) \qquad f(1) = \max_{X=1,2,3} \left\{\frac{a_X}{p_X}\right\} = \max \left\{\frac{435}{0,2} \,;\, \frac{790}{0,4} \,;\, \frac{1050}{0,6}\right\} = 2175.$$

De maximaliserende beslissing is $X = 1$. Vervolgens gaan wij op precies dezelfde wijze te werk als onder punt a) na relatie (6.52); wij vinden dan dat strategie $z^* \in \mathcal{Z}$ met $z^*(0) = 0$ en $z^*(1) = 1$ optimaal is.

Soms is het mogelijk uit de optimaliteitsvergelijking (6.1) resp. (6.2) direct een optimale strategie af te leiden (dus zonder eerst een oplossing f(S) te bepalen). Wij zullen dit illustreren aan de hand van voorbeeld 6.1 (zie blz. 28).

c) Oplossing met behulp van een eenvoudige deelklasse van strategieën.

In de praktijk moet men zich meestal beperken tot eenvoudig toepasbare strategieën, die tot een deelklasse $\mathcal{Z}'$ van de klasse $\mathcal{Z}$ van strategieën behoren. Dit wordt vooral vaak gedaan in de voorraadtheorie. Soms is het mogelijk de klasse $\mathcal{Z}'$ zó te kiezen, dat tenminste één optimale strategie uit $\mathcal{Z}$ ook tot de deelklasse $\mathcal{Z}'$ behoort (vgl. voorbeeld 6.2; blz. 40).

Stel dat elke strategie z uit de deelklasse $\mathcal{Z}'$ geïdentificeerd kan worden met een punt $(p_1,\dots,p_r)$ in een r-dimensionale Cartesische ruimte. Neem aan dat het mogelijk is om de kriteriumfunktie $f(S;z)$ voor $z \in \mathcal{Z}'$ te bepalen zonder de strategieparameters $p_1,\dots,p_r$ te specificeren. Dan kunnen wij de beste strategie uit de klasse $\mathcal{Z}'$ bepalen door de kriteriumfunktie $f(S;z)$, waarbij $z \in \mathcal{Z}'$, naar de parameters $p_1,\dots,p_r$ te maximaliseren.

Wij zullen deze aanpak aan de hand van voorbeeld 6.2 nader toelichten.

d) Iteratiemethode voor strategie-verbetering.

Beschouw weer de klasse $\mathcal{Z}$ van strategieën. Wij nemen aan dat de opbrengsten verdisconteerd worden met een factor α kleiner dan 1.

Verder veronderstellen wij dat voor elke strategie $z \in \mathcal{Z}$ en elke toestand S geldt dat een beslissing $X' \in \mathcal{X}(S)$ gevonden kan worden, zodat voldaan is aan

$$(6.60)\qquad h(S,X') + \alpha f(T(S,X');z) = \max_{X \in \mathcal{X}(S)} \{h(S,X) + \alpha f(T(S,X);z)\}$$

in het deterministische beslissingsprobleem resp. aan

$$(6.61)\qquad h(S,X') + \alpha \int_{S' \in \mathcal{S}} f(S';z)dP(S'|S,X') = \max_{X \in \mathcal{X}(S)} \{h(S,X) + \alpha \int_{S' \in \mathcal{S}} f(S';z)dP(S'|S,X)\}$$

in het stochastische beslissingsprobleem. Deze laatste veronderstelling is automatisch vervuld als voor alle $S \in \mathcal{S}$ de verzameling $\mathcal{X}(S)$ uit eindig veel beslissingen bestaat.*)

Wij zullen nu een iteratiemethode geven, die tot steeds betere strategieën uit $\mathcal{Z}$ leidt. Vervolgens zullen wij voor een belangrijke klasse van beslissingsproblemen aantonen, dat de iteratiemethode na een eindig aantal stappen tot een optimale strategie leidt.

De m^{de} stap uit de iteratiemethode luidt als volgt (de eerste stap beginnen wij met een willekeurige strategie $z_1 \in \mathcal{Z}$):

<u>De m^{de} stap van de iteratiemethode voor strategie-verbetering.</u>

a) Zij strategie $z_m \in \mathcal{Z}$ verkregen aan het eind van de $(m-1)^{ste}$ stap. Bepaal voor het deterministische resp. stochastische ∞-stapsbeslissingsprobleem de unieke oplossing $f(S;z_m)$ van de funktionaalvergelijking (vgl. (6.3) resp. (6.4), blz. 5)

$$f(S;z_m) = h(S,z_m(S)) + \alpha f(T(S,z_m(S));z_m) \qquad \text{voor alle } S \in \mathcal{S} \tag{6.62}$$

resp.

$$f(S;z_m) = h(S,z_m(S)) + \alpha \int_{S' \in \mathcal{S}} f(S';z_m) dP(S'|S,z_m(S)) \quad \text{voor alle } S \in \mathcal{S}. \tag{6.63}$$

b) Bepaal voor elke $S \in \mathcal{S}$ een beslissing $X \in \mathcal{X}(S)$ welke

$$h(S,X) + \alpha f(T(S,X);z_m) \tag{6.64}$$

resp.

$$h(S,X) + \alpha \int_{S' \in \mathcal{S}} f(S';z_m) dP(S'|S,X) \tag{6.65}$$

maximaliseert en kies $z_{m+1}(S)$ gelijk aan zo'n maximaliserende beslissing X. Hierbij spreken wij af dat wij $z_{m+1}(S) = z_m(S)$ kiezen als $X = z_m(S)$ de uitdrukking (6.64) resp. (6.65) maximaliseert.

<u>Einde van de m^{de} stap.</u>

---

*) Een niet-eindige verzameling van getallen behoeft geen getal te bevatten dat groter is dan of gelijk aan elk ander getal uit die verzameling; een voorbeeld vormt de aftelbare verzameling van getallen $\{-\frac{1}{n}, n \geq 1\}$.

Uit de constructie van strategie z_{m+1} en de relaties (6.5) en (6.6) volgt dat het maximum van (6.64) resp. (6.65) gelijk is aan $f(S;(z_{m+1})z_m)$. Voor alle $S \in \mathcal{S}$ geldt dat $f(S;(z_{m+1})z_m) \geqq f(S;z_m)$ (vgl. (6.62) t/m (6.65)). Uit de afspraak betreffende de keuze van $z_{m+1}(S)$ volgt dat $f(S;(z_{m+1})z_m) >$ $> f(S;z_m)$ dan en slechts dan als $z_{m+1}(S) \neq z_m(S)$. Met behulp van stelling 6.1 kunnen wij nu concluderen dat voor alle $m \geqq 1$ geldt

$$(6.66) \quad \begin{cases} f(S;z_{m+1}) \geqq f(S;z_m) & \text{voor alle } S \in \mathcal{S}, \\ \\ f(S;z_{m+1}) > f(S;z_m) & \text{als } z_{m+1}(S) \neq z_m(S). \end{cases}$$

Het is <u>niet</u> zonder meer zeker dat de rij $\{z_m, m \geqq 1\}$ naar een optimale strategie uit $\mathcal{Z}$ convergeert.

Wel kunnen wij aantonen dat, wanneer de iteratiemethode na een eindig aantal stappen (stel M) leidt tot

$$(6.67) \qquad z_M = z_{M+1},$$

een optimale strategie gevonden is. Het bewijs verloopt als volgt.

Uit (6.64) en (6.3) resp. (6.65) en (6.4) volgt

$$(6.68) \qquad f(S;(z)z_M) \leqq f(S;z_M) \qquad \text{voor alle } S \in \mathcal{S} \text{ en alle } z \in \mathcal{Z}.$$

Uit stelling 6.1 volgt nu dat $f(S;z) \leqq f(S;z_M)$ voor alle $S \in \mathcal{S}$ en alle $z \in \mathcal{Z}$; m.a.w. strategie $z_M \in \mathcal{Z}$ is optimaal.

Wij zullen nu voor een belangrijke klasse van ∞-stapsbeslissingsproblemen aantonen dat de klasse $\mathcal{Z}$ een optimale strategie bevat en dat de iteratiemethode na een eindig aantal stappen tot een optimale strategie leidt.

Beschouw de volgende klasse van stochastische ∞-stapsbeslissingsproblemen (de deterministische beslissingsproblemen zijn hier uiteraard in bevat). Het systeem kan slechts <u>eindig</u> veel toestanden aannemen. Laten wij de toestanden nummeren als $S = 1,\ldots,n$. Voor elke $S \in \mathcal{S}$ bestaat de verzameling $\mathcal{X}(S)$ uit <u>eindig</u> veel beslissingen. Omdat de toestandsruimte discreet is, is voor elke $X \in \mathcal{X}(S)$ en $S \in \mathcal{S}$ de funktie $P(S'|S,X)$ de verdelingsfunktie van een discrete kansverdeling $\{p(S'|S,X), S' = 1,\ldots,n\}$. De funktionaalvergelijking (6.4) gaat voor deze klasse van beslissingsproblemen over in

$$(6.69)\quad f(S;z) = h(S,z(S)) + \alpha \sum_{S'=1}^{n} f(S';z)p(S'|S,z(S)) \qquad \text{voor } S = 1,\dots,n.$$

Zoals reeds is opgemerkt, heeft de funktionaalvergelijking (6.4) een unieke oplossing. Niettemin laten wij zien hoe wij met matrixtheorie kunnen aantonen dat het stelsel lineaire vergelijkingen (6.69) een unieke oplossing bezit. Daartoe voeren wij de vektoren $F(z) = (f(1;z),\dots,f(n;z))$ en $H(z) = (h(1,z(1)),\dots,h(n,z(n)))$ in en de matrix $\mathfrak{P}(z)$ met als elementen de overgangskansen $p(S'|S,z(S))$. In matrixnotatie geschreven ziet het stelsel van lineaire vergelijkingen (6.69) er als volgt uit:

$$(6.70)\qquad \left[\mathfrak{J} - \alpha\mathfrak{P}(z)\right] F(z) = H(z),$$

waarin $\mathfrak{J}$ de eenheidsmatrix voorstelt. Het stelsel (6.69) heeft een unieke oplossing als de matrix $\mathfrak{J} - \alpha\mathfrak{P}(z)$ een inverse heeft. Om het bestaan van deze inverse aan te tonen, beschouwen wij de voor alle $k \geqq 1$ geldende identiteit

$$(6.71)\qquad \begin{aligned} \left[\mathfrak{J} + \sum_{m=1}^{k} \alpha^m \mathfrak{P}^m(z)\right]\left[\mathfrak{J} - \alpha\mathfrak{P}(z)\right] &= \left[\mathfrak{J} - \alpha\mathfrak{P}(z)\right]\left[\mathfrak{J} + \sum_{m=1}^{k} \alpha^m \mathfrak{P}^m(z)\right] = \\ &= \mathfrak{J} - \alpha^{k+1}\mathfrak{P}^{k+1}(z), \end{aligned}$$

waarin de matrix $\mathfrak{P}^m(z)$ het m-voudig matrixprodukt van $\mathfrak{P}(z)$ met zichzelf is. Omdat $0 \leq \alpha < 1$ is en de elementen van de matrix $\mathfrak{P}^k(z)$ kansen zijn, convergeert elk element van de matrix $\alpha^k\mathfrak{P}^k(z)$ naar 0 als k naar ∞ gaat. Uit (6.71) volgt nu dat de matrix $\mathfrak{J} - \alpha\mathfrak{P}(z)$ een inverse heeft, die gegeven wordt door

$$(6.72)\qquad (\mathfrak{J} - \alpha\mathfrak{P}(z))^{-1} = \mathfrak{J} + \sum_{m=1}^{\infty} \alpha^m \mathfrak{P}^m(z).$$

Hiermee is aangetoond dat het stelsel lineaire vergelijkingen (6.69) een unieke oplossing heeft.

Om de iteratiemethode op te stellen, hebben wij verondersteld dat de voorwaarde (6.61) vervuld is. Aan deze voorwaarde is automatisch voldaan, omdat voor elke $S \in \mathfrak{S}$ de verzameling $\mathfrak{X}(S)$ eindig is.

De klasse $\mathfrak{Z}$ bestaat uit eindig veel strategieën, omdat zowel het aantal toestanden als het aantal toegelaten beslissingen eindig is. De eindigheid van de klasse $\mathfrak{Z}$, het uniek zijn van de oplossing $f(S;z)$ van (6.69) en relatie (6.66) hebben tot gevolg dat de iteratiemethode na een eindig aantal stappen (stel M) leidt tot $z_M = z_{M+1}$. Op blz. 20 (vgl. (6.67)) hebben wij bewezen dat de strategie $z_M \in \mathfrak{Z}$ optimaal is; hiermee is tevens aangetoond dat de eindige klasse $\mathfrak{Z}$ een optimale strategie bevat. Wij vatten de gevonden resultaten samen in de volgende stelling:

Stelling 6.5

Stel dat de opbrengsten verdisconteerd worden met een factor α kleiner dan 1 en dat zowel het aantal toestanden als het aantal toegelaten beslissingen eindig is. Dan bevat de klasse $\mathfrak{Z}$ een optimale strategie en de op blz. 19 beschreven iteratiemethode convergeert na een eindig aantal stappen naar een optimale strategie uit $\mathfrak{Z}$.

e) Oplossing als lineair programmeringsprobleem.

Wij veronderstellen dat het systeem slechts n verschillende toestanden kan aannemen en verder gaan wij ervan uit dat in elke toestand het aantal toegelaten beslissingen eindig is. Laat N_i het aantal toegelaten beslissingen in toestand i zijn. Ter wille van de notatie zullen wij de toestanden onderscheiden door de index i die de waarden $1,\dots,n$ aanneemt en zullen wij $p(j|i,X)$ - de kans dat de toestand op het volgende beslissingstijdstip j is als nu in toestand i de beslissing X wordt genomen - noteren als $p_{ij}(X)$. Vanzelfsprekend moet gelden

$$(6.73) \qquad p_{ij}(X) \geq 0 \quad \text{en} \quad \sum_{j=1}^{n} p_{ij}(X) = 1 \quad \text{voor elke } X \in \mathfrak{X}(i);\ i,j = 1,\dots,n.$$

De opbrengsten worden verdisconteerd en de vaste verdisconteringsfactor α voldoet aan $0 \leq \alpha < 1$.

Uit de stellingen 6.4 (blz. 9) en 6.5 (blz.22) volgt dat

$$f(i) = \max_{z \in \mathcal{Z}} f(i;z) \qquad \text{voor } i = 1,\ldots,n, \tag{6.74}$$

de unieke oplossing vormt van de optimaliteitsvergelijking

$$f(i) = \max_{X \in \mathcal{X}(i)} \{h(i,X) + \alpha \sum_{j=1}^{n} p_{ij}(X) f(j)\} \qquad \text{voor } i = 1,\ldots,n. \tag{6.75}$$

Uit (6.75) volgt dat

$$f(i) \geqq h(i,X) + \alpha \sum_{j=1}^{n} p_{ij}(X) f(j) \quad \text{voor alle } X \in \mathcal{X}(i);\ i = 1,\ldots,n. \tag{6.76}$$

Beschouw nu eens het volgende eindige stelsel ongelijkheden in de vrije variabelen $(u_1,\ldots,u_n)$:

$$u_i - \alpha \sum_{j=1}^{n} p_{ij}(X)\, u_j \geqq h(i,X) \qquad \text{voor alle } X \in \mathcal{X}(i);\ i = 1,\ldots,n. \tag{6.77}$$

Uit (6.76) volgt dat de vektor $(u_1,\ldots,u_n) = (f(1),\ldots,f(n))$ aan het stelsel ongelijkheden (6.77) voldoet. Deze oplossing heeft de eigenschap dat bij elke $i = 1,\ldots,n$ een $X_i \in \mathcal{X}(i)$ te vinden is, zodat [vgl. (6.75)]

$$f(i) - \alpha \sum_{j=1}^{n} p_{ij}(X_i) f(j) = h(i,X_i) \qquad \text{voor } i = 1,\ldots,n. \tag{6.78}$$

Voer de vektor

$$F = (f(1),\ldots,f(n)) \tag{6.79}$$

in. We zullen nu aantonen dat voor elke oplossing $(u_1,\ldots,u_n)$ van (6.77) geldt dat

$$u_i \geqq f(i) \qquad \text{voor } i = 1,\ldots,n. \tag{6.80}$$

Zij $U = (u_1,\ldots,u_n)$ een willekeurige oplossing van (6.77). Wij definiëren de vektor $A = (a_1,\ldots,a_n)$ door

$$a_i = u_i - f(i). \tag{6.81}$$

Aangezien U = F + A voldoet aan (6.77), geldt

$$f(i) + a_i - \alpha \sum_{j=1}^{n} p_{ij}(X)\{f(j) + a_j\} \geqq h(i,X) \tag{6.82}$$

voor alle $X \in \mathfrak{X}(i)$; $i = 1,\ldots,n$.

Uit (6.78) en (6.82) volgt dat

$$a_i - \alpha \sum_{j=1}^{n} p_{ij}(X_i)a_j \geqq 0 \qquad \text{voor } i = 1,\ldots,n. \tag{6.83}$$

Definiëren wij $\mathfrak{P}$ als de n×n-matrix met als elementen $p_{ij}(X_i)$, $i,j=1,\ldots,n$, dan bestaat een vektor $D \geqq 0$ zodat (6.83) equivalent is met $(\mathfrak{I} - \alpha\mathfrak{P})A = D$. Omdat $\mathfrak{P}$ een kansenmatrix is en $0 \leqq \alpha < 1$ is, volgt uit (6.72) dat de inverse $(\mathfrak{I} - \alpha\mathfrak{P})^{-1}$ bestaat en dat alle elementen van deze matrix niet-negatief zijn. Dus alle componenten van de vektor $A = (\mathfrak{I} - \alpha\mathfrak{P})^{-1}D$ zijn niet-negatief. Hiermee is aangetoond dat (6.80) juist is.

Uit (6.80) en het feit dat $F = (f(1),\ldots,f(n))$ voldoet aan (6.77) volgt dat het lineaire programmeringsprobleem

$$\text{"minimaliseer } \sum_{j=1}^{n} u_j \text{ onder de voorwaarden (6.77); } u_i \text{ vrij"} \tag{6.84}$$

precies één optimale oplossing bezit en dat deze oplossing gegeven wordt door U = F. Uit de optimale oplossing F kunnen wij direct een optimale strategie bepalen. Bij elke i vinden wij tenminste één $X_i \in \mathfrak{X}(i)$ zodat (6.78) geldt. Elke X_i met deze eigenschap is een optimale beslissing voor toestand i (vgl. stelling 6.4).

Het lineaire programmeringsprobleem (6.84) heeft n vrije variabelen en $\sum_{i=1}^{n} N_i$ bijvoorwaarden. Aangezien het aantal bijvoorwaarden groter is dan het aantal variabelen, is het rekentechnisch voordeliger om het duale probleem van het lineaire programmeringsprobleem op te lossen. Het duale probleem (vgl. deel 6a, blz. 113) luidt

$$\text{"maximaliseer } \sum_{i=1}^{n} \sum_{X \in \mathfrak{X}(i)} h(i,X)\, v_{iX} \tag{6.85}$$

onder de bijvoorwaarden

$$(6.86)\begin{cases}\sum\limits_{X\in\mathfrak{X}(j)} v_{jX} - \alpha\sum\limits_{i=1}^{n}\sum\limits_{X\in\mathfrak{X}(i)} p_{ij}(X)\, v_{iX} = 1 & \text{voor } j = 1,\ldots,n, \\ v_{jX} \geq 0 \quad \text{voor elke } X\in\mathfrak{X}(j);\ j = 1,\ldots,n." \end{cases}$$

Uit (6.86) volgt

$$(6.87)\qquad \sum_{X\in\mathfrak{X}(j)} v_{jX} \geq 1 \qquad \text{voor } j = 1,\ldots,n.$$

Dus bij elke j behoort tenminste één $X\in\mathfrak{X}(j)$ met $v_{jX} > 0$. Aangezien het duale probleem n bijvoorwaarden bezit, heeft een toegelaten basisoplossing van (6.86) hoogstens n positieve componenten. Wij kunnen nu concluderen dat bij elke j precies één $X\in\mathfrak{X}(j)$ behoort met $v^*_{jX} > 0$ als $V^*=(v^*_{jX}, X\in\mathfrak{X}(j); 1\leq j\leq n)$ een toegelaten basisoplossing is. Uit een optimale basisoplossing van het lineaire programmeringsprobleem (6.85) kunnen wij direct een optimale strategie aflezen. Als van een optimale basisoplossing $\bar{V}$ de component $\bar{v}_{iX_i} > 0$ is, dan is X_i een optimale beslissing in toestand i. Immers uit de "complementary slackness"-eigenschap volgt dat relatie (6.78) geldt (vgl. deel 6a, blz. 117).

Het lineaire programmeringsprobleem (6.85) telt $\sum_{i=1}^{n} N_i$ variabelen waarbij N_i het aantal toegelaten beslissingen in toestand i aangeeft. Dit aantal variabelen kan zó groot worden dat het rekentechnisch niet mogelijk is om het lineaire programmeringsprobleem op te lossen. Wanneer het beslissingsprobleem echter bepaalde eigenschappen bezit, kan een gereduceerd lineair programmeringsprobleem afgeleid worden waarvan het aantal variabelen in het algemeen aanzienlijk minder is. Een optimale basisoplossing van dit probleem bepaalt ook een optimale strategie.

Wij geven eerst de voorwaarden waaraan het beslissingsprobleem moet voldoen. Vervolgens formuleren wij het gereduceerde lineaire programmeringsprobleem en geven wij de regel die uit een optimale basisoplossing een optimale strategie destilleert.*)

---

*) Voor bewijzen wordt de geïnteresseerde lezer verwezen naar het artikel van G.T. DE GHELLINCK en G.D. EPPEN, Linear programming solutions for separable Markovian decision problems, Management Science, series A, 13, (1967), p. 371-395.

Wij veronderstellen dat bij elke i een (mogelijk lege) deelverzameling $\mathcal{A}(i)$ van $\mathcal{X}(i)$ te vinden is, zodat geldt

(6.88) $\qquad \mathcal{A}(i) \supseteq \mathcal{A}(i+1) \qquad$ voor $i = 1,2,\dots,n-1$,

(6.89) $\qquad h(i,X) = c(i) + d(X) \qquad$ voor elke $X \in \mathcal{A}(i)$; $i = 1,\dots,n$,

(6.90) $\qquad p_{ij}(X) = p(j,X) \qquad$ voor elke $X \in \mathcal{A}(i)$; $i,j = 1,\dots,n$.

Vanzelfsprekend nemen wij aan dat $\mathcal{A}(1)$ niet-leeg is. Wij definiëren

(6.91) $\qquad \mathcal{B}(i) = \mathcal{X}(i) - \mathcal{A}(i) \qquad$ voor $i = 1,\dots,n$;

in woorden: $\mathcal{B}(i)$ is de verzameling van dié beslissingen uit $\mathcal{X}(i)$, welke niet tot $\mathcal{A}(i)$ behoren. Sommige van de verzamelingen $\mathcal{A}(i)$ kunnen leeg zijn. Indien $\mathcal{A}(i_0)$ leeg is, dan volgt uit (6.88) dat $\mathcal{A}(i)$ leeg is voor elke $i_0 \leq i \leq n$. Wij definiëren nu k als de grootste index i waarvoor $\mathcal{A}(i)$ niet-leeg is. Onder de niet-lege verzamelingen $\mathcal{A}(1),\dots,\mathcal{A}(k)$ kunnen gelijke voorkomen. Wij definiëren m als het aantal verschillende verzamelingen onder $\mathcal{A}(1),\dots,\mathcal{A}(k)$. Deze m verzamelingen ordenen wij naar afnemende grootte, zodat geldt

(6.92) $\qquad \mathcal{C}(1) \supset \dots \supset \mathcal{C}(m)$.

Dus $\mathcal{C}(j)$ is gelijk aan de verzameling(en) $\mathcal{A}(i)$ die, wat grootte betreft, op de j^{de} plaats komt (komen). Bij elke i waarvoor $\mathcal{A}(i)$ niet-leeg is, behoort een ondubbelzinnig bepaalde index $j(i)$. Wij definiëren $j(i)$ als dié j waarvoor $\mathcal{C}(j) = \mathcal{A}(i)$. Tenslotte definiëren wij de niet-lege verzamelingen

(6.93) $$\mathcal{D}(i) = \begin{cases} \mathcal{C}(i) - \mathcal{C}(i+1) & \text{voor } i = 1,\dots,m-1, \\ \mathcal{C}(m) & \text{voor } i = m. \end{cases}$$

Uit de gegeven definities volgt

(6.94) $$\mathcal{A}(i) = \bigcup_{r=j(i)}^{m} \mathcal{D}(r) \qquad \text{voor } i = 1,\dots,k.$$

Wij kunnen nu het gereduceerde lineaire programmeringsprobleem formuleren. Het probleem ziet er als volgt uit:

"maximaliseer

$$(6.95)\qquad \sum_{i=1}^{n} \sum_{X\in\mathfrak{B}(i)} h(i,X)v_{iX} + \sum_{i=1}^{k} c(i)r_i + \sum_{j=1}^{m} \sum_{X\in\mathfrak{D}(j)} d(X)w_{jX}$$

onder de voorwaarden

$$(6.96)\qquad \begin{cases} \sum_{X\in\mathfrak{B}(i)} v_{iX} - \alpha \sum_{h=1}^{n} \sum_{X\in\mathfrak{B}(h)} p_{hi}(X)v_{hX} + r_i - \alpha \sum_{j=1}^{m} \sum_{X\in\mathfrak{D}(j)} p(i,X)w_{jX} = 1 & \text{voor } i = 1,\dots,n, \\ -\sum_{\substack{i \text{ met} \\ j(i)=j}} r_i + \sum_{X\in\mathfrak{D}(j)} w_{jX} + t_j - t_{j-1} = 1 & \text{voor } j = 1,\dots,m, \end{cases}$$

waarbij

$$\begin{cases} r_{k+1} = \dots = r_n = 0;\ t_0 = t_m = 0, & \\ r_i \geqq 0,\ v_{iX} \geqq 0 & \text{voor elke } X\in\mathfrak{B}(i);\ i = 1,\dots,n, \\ t_j \geqq 0,\ w_{jX} \geqq 0 & \text{voor elke } X\in\mathfrak{D}(j);\ j = 1,\dots,n." \end{cases}$$

Ter wille van de notatie hebben wij in de formulering van bovenstaand lineair programmeringsprobleem "variabelen" $r_{k+1},\dots,r_n,t_0$ en t_m gebruikt. Men leze echter voor elk van deze "variabelen" een nul. Het gereduceerde lineaire programmeringsprobleem bezit m bijvoorwaarden meer dan het lineaire programmeringsprobleem (6.85), maar heeft minder variabelen. De reductie in het aantal variabelen hangt sterk af van het aantal gelijke onder de verzamelingen $\mathfrak{A}(i)$.

Het lineaire programmeringsprobleem (6.95) bezit een eindige oplossing. Op analoge wijze als bij het lineaire programmeringsprobleem (6.85) kan voor een optimale basisoplossing $(\bar{V},\bar{R},\bar{W},\bar{T})$ van (6.95) aangetoond worden dat voor elke i precies één van de componenten $\bar{r}_i,\bar{v}_{iX}$ met $X\in\mathfrak{B}(i)$ positief is en dat voor elke j precies één van de componenten $\bar{t}_j,\bar{w}_{jX}$ met $X\in\mathfrak{D}(j)$ positief is. Aangezien de "variabele" $t_m = 0$ is, volgt uit (6.96)

$$(6.97)\qquad \sum_{X\in\mathfrak{D}(m)} \bar{w}_{mX} \geqq 1.$$

Dus precies één van de $\bar{w}_{mX}$ met $X \in \mathfrak{D}(m)$ is positief. Met de volgende regel kunnen wij uit een optimale basisoplossing *) $(\bar{V},\bar{R},\bar{W},\bar{T})$ een optimale strategie afleiden:

A) als $\bar{v}_{iX} > 0$ is voor een $X \in \mathfrak{B}(i)$, dan is X een optimale beslissing in toestand i;

B) als $\bar{r}_i > 0$ is (dus $\bar{v}_{iX} = 0$ voor alle $X \in \mathfrak{B}(i)$), dan bepalen wij de kleinste $r \geqq j(i)$ met $\bar{w}_{rX} > 0$ voor een $X \in \mathfrak{D}(r)$. Zo'n r bestaat op grond van (6.97). De op deze wijze gevonden beslissing X is optimaal in toestand i.

Wij zullen het een en ander toelichten aan de hand van voorbeeld 6.4 (vgl. blz. 63).

Voorbeeld 6.1 Een vissersprobleem (methode b).

Een hartstochtelijk sportvisser heeft in een cafe horen spreken over een boer, die twee vennetjes bezit met bijzonder mooie en grote vissen. Helaas heeft de boer, zoals gebruikelijk, het vissen in beide vennetjes verboden. In het vennetje V_1 zouden zich S_1^0 en in V_2 S_2^0 vissen bevinden. Alle sprekers in het cafe waren het erover eens dat de kans om bij vennetje V_X ($X = 1,2$) gesnapt te worden, gelijk is aan p_X ($0 < p_X < 1$).

De sportvisser, die door deze mededelingen nieuwsgierig is geworden, besluit eens op verkenning uit te gaan. Op grond van zijn ervaring meent hij te mogen vaststellen dat voor iedere vis in V_X de kans op een noodlottige beet in zijn aas tijdens de vispartij gelijk is aan ρ_X ($0 < \rho_X < 1$).

Thuisgekomen overweegt hij, of hij als vreemdeling in deze omgeving geacht wordt het visverbod te kennen. Een ∞-stapsbeslissingsprobleem ontstaat zodra de sportvisser besluit elke vrije zaterdag in één van de vennetjes te gaan vissen, zich daarbij elke zaterdagmorgen opnieuw afvragend in welk ven hij zijn hengel zal uitwerpen; wel zal hij het vissen staken

---

*) Wij vermelden zonder bewijs dat van de corresponderende duale oplossing (vgl. deel 6a, stelling 7.3) de eerste n componenten gegeven worden door f(1),...,f(n).

zodra hij wordt gesnapt. Het spreekt haast vanzelf dat het deze sportvisser erom te doen is zijn in totaal te verwachten vangst zo groot mogelijk te maken.

Oplossing

De toestand S van het systeem wordt elke zaterdagmorgen bepaald door de aantallen vissen S_X in V_X ($X = 1,2$); dus $S = (S_1,S_2)$. Zodra de visser is betrapt, stellen wij $S = (0,0)$. Met betrekking tot de totaal te verwachten vangst maakt het nu niet meer uit of de visser na zijn ontmaskering doorgaat met vissen of niet.

De verzameling van toegelaten beslissingen $\mathfrak{X}(S)$ is voor iedere toestand S dezelfde en bestaat uit de volgende twee beslissingen:

$$(6.98) \qquad \begin{cases} X = 1 & \text{betekent "hengel uitwerpen in } V_1\text{"} \\ X = 2 & \text{betekent "hengel uitwerpen in } V_2\text{".} \end{cases}$$

Indien voor elke vis uit V_X de kans om tijdens een vispartij gevangen te worden steeds ρ_X bedraagt en indien zich in V_X S_X vissen bevinden, is de kans op een vangst van k vissen in ven V_X gelijk aan (binomiale verdeling; vgl. deel 2, blz. 49)

$$(6.99) \qquad p_X(k|S_X) = \binom{S_X}{k} \rho_X^k (1-\rho_X)^{S_X-k} \qquad \text{voor } 0 \leq k \leq S_X.$$

De verwachting van het aantal te vangen vissen in ven V_X wordt dus gegeven door $\rho_X\, S_X$. Indien wordt aangenomen dat de visser, wanneer hij betrapt wordt, de gevangen vissen moet afgeven, ziet de opbrengstfunktie $h(S,X)$ er als volgt uit:

$$(6.100) \qquad h(S,X) = (1-p_X)\, \rho_X\, S_X.$$

De voorwaardelijke kansverdeling van de toestand $\underline{S}' = (\underline{S}_1',\underline{S}_2')$ op de volgende zaterdagmorgen wordt gegeven door de kansen

$$
(6.101)\quad
\begin{cases}
p((0,0)|(S_1,S_2),X) = \begin{cases} 1 & \text{als } (S_1,S_2) = (0,0) \\ p_X & \text{als } S_1 > 0 \text{ en } S_2 > 0 \\ p_X & \text{als } S_X = 0 \text{ en } (S_1,S_2) \neq (0,0) \\ p_X+(1-p_X)\rho_X^{S_X} & \text{als } S_X > 0 \text{ en } S_1S_2 = 0 \end{cases} \\
p((S_1',S_2)|(S_1,S_2),1) = (1-p_1)p_1(S_1-S_1'|S_1) \text{ als } S_1 \geqq S_1' \text{ en } (S_1',S_2) \neq (0,0) \\
p((S_1,S_2')|(S_1,S_2),2) = (1-p_2)p_2(S_2-S_2'|S_2) \text{ als } S_2 \geqq S_2' \text{ en } (S_1,S_2') \neq (0,0) \\
p((S_1',S_2')|(S_1,S_2),X) = 0 \quad \text{anders.}
\end{cases}
$$

De opbrengsten worden niet verdisconteerd, zodat $\alpha = 1$ is. De nodige instrumenten voor de formulering van het ∞-stapsbeslissingsprobleem zijn nu gedefinieerd.

Laat het maximum van de in totaal te verwachten vangst aangeduid worden door $f(S_1,S_2)$, als de begintoestand (S_1,S_2) is. Het zal duidelijk zijn dat

$$(6.102)\qquad f(0,0) = 0.$$

Aan de "uitputtingsvoorwaarde" (6.37) is voldaan; welke strategie de visser ook toepast, het systeem komt op den duur noodzakelijkerwijs terecht in toestand $S^* = (0,0)$, waarin het dan voor altijd blijft. De funktie $f(S_1,S_2)$ is voor gehele en niet-negatieve S_1 en S_2 gedefinieerd. Men gaat eenvoudig na, dat deze funktie voldoet aan de funktionaalvergelijking

$$
(6.103)\qquad
\begin{aligned}
f(S_1,S_2) = \max\Big[&(1-p_1)\rho_1S_1 + (1-p_1)\sum_{k=0}^{S_1} f(S_1-k,S_2)p_1(k|S_1); \\
&(1-p_2)\rho_2S_2 + (1-p_2)\sum_{k=0}^{S_2} f(S_1,S_2-k)p_2(k|S_2)\Big]
\end{aligned}
$$

waarin S_1 en S_2 geheel en niet-negatief zijn. Onder de randvoorwaarde $f(0,0) = 0$ is de funktie $f(S_1,S_2)$ ondubbelzinnig vastgelegd door (6.103).

Om een analytische oplossing van het vissersprobleem te kunnen bepalen, moeten wij een continue versie van het gestelde probleem beschouwen. Wij zullen $f(S_1,S_2)$ generaliseren tot een funktie $f^*(S_1,S_2)$, die voor alle reële $S_1,S_2 > -1$ gedefinieerd is en gelijk aan $f(S_1,S_2)$ is voor $S_1,S_2 \geqq 0$ en geheel. Daartoe voeren wij voor elke k ($k \geqq 0$ en geheel) funkties $p_1^*(k|y)$ en $p_2^*(k|y)$ in, die voor alle reële $y > -1$ gedefinieerd en continu zijn en die tevens voldoen aan

(6.104) $\quad 0 \leqq p_i^*(k|y) \leqq 1 \qquad$ voor alle $y > -1$ en $i = 1,2$;

(6.105) $\quad p_i^*(k|n) = p_i(k|n) \qquad$ voor $n = k,k+1,\ldots$ en $i = 1,2$;

(6.106) $\quad p_i^*(k|k-1) = 0 \qquad$ voor $k = 1,2,\ldots$ en $i = 1,2$.

Dergelijke funkties $p_i^*(k|y)$ bestaan, omdat $p_i(k|n)$ de kans is op een vangst van k vissen in ven i als daarin n vissen voorkomen.

Voorts definiëren wij voor $y > -1$

(6.107) $\quad n(y) \overset{\text{def}}{=}$ kleinste gehele getal $\geqq y$.

Wij voeren nu voor reële $S_1,S_2 > -1$ de hulpfunktie $f^*(S_1,S_2)$ in, die gegeven wordt door

(6.108) $\quad f^*(S_1,S_2) = 0 \qquad$ voor $-1 < S_1,S_2 \leqq 0$,

terwijl de funktie voor $S_1 > 0$, $S_2 > -1$ en voor $S_1 > -1$, $S_2 > 0$ wordt bepaald als oplossing van de funktionaalvergelijking

$$f^*(S_1,S_2) = \max \Big[h^*(S_1,1) + (1-p_1) \sum_{k=0}^{n(S_1)} f^*(S_1-k,S_2)p_1^*(k|S_1);$$

(6.109)

$$h^*(S_2,2) + (1-p_2) \sum_{k=0}^{n(S_2)} f^*(S_1,S_2-k)p_2^*(k|S_2)\Big],$$

waarbij

(6.110) $\quad h^*(S_X,X) = \max\,[0;(1-p_X)\rho_X S_X] \qquad$ voor $S_X > -1$ en $X = 1,2$.

De relaties (6.103), (6.105) en (6.107) t/m (6.110) impliceren

(6.111) $\quad f^*(S_1,S_2) = f(S_1,S_2) \qquad$ voor $S_1,S_2 \geqq 0$ en geheel.

De relaties (6.108) en (6.109) kunnen worden samengevat tot de met deze twee relaties equivalente funktionaalvergelijking

$$(6.112) \qquad f^*(S_1,S_2) = \max\left[g_1^*(S_1,S_2)\ ;\ g_2^*(S_1,S_2)\right] \qquad \text{voor } S_1,S_2 > -1,$$

waarbij voor $S_1,S_2 > -1$ geldt

$$(6.113) \qquad g_1^*(S_1,S_2) = \frac{h^*(S_1,1) + (1-p_1)\sum\limits_{k=1}^{n(S_1)} f^*(S_1-k,S_2)p_1^*(k|S_1)}{1-(1-p_1)p_1^*(0|S_1)}$$

en

$$(6.114) \qquad g_2^*(S_1,S_2) = \frac{h^*(S_2,2) + (1-p_2)\sum\limits_{k=1}^{n(S_2)} f^*(S_1,S_2-k)p_2^*(k|S_2)}{1-(1-p_2)p_2^*(0|S_2)}$$

Als in (6.113) resp. (6.114) de sommatiegrens $n(S_1)$ resp. $n(S_2)$ gelijk aan nul is, dan is de bijbehorende som als nul gedefinieerd; dus $g_1^*(S_1,S_2)$ en $g_2^*(S_1,S_2)$ zijn nul voor $S_1,S_2 \leq 0$. Verder is zowel de noemer van het rechterlid van (6.113) als van (6.114) positief en kleiner dan 1.

Om aan te tonen dat $f^*(S_1,S_2)$ op het definitiegebied continu is, maken wij eerst de volgende twee opmerkingen:

a) als de funkties $g_1(U)$ en $g_2(U)$ continu zijn in een punt U^*, dan is ook $\max\left[g_1(U);g_2(U)\right]$ continu in U^*.

b) Laat (S_1^*,S_2^*) een punt zijn met $S_1^*,S_2^* > -1$. Als $n(S_1^*) = 0$ is, dan kunnen wij met behulp van (6.107), de relatie $p_1^*(1|0) = 0$ en de continuïteit van de funkties $p_1^*(0|S_1)$, $p_1^*(1|S_1)$ en $h^*(S_1,1)$ eenvoudig aantonen dat $g_1^*(S_1,S_2)$ continu is in (S_1^*,S_2^*). Als $n(S_1^*) \geq 1$ is, dan volgt uit (6.106) en (6.107) en uit de continuïteit van de funkties $p_1^*(k|S_1)$ en $h^*(S_1,1)$, dat $g_1^*(S_1,S_2)$ continu is in (S_1^*,S_2^*), indien $f^*(S_1,S_2)$ continu is in de punten $(S_1^*-1,S_2^*),\ldots,(S_1^*-n(S_1^*),S_2^*)$. Voor de funktie $g_2^*(S_1,S_2)$ geldt uiteraard een analoge bewering.

Uit (6.108) volgt dat $f^*(S_1,S_2)$ continu is op het gebied $\{-1 < S_1 < 0, -1 < S_2 < 0\}$. De continuïteit van $f^*(S_1,S_2)$ op dit gebied en opmerking b) hebben tot gevolg dat $g_1^*(S_1,S_2)$ en $g_2^*(S_1,S_2)$ continu zijn op zowel het gebied $\{0 \leqq S_1 < 1, -1 < S_2 \leqq 0\}$ als het gebied $\{-1 < S_1 \leqq 0, 0 \leqq S_2 < 1\}$. Uit opmerking a) volgt dan dat ook $f^*(S_1,S_2)$ continu is op deze twee gebieden. Vervolgens zijn wij in staat om met behulp van de opmerkingen a) en b) aan te tonen dat $f^*(S_1,S_2)$ continu is op het gebied $\{0 \leqq S_1 \leqq 1, 0 \leqq S_2 \leqq 1\}$. Door stapsgewijs verder te gaan ("polder steeds één eenheidsvierkant van het (S_1,S_2)-vlak in") vinden wij de volgende stelling:

<u>Stelling 6.6</u>

De funktie $f^*(S_1,S_2)$ is in elk punt (S_1,S_2) met $S_1,S_2 > -1$ ondubbelzinnig gedefinieerd en continu. Bovendien is $f^*(S_1,S_2) \geqq 0$.

De funktie $f^*(S_1,S_2)$ kunnen wij interpreteren als het maximum van de in totaal te verwachten opbrengst in een ander, nu te bespreken, ∞-stapsbeslissingsprobleem. Dit probleem volgt uit het bestaande als men voor de toestandscomponenten S_1 en S_2 elke reële waarde > -1 toelaat en $h^*(S_X,X)$ als opbrengstfunktie neemt. De component S_X geeft wederom het "aantal vissen" aan dat zich op een zaterdagmorgen in ven V_X bevindt en dit aantal mag nu niet-geheel of negatief zijn. Stel dat in toestand (S_1,S_2) de beslissing $X = 1$ [resp. $X = 2$] genomen wordt. De kans dat de sportvisser betrapt wordt, is wederom gelijk aan p_1 [resp. p_2]. Als de visser niet gesnapt wordt, is met "kans" $p_1^*(k|S_1)$, $k = 0,\dots,n(S_1)$, [resp. $p_2^*(k|S_2)$, $k = 0,\dots,n(S_2)$] de toestand op het volgende beslissingstijdstip gelijk aan (S_1-k,S_2) [resp. (S_1,S_2-k)]. Het systeem neemt de toestand (0,0) aan als de visser betrapt wordt en wanneer het systeem eenmaal deze toestand heeft aangenomen, dan houdt het die voor altijd.

Het nieuwe probleem (*) "omvat" het oude probleem. Het is zelfs identiek ermee als de componenten van de begintoestand geheel en $\geqq 0$ zijn. Immers als dit laatste het geval is, zijn de componenten van de toestand op latere tijdstippen ook geheel en $\geqq 0$.

Laten wij eens nagaan voor welke punten (S_1,S_2) met $S_1,S_2 \geqq 0$ het in het nieuwe probleem onverschillig is of men eerst de beslissing $X = 1$ kiest, daarna $X = 2$ om vanaf het derde tijdstip optimaal verder te gaan, dan wel dat men begint met $X = 2$ en vervolgt met $X = 1$, alvorens optimaal verder te beslissen. Merk op dat in beide alternatieven de eerste twee beslissingen niet optimaal behoeven te zijn. De te verwachten opbrengst zullen wij aangeven met $f^*_{12}(S_1,S_2)$ resp. $f^*_{21}(S_1,S_2)$. Men kan gemakkelijk nagaan dat

$$\begin{aligned} f^*_{12}(S_1,S_2) &= (1-p_1)\rho_1 S_1 + (1-p_1)(1-p_2)\rho_2 S_2 + \\ &+ (1-p_1)(1-p_2) \sum_{h=0}^{n(S_1)} \sum_{k=0}^{n(S_2)} f^*(S_1-h,S_2-k)p^*_1(h|S_1)p^*_2(k|S_2) \end{aligned} \tag{6.115}$$

en

$$\begin{aligned} f^*_{21}(S_1,S_2) &= (1-p_2)\rho_2 S_2 + (1-p_1)(1-p_2)\rho_1 S_1 + \\ &+ (1-p_1)(1-p_2) \sum_{h=0}^{n(S_1)} \sum_{k=0}^{n(S_2)} f^*(S_1-h,S_2-k)p^*_1(h|S_1)p^*_2(k|S_2). \end{aligned} \tag{6.116}$$

De relatie $f^*_{12}(S_1,S_2) = f^*_{21}(S_1,S_2)$ is dus equivalent met

$$(1-p_1)\rho_1 p_2 S_1 = (1-p_2)\rho_2 p_1 S_2. \tag{6.117}$$

Uit (6.117) volgt (zie fig. 6.1) dat de gezochte punten liggen op een rechte L door de oorsprong vàn een 2-dimensionaal orthogonaal assenstelsel (S_1,S_2). In fig. 6.1 is tevens aangegeven voor welke punten (S_1,S_2) met $S_1,S_2 \geqq 0$ geldt dat $f^*_{12}(S_1,S_2)$ kleiner dan resp. groter dan $f^*_{21}(S_1,S_2)$ is.

Wij zullen nu aantonen dat de rechte L een optimale strategie bepaalt. Daartoe zullen wij mede gebruik maken van de relatie (6.112) en van de continuïteit van de niet-negatieve funkties $g^*_1(S_1,S_2)$ en $g^*_2(S_1,S_2)$.

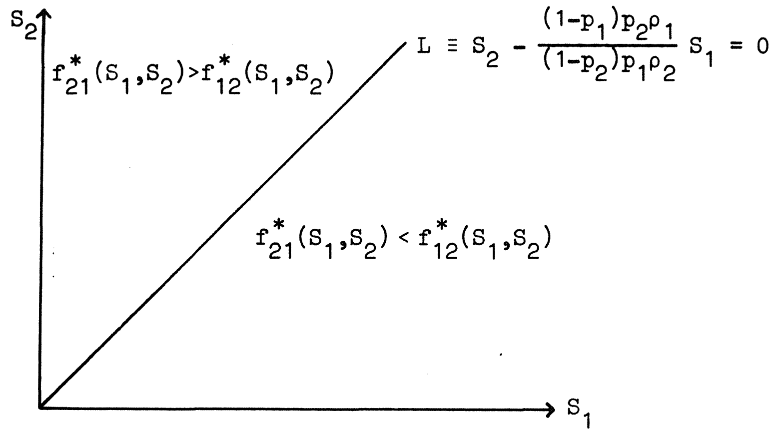

fig. 6.1

Definiëren wij voor $S_1 > 0$, $S_2 > -1$ en voor $S_1 > -1$, $S_2 > 0$

(6.118) $\quad v^*(S_1,S_2) = g^*_1(S_1,S_2) - g^*_2(S_1,S_2)$

dan volgt uit (6.112) dat

(6.119) $\quad v^*(S_1,S_2) > 0 \Longleftrightarrow$ alleen $X = 1$ is optimaal in (S_1,S_2),

(6.120) $\quad v^*(S_1,S_2) < 0 \Longleftrightarrow$ alleen $X = 2$ is optimaal in (S_1,S_2),

(6.121) $\quad v^*(S_1,S_2) = 0 \Longleftrightarrow$ zowel $X = 1$ als $X = 2$ is optimaal in (S_1,S_2).

Wij zullen nu aantonen dat alleen de beslissing $X = 1$ optimaal is voor de punten (S_1,S_2) die onder L liggen en waarvoor $S_1 > 0$ is. Daartoe merken wij eerst op dat in toestand (S_1,S_2) met $S_1 > 0$, $-1 < S_2 \leqq 0$ alleen $X = 1$ optimaal is, omdat [vgl. (6.113), (6.114) en stelling 6.6]

(6.122) $\quad g^*_1(S_1,S_2) > g^*_2(S_1,S_2) = 0 \qquad$ voor $S_1 > 0$, $-1 < S_2 \leqq 0$.

Beschouw nu eens een vast punt (S^*_1,S^*_2) met $S^*_1,S^*_2 > 0$, dat onder de lijn L ligt. Neem aan dat $X = 2$ optimaal is in (S^*_1,S^*_2). Wij zullen laten zien dat deze aanname tot een tegenspraak leidt. De veronderstelling dat $X = 2$ optimaal is in (S_1^*,S_2^*) impliceert

(6.123) $$v^*(S_1^*,S_2^*) \leqq 0.$$

Uit (6.122) volgt

(6.124) $$v^*(S_1^*,0) > 0.$$

Stelling 6.6 en opmerking b) op blz. 33 resp. 32 hebben tot gevolg dat bij vaste S_1^* de funktie $v^*(S_1^*,S_2)$ continu is voor $S_2 \geqq 0$. Dus bestaat er een getal S_2^{**}, waarvoor geldt

(6.125) $$0 < S_2^{**} \leqq S_2^*,$$

(6.126) $$v^*(S_1^*,S_2^{**}) = 0,$$

(6.127) $$v^*(S_1^*,S_2) > 0 \qquad \text{voor } 0 \leqq S_2 < S_2^{**}.$$

Uit (6.119), (6.121), (6.122), (6.126) en (6.127) volgt dat X = 1 de enige beslissing is die optimaal is in de toestanden (S_1^*,S_2) met $-1 < S_2 < S_2^{**}$ en dat zowel X = 1 als X = 2 optimaal is in (S_1^*,S_2^{**}). Als in toestand (S_1^*,S_2^{**}) de beslissing X = 2 wordt genomen en "de visser wordt niet betrapt", dan geldt voor de toestand (S_1^*,S_2) op het volgende tijdstip dat $-1 < S_2 \leqq S_2^{**}$ is (als de visser betrapt wordt, neemt het systeem toestand (0,0) aan, waarin zowel X = 1 als X = 2 optimaal is). Dus na de optimale keuze van X = 2 in toestand (S_1^*,S_2^{**}) is het optimaal om op het volgende beslissingstijdstip X = 1 te kiezen. Bijgevolg geldt

(6.128) $$f^*(S_1^*,S_2^{**}) = f_{21}^*(S_1^*,S_2^{**}).$$

Uit (6.125) volgt dat het punt (S_1^*,S_2^{**}) onder de lijn L ligt, zodat (vgl. fig. 6.1)

(6.129) $$f_{21}^*(S_1^*,S_2^{**}) < f_{12}^*(S_1^*,S_2^{**}).$$

De relaties (6.128) en (6.129) zijn met elkaar in tegenspraak, omdat

$f^*(S_1^*,S_2^{**})$ de maximaal te verwachten opbrengst is. Dus alleen de beslissing $X = 1$ is optimaal in (S_1^*,S_2^*).

Op dezelfde wijze als aangetoond is dat slechts de beslissing $X = 1$ optimaal is in de punten (S_1,S_2) die onder de lijn L liggen en waarvoor $S_1 > 0$ is, kan bewezen worden dat alleen de beslissing $X = 2$ optimaal is in de punten (S_1,S_2) die boven L liggen en waarvoor $S_2 > 0$ is. Voor de punten óp L is zowel $X = 1$ als $X = 2$ optimaal.

Voor het (*)-probleem is de optimale strategie nu bepaald. Aangezien het (*)-probleem slechts een uitbreiding is van het oorspronkelijke probleem, is nu ook de optimale strategie van dat probleem gevonden. In fig. 6.2 wordt de optimale strategie, die dus bepaald is zonder $f(S_1,S_2)$ expliciet te kennen, nogmaals aangegeven.

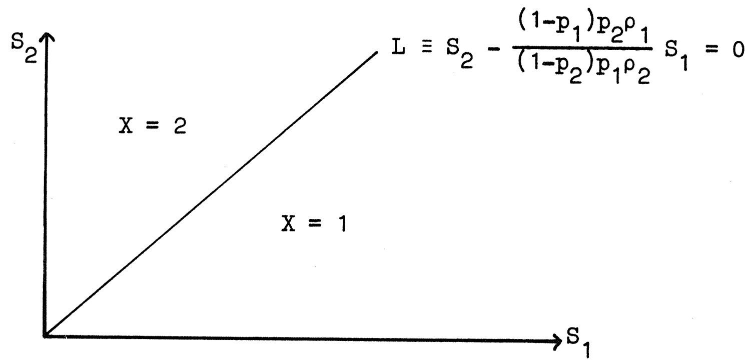

fig. 6.2

De optimale strategie

In het hierna te behandelen voorbeeld 6.2 zullen wij een funktionaalvergelijking van het volgende type ontmoeten:

$$(6.130)\quad g(S) = k(S) + \alpha \int_0^{S-a} g(S-u)w(u)du + \alpha \int_{S-a}^{\infty} g(b)w(u)du \quad \text{voor } S \geq a,$$

waarin $0 \leq \alpha < 1$ en $a \leq b$ is. De funktie $k(S)$ is gegeven en begrensd op ieder eindig interval. De funktie $w(u)$ stelt de kansdichtheid van een nietnegatieve stochastische variabele voor; dus geldt

$$(6.131)\qquad \begin{cases} w(u) \geqq 0, \\ \int_0^\infty w(u)du = 1. \end{cases}$$

Bovenstaande funktionaalvergelijking in g(S) kan voor $S \geqq a$ geschreven worden als

$$(6.132)\qquad g(S) - g(b) = k(S) - (1-\alpha)g(b) + \alpha \int_0^{S-a} \{g(S-u)-g(b)\}w(u)du.$$

Wij definiëren de volgende funkties:

$$(6.133)\qquad g^*(S) = g(S) - g(b) \qquad \text{voor } S \geqq a,$$

$$(6.134)\qquad k^*(S) = k(S) - (1-\alpha)g(b) \qquad \text{voor } S \geqq a,$$

$$(6.135)\qquad w_j(u) = \begin{cases} w(u) & \text{voor } j = 1;\ u \geqq 0, \\ \int_0^u w_{j-1}(u-v)w(v)dv & \text{voor } j \geqq 2;\ u \geqq 0, \end{cases}$$

en

$$(6.136)\qquad w(u;\alpha) = \sum_{j=1}^{\infty} \alpha^j w_j(u) \qquad \text{voor } u \geqq 0.$$

De funktie $w_j(u)$ kan geïnterpreteerd worden als de kansdichtheid van de som van j onderling onafhankelijke stochastische variabelen, die elk w(u) als dichtheid hebben [vgl. deel 2; stelling 11.2 op blz. 98]. Wij merken op dat het rechterlid van (6.136) gedefinieerd is.

Wij kunnen nu (6.132) schrijven als [vgl. (6.133) en (6.134)]

$$(6.137)\qquad g^*(S) = k^*(S) + \alpha \int_0^{S-a} g^*(S-u)w(u)du \qquad \text{voor } S \geqq a.$$

Uit een herhaalde toepassing van (6.137) en uit $\lim_{j\to\infty} \alpha^j = 0$ volgt dat (6.137) een unieke oplossing bezit, die gegeven wordt door

$$(6.138)\qquad g^*(S) = k^*(S) + \int_0^{S-a} k^*(S-u)w(u;\alpha)du \qquad \text{voor } S \geqq a.$$

Uit (6.133) volgt dat $g^*(b) = 0$, zodat geldt

$$(6.139)\qquad k^*(b) + \int_0^{b-a} k^*(b-u)w(u;\alpha)du = 0.$$

Hieruit volgt [vgl. (6.134)]

$$(6.140)\quad k(b) - (1-\alpha)g(b) + \int_0^{b-a} k(b-u)w(u;\alpha)du - (1-\alpha)g(b)\int_0^{b-a} w(u;\alpha)du = 0,$$

waardoor we uitkomen op de volgende relatie:

$$(6.141)\qquad g(b) = \frac{k(b) + \int_0^{b-a} k(b-u)w(u;\alpha)du}{(1-\alpha)[1 + \int_0^{b-a} w(u;\alpha)du]}.$$

De oplossing van de funktionaalvergelijking (6.130) wordt dus voor $S \geqq a$ gegeven door [vgl. (6.133), (6.134) en (6.138)]

$$(6.142)\quad g(S) = k(S) + \{\alpha-(1-\alpha)\int_0^{S-a} w(u;\alpha)du\}g(b) + \int_0^{S-a} k(S-u)w(u;\alpha)du.$$

Op overeenkomstige wijze bepalen wij de oplossing van de funktionaalvergelijking

$$(6.143)\quad g(S) = k(S) + \alpha\sum_{i=0}^{S-a} g(S-i)p(i) + \alpha\sum_{i=S-a+1}^{\infty} g(b)p(i) \quad \text{voor } S = a,a+1,\ldots$$

waarin $0 \leqq \alpha < 1$ is, a en b geheelwaardig zijn en voldoen aan $a \leqq b$, en de

$p(i)$ voor alle niet-negatieve i kansen voorstellen met $\sum_{i=0}^{\infty} p(i) = 1$. Wij definiëren

$$(6.144) \qquad p_j(i) = \begin{cases} p(i) & \text{voor } j = 1;\ i \geq 0, \\ \sum_{r=0}^{i} p_{j-1}(i-r)p(r) & \text{voor } j \geq 2;\ i \geq 0, \end{cases}$$

en

$$(6.145) \qquad p(i;\alpha) = \sum_{j=1}^{\infty} \alpha^j p_j(i) \qquad \text{voor } i \geq 0.$$

De oplossing van de funktionaalvergelijking (6.143) wordt dan voor $S = a,a+1,\ldots$ gegeven door

$$(6.146) \qquad g(S) = k(S) + \{\alpha-(1-\alpha)\sum_{i=0}^{S-a} p(i;\alpha)\}g(b) + \sum_{i=0}^{S-a} k(S-i)p(i;\alpha),$$

waarin

$$(6.147) \qquad g(b) = \frac{k(b) + \sum_{i=0}^{b-a} k(b-i)p(i;\alpha)}{(1-\alpha)(1 + \sum_{i=0}^{b-a} p(i;\alpha))}.$$

<u>Voorbeeld 6.2</u> <u>(S^+,s_-)-strategieën voor diverse voorraadmodellen (methode c).</u>

In eerste instantie beschouwen wij een (S^+,s_-)-voorraadmodel met <u>nalevering</u>, <u>levertijd nul</u> en een <u>continu verdeelde vraag</u>. De beschrijving van dit model luidt als volgt.

Een voorraadbeheerder kan zijn voorraad slechts aanvullen op tevoren gegeven equidistante tijdstippen, welke liggen aan het begin van de perioden 1,2,... . De aanvullingen worden zonder levertijd verkregen. De behoefte aan goederen in een periode gedraagt zich als een stochastische grootheid. De kansverdelingen van de vraag in de verschillende perioden worden identiek en onafhankelijk verondersteld te zijn. Wij nemen aan dat de behoefte $\underline{b}$ een positieve kansdichtheid $w(u)$ bezit. De vraag wordt verondersteld plaats te vinden aan het eind van elke periode.

Indien de voorraad ontoereikend is voor de vraag, zal door de voorraadbeheerder worden nageleverd, zodra door een bestelling voorraad ter beschikking komt. Dit brengt uiteraard boetekosten met zich mee.

De inkoopkosten zijn een gegeven funktie $\phi(q)$ van de bestelgrootte q. Wij nemen aan dat de verwachting van de voorraadkosten en de boetes, welke worden opgelopen in een periode tussen twee equidistante tijdstippen, een gegeven funktie $h(S)$ is van de voorraad S, die na een eventuele aanvulling aan het begin van de beschouwde periode aanwezig is. Wij laten voor de voorraad ook negatieve waarden toe door een negatieve voorraad te interpreteren als achterstand in de levering.

Wij zullen alleen strategieën van het volgende type beschouwen: "Vul aan het begin van een periode de voorraad allééén dan aan als deze kleiner is dan s_- en vul in dat geval aan tot S^+". Wij noteren een dergelijke strategie als $z = (S^+,s_-)$. In figuur 6.3 is het verloop van de voorraad voor een aantal opeenvolgende perioden aangegeven, als een (S^+,s_-)-strategie wordt toegepast.

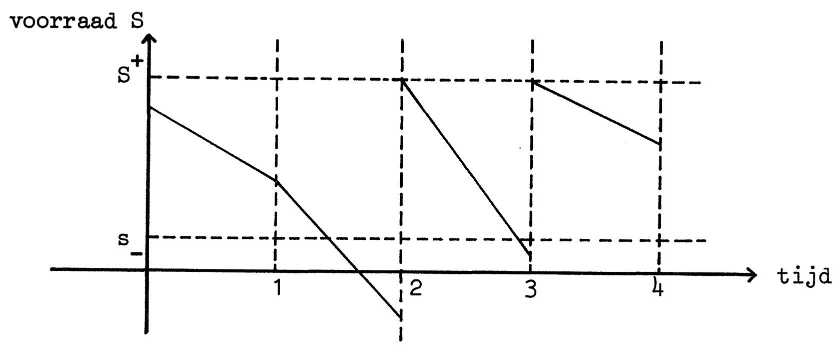

fig. 6.3

Een (S^+,s_-)-strategie

De te beschouwen strategieën verschillen alleen in de keuze van de grootheden S^+ en s_-. Vanzelfsprekend moet gelden

(6.148) $$S^+ \geqq s_- .$$

De (S^+,s_-)-strategieën worden vanwege hun eenvoud in de praktijk vaak toegepast, ook al is de optimale strategie niet van het type (S^+,s_-). Aansluitend op door H. SCARF *) gevonden resultaten heeft D.L. IGLEHART **) echter aangetoond dat de optimale strategie een (S^+,s_-)-strategie is, als $h(S)$ convex is, $\phi(0) = 0$ en $\phi(q) = cq + K$ voor $q > 0$, waarbij $K \geqq 0$ is. Voor $K = 0$, zijn de optimale s_- en S^+ aan elkaar gelijk. Wij zien dus dat onder algemene voorwaarden de optimale strategie behoort tot de te beschouwen klasse van (S^+,s_-)-strategieën. De bepaling van de beste strategie uit deze klasse kan herleid worden tot een één-stapsbeslissingsprobleem, zodra wij in staat zijn een kriteriumfunktie op te stellen waarin de parameters S^+ en s_- expliciet voorkomen. In de hierna te geven oplossing wordt een dergelijke kriteriumfunktie afgeleid.

Oplossing

De toestand van het systeem wordt gedefinieerd door de omvang van de voorraad (negatieve voorraad is achterstand!); de toestand wordt alleen vlak voor een eventuele aanvulling gemeten. De beslissing X identificeren wij met de voorraad vlak na een eventuele aanvulling. De verzameling van toegelaten beslissingen in toestand S is dus

(6.149) $$\mathfrak{X}(S) = \{X \mid X \geqq S\}.$$

Aangezien in dit probleem alleen kosten voorkomen, geeft $h(S,X)$ de verwachting aan van de kosten in het komende tijdsinterval. In $h(S,X)$ zijn de inkoopkosten en de verwachting van de voorraad- en boetekosten opgenomen. De funktie $h(S,X)$ wordt dus gegeven door

---

*) H. SCARF, The optimality of (S,s) policies in the dynamic inventory problem, ch. 13 (p. 196) uit K.J. ARROW, S. KARLIN and P. SUPPES (ed.), Mathematical methods in the social sciences, (1960), Stanford Univ. Press, Stanford (Calif.).

**) D.L. IGLEHART, Optimality of (S,s) policies in the infinite horizon dynamic inventory problem, Management Science, 9, (1963), p. 259-268.

$$(6.150)\qquad h(S,X) = \begin{cases} h(S) & \text{als } X = S, \\ h(X) + \phi(X-S) & \text{als } X > S. \end{cases}$$

De voorwaardelijke kansdichtheid van de toestand S' op het volgende equidistante tijdstip wordt gegeven door

$$(6.151)\qquad g(S'|S,X) = \begin{cases} w(X-S') & \text{als } S' \leqq X, \\ 0 & \text{als } S' > X. \end{cases}$$

Door (6.149), (6.150) en (6.151) is het stochastische ∞-stapsbeslissingsprobleem gedefinieerd. Wij nemen aan, dat de verdisconteringsfactor α kleiner dan 1 is.

Past de voorraadbeheerder strategie $z = (S^+,s_-)$ toe en is de begintoestand S, dan wordt de verwachting van de totale verdisconteerde kosten gegeven door de funktie $f(S;z)$. Deze funktie voldoet aan de funktionaalvergelijking

$$(6.152)\qquad f(S;z) = \begin{cases} \phi(S^+-S) + f(S^+;z) & \text{voor } S < s_-, \\ h(S) + \alpha \int_0^{S-s_-} f(S-u;z)w(u)du + \alpha \int_{S-s_-}^{\infty} \{f(S^+;z) + \phi(S^+-S+u)\}w(u)du & \text{voor } S \geqq s_-. \end{cases}$$

Dus voor $S \geqq s_-$ geldt

$$(6.153)\qquad \begin{aligned} f(S;z) = h(S) &+ \alpha \int_{S-s_-}^{\infty} \phi(S^+-S+u)w(u)du + \alpha \int_0^{S-s_-} f(S-u;z)w(u)du + \\ &+ \alpha \int_{S-s_-}^{\infty} f(S^+;z)w(u)du. \end{aligned}$$

Schrijven wij

$$(6.154)\qquad k(S) = h(S) + \alpha \int_{S-s_-}^{\infty} \phi(S^+-S+u)w(u)du$$

en

$$(6.155)\qquad g(S) = f(S;z)\ ;\ a = s_-\ ;\ b = S^+,$$

dan gaat (6.153) over in (6.130).

Bijgevolg geldt [vgl. (6.141) en (6.142)]

$$(6.156)\qquad f(S;z)=\begin{cases} h(S) + \alpha \displaystyle\int_{S-s_-}^{\infty} \phi(S^+-S+u)w(u)du + \{\alpha-(1-\alpha)\displaystyle\int_{0}^{S-s_-} w(v;\alpha)dv\}f(S^+;z) + \\ + \displaystyle\int_{0}^{S-s_-} \{h(S-u) + \alpha \displaystyle\int_{S-s_--u}^{\infty} \phi(S^+-S+u+v)w(v)dv\}w(u;\alpha)du & \text{voor } S \geqq s_-, \\ \\ \phi(S^+-S) + f(S^+;z) & \text{voor } S < s_-, \end{cases}$$

waarin

$$(6.157)\qquad f(S^+;z)= \frac{h(S^+)+\alpha \displaystyle\int_{S^+-s_-}^{\infty} \phi(u)w(u)du+ \displaystyle\int_{0}^{S^+-s_-} \{h(S^+-u)+\alpha \displaystyle\int_{S^+-s_--u}^{\infty} \phi(u+v)w(v)dv\}w(u;\alpha)du}{(1-\alpha)(1+\displaystyle\int_{0}^{S^+-s_-} w(u;\alpha)du)}\ .$$

De gezochte kriteriumfunktie voor de beste (S^+,s_-)-strategie wordt voor iedere <u>vaste</u> begintoestand S gegeven door (6.156). Wij hebben dus voor iedere begintoestand S een kriteriumfunktie in S^+ en s_- gevonden, waarbij men echter wel moet bedenken dat deze funktie voor $s_- \leqq S$ anders gedefinieerd is dan voor $s_- > S$ (S vast!). De beste (S^+,s_-)-strategie, ingeval de begintoestand S is, vinden wij door $f(S;z)$ te minimaliseren naar S^+ en s_-. Wij kunnen niet zonder meer stellen dat er een (S^+,s_-)-strategie bestaat die voor iedere begintoestand de beste is. Als wij echter aannemen dat $h(S)$ een convexe en $\phi(q)$ een lineaire funktie is, bestaat er een (S^+,s_-)-strategie die voor iedere begintoestand optimaal is (vgl. bijv. het in de voetnoot op blz. 42 genoemde artikel van IGLEHART).

Laten wij vanaf nu aanemen dat $h(S)$ convex en differentieerbaar is en dat de inkoopfunktie gegeven wordt door

$$(6.158) \qquad \phi(q) = \begin{cases} cq + K & \text{voor } q > 0, \\ 0 & \text{voor } q = 0, \end{cases}$$

waarin $K \geqq 0$ is. Verder nemen wij aan dat de funktie $h(x) + c(1-\alpha)x$ voor een eindige waarde van x een minimum aanneemt *). Deze laatste voorwaarde zal in de praktijk altijd vervuld zijn.

Met betrekking tot K kunnen wij de volgende twee gevallen onderscheiden:

<u>Geval I</u>

$$(6.159) \qquad K = 0.$$

Voor de optimale (S^+,s_-)-strategie geldt dan (vgl. IGLEHART; voetnoot blz. 42)

$$(6.160) \qquad s_- = S^+.$$

Wij stellen nu de variabelen S^+ en s_- aan elkaar gelijk. Voor iedere begintoestand S is de bijbehorende kriteriumfunktie $f(S;z)$ nu alleen een funktie van s_-. Na enig rekenwerk vinden wij dat zowel voor $s_- < S$ als voor $s_- > S$ (S vast!) uit

*) Als dit niet het geval is, dan is "nimmer bestellen" optimaal.

(6.161) $$\frac{df(S;z)}{ds_-} = 0$$

volgt dat

(6.162) $$h'(s_-) + c(1-\alpha) = 0.$$

Deze vergelijking heeft tenminste één oplossing; we nemen de kleinste oplossing als de optimale s_-.

Geval II

(6.163) $$K > 0.$$

Voor een optimale (S^+,s_-)-strategie geldt nu (vgl. IGLEHART)

(6.164) $$S^+ > s_-.$$

Laten wij (S^+-s_-) en S^+ als variabelen nemen i.p.v. s_- en S^+. Wij vinden na enig rekenwerk dat zowel voor $s_- < S$ als voor $s_- > S$ (S vast) de relatie

(6.165) $$\frac{\partial f(S;z)}{\partial(S^+-s_-)} = 0$$

gelijkwaardig is met

(6.166) $$f(S^+;z) = \frac{h(s_-') - \phi(S^+-s_-) + \alpha\int_0^\infty \phi(S^+-s_-+u)w(u)du}{1-\alpha}.$$

Aangezien $\phi(q) = cq + K$ voor $q > 0$, kunnen wij (6.166) vereenvoudigen tot

(6.167) $$f(S^+;z) = \frac{h(s_-) + \alpha c\mu}{1-\alpha} - c(S^+-s_-) - K,$$

waarin

(6.168) $$\mu \overset{\text{def}}{=} \int_0^\infty uw(u)du.$$

Na enig rekenwerk vinden wij dat zowel voor $s_- < S$ als voor $s_- > S$ (S vast) uit

$$\frac{\partial f(S;z)}{\partial S^+} = 0 \tag{6.169}$$

volgt dat

$$\frac{h'(S^+) + \int_0^{S^+-s_-} h'(S^+-u)w(u;\alpha)du}{(1-\alpha)(1+\int_0^{S^+-s_-} w(u;\alpha)du)} + c = 0. \tag{6.170}$$

Uit (6.135) en (6.136) volgt voor $u \geqq 0$

$$w(u;\alpha) = \alpha\, w(u) + \alpha \int_0^u w(u-v;\alpha)w(v)dv. \tag{6.171}$$

Met behulp van (6.158) en (6.171) kunnen we na enig rekenwerk formule (6.157) vereenvoudigen tot

$$f(S^+;z) = \frac{h(S^+) + \int_0^{S^+-s_-} h(S^+-u)w(u;\alpha)du + K + c(\alpha-1)\int_0^{S^+-s_-} uw(u;\alpha)du}{(1-\alpha)(1+\int_0^{S^+-s_-} w(u;\alpha)du)} + \frac{\alpha c\mu}{1-\alpha} - K. \tag{6.172}$$

Stellen wij voor alle S

$$h_\alpha(S) = h(S) + c(1-\alpha)S, \tag{6.173}$$

dan volgt uit (6.167), (6.170) en (6.172), dat de optimale s_- en S^+ gevonden kunnen worden uit de volgende twee betrekkingen:

$$h_\alpha(s_-)(1+\int_0^{S^+-s_-} w(u;\alpha)du) = h_\alpha(S^+) + \int_0^{S^+-s_-} h_\alpha(S^+-u)w(u;\alpha)du + K \tag{6.174}$$

en

$$(6.175)\qquad h'_\alpha(S^+) + \int_0^{S^+-s_-} h'_\alpha(S^+-u)w(u;\alpha)du = 0.$$

Deze twee niet-lineaire vergelijkingen kunnen in het algemeen alleen numeriek opgelost worden.

Toepassing op een exponentieel verdeelde vraag.

In deze toepassing heeft de vraag $\underline{b}$ een exponentiële kansdichtheid, zodat geldt

$$(6.176)\qquad w(u) = \lambda e^{-\lambda u}.$$

Wij nemen verder aan dat voor de inkoopkosten geldt

$$(6.177)\qquad \phi(q) = \begin{cases} cq + K & \text{voor } q > 0, \\ 0 & \text{voor } q = 0, \end{cases}$$

waarin $K \geqq 0$ is; we veronderstellen dat de voorraadkosten in een periode gegeven worden door

$$(6.178)\qquad h_1(y) = \begin{cases} c_1 y & \text{voor } y > 0, \\ 0 & \text{anders}, \end{cases}$$

als, nadat de vraag heeft plaatsgevonden, de voorraad aan het eind van die periode y is. Bovendien nemen wij aan dat de boetekosten in een periode gegeven worden door

$$(6.179)\qquad h_2(y) = \begin{cases} -c_2 y & \text{voor } y < 0, \\ 0 & \text{anders}, \end{cases}$$

als de achterstand aan het eind van die periode $-y$ is.

Voor de funktie h(S), zoals die gedefinieerd is op blz. 41, vinden wij

$$h(S) = c_1 \int_0^S (S-u)\lambda e^{-\lambda u} du + c_2 \int_S^\infty (u-S)\lambda e^{-\lambda u} du =$$

(6.180)

$$= c_1 S + \frac{c_1+c_2}{\lambda} e^{-\lambda S} - \frac{c_1}{\lambda} \qquad \text{voor } S > 0$$

en

(6.181) $$h(S) = c_2 \int_0^\infty (u-S)\lambda e^{-\lambda u} du = \frac{c_2}{\lambda} - c_2 S \qquad \text{voor } S \leqq 0.$$

Wij veronderstellen dat $c_2 > c(1-\alpha)$ is. De funktie h(S) is convex en differentieerbaar (vgl. fig. 6.4). De funktie $h_\alpha(S) = h(S) + c(1-\alpha)S$ neemt voor een eindige waarde van S een minimum aan. Dus een (S^+,s_-)-strategie is optimaal (vgl. IGLEHART).

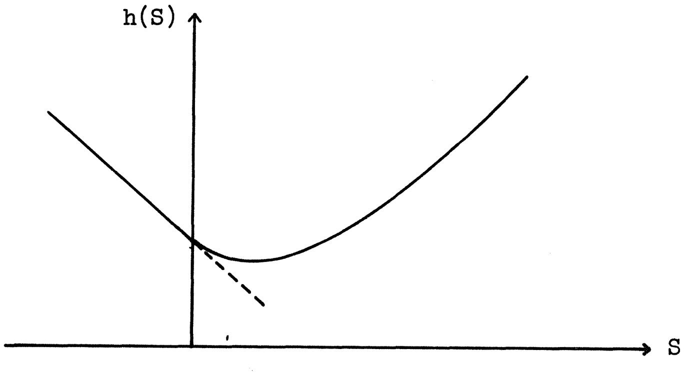

fig. 6.4

De funktie h(S)

Geval I (K = 0)

Uit (6.162) volgt dat de optimale s_- en S^+ gegeven worden door

(6.182) $$s_- = S^+ = \frac{1}{\lambda} \log \frac{c_1 + c_2}{c_1 + c(1-\alpha)} .$$

Geval II (K > 0)

Laten wij veronderstellen dat α, λ, K, c, c_1 en c_2 zodanige waarden hebben, dat de optimale $s_- > 0$ is. De kansdichtheid $w_j(u)$ is de dichtheid van de som van j onderling onafhankelijke stochastische variabelen, elk met dichtheid $\lambda e^{-\lambda u}$. Bijgevolg is $w_j(u)$ de dichtheid van een gamma-verdeelde stochastische variabele (vgl. deel 2, blz. 55), zodat geldt

$$w_j(u) = e^{-\lambda u}\, \lambda^j \frac{u^{j-1}}{(j-1)!} \qquad \text{voor } j \geqq 1. \tag{6.183}$$

Hieruit volgt

$$w(u;\alpha) = \alpha\lambda e^{-\lambda u(1-\alpha)}. \tag{6.184}$$

Na enig rekenwerk volgt uit (6.174)

$$\begin{aligned} &\frac{\lambda K(1-\alpha)^2}{d_\alpha} + e^{-\lambda(S^+-s_-)(1-\alpha)} \{\alpha + \frac{(1-\alpha)(c_1+c_2)}{d_\alpha} e^{-\lambda s_-}\} + \\ &+ \lambda(S^+-s_-)(1-\alpha) - \{\alpha + \frac{(1-\alpha)(c_1+c_2)}{d_\alpha} e^{-\lambda s_-}\} = 0, \end{aligned} \tag{6.185}$$

waarin

$$d_\alpha \overset{\text{def}}{=} c_1 + c(1-\alpha). \tag{6.186}$$

Uit (6.175) volgt na enig rekenwerk

$$e^{-\lambda(S^+-s_-)(1-\alpha)} \{\alpha + \frac{(1-\alpha)(c_1+c_2)}{d_\alpha} e^{-\lambda s_-}\} = 1. \tag{6.187}$$

Stel nu

$$A = \alpha + \frac{(1-\alpha)(c_1+c_2)}{d_\alpha} e^{-\lambda s_-}. \tag{6.188}$$

Bijgevolg geldt

$$s_- = \frac{1}{\lambda} \log \frac{(c_1+c_2)(1-\alpha)}{d_\alpha(A-\alpha)}. \tag{6.189}$$

Uit (6.187) en (6.188) volgt

(6.190) $$S^{+} - s_{-} = \frac{1}{\lambda(1-\alpha)} \log A.$$

Uit (6.185) t/m (6.190) volgt nu

(6.191) $$A = 1 + \frac{\lambda K(1-\alpha)^2}{d_\alpha} + \log A.$$

Vergelijking (6.191) heeft twee nulpunten als $K > 0$ is; de ene wortel is een getal tussen 0 en 1, terwijl de andere groter is dan 1. Wij vinden de optimale S^{+} en s_{-} door het grootste nulpunt van (6.191) te berekenen en de gevonden waarde vervolgens in (6.189) en (6.190) te substitueren. Wij moeten het grootste nulpunt bepalen, omdat voor het andere nulpunt log A negatief is, wat in tegenspraak is met $S^{+} \geqq s_{-}$ [vgl. (6.148)]

Hieronder zullen wij een numeriek voorbeeld geven. Stel

(6.192) $$\alpha = 0{,}96 \quad ; \quad \lambda = 0{,}02 \quad ; \quad c = c_1 = 1 \quad ; \quad c_2 = 15 \quad \text{en} \quad K = 30.$$

Met behulp van een rekenautomaat vinden wij

(6.193) $$A = 1{,}0436 \quad ; \quad S^{+} = 153{,}146 \quad \text{en} \quad s_{-} = 99{,}819.$$

Wij kunnen het grootste nulpunt van (6.191) benaderend bepalen, indien

(6.194) $$\frac{\lambda K(1-\alpha)^2}{d_\alpha} \approx 0.$$

Aangezien de vergelijking $A = 1 + \log A$ als enige oplossing $A = 1$ bezit, liggen de wortels van (6.191) dicht bij 1 als $\frac{\lambda K(1-\alpha)^2}{d_\alpha}$ ongeveer gelijk is aan nul. Voor $A \approx 1$ geldt *)

(6.195) $$\log A \approx (A-1) - \frac{(A-1)^2}{2}.$$

*) $$\log x \overset{\text{def}}{=} \int_1^x \frac{dt}{t} = \int_0^{x-1} \frac{du}{1+u} \approx \int_0^{x-1} (1-u)du \quad \text{als } x \approx 1.$$

Uit (6.191) en (6.195) volgt

(6.196) $$A \approx 1 + (1-\alpha)\sqrt{\frac{2\lambda K}{d_\alpha}}\ .$$

Uit (6.189), (6.190), (6.191) en (6.196) volgt nu

(6.197) $$s_- \approx \frac{1}{\lambda} \log \frac{c_1 + c_2}{d_\alpha + \sqrt{2\lambda K d_\alpha}}$$

en

(6.198 $$S^+ - s_- \approx \sqrt{\frac{2K}{\lambda d_\alpha}} - \frac{(1-\alpha)K}{d_\alpha}\ .$$

Wij zullen hieronder bij het voorafgaande nog enige opmerkingen plaatsen, die betrekking hebben op diverse andere voorraadmodellen.

Opmerking 6.2 (S^+,s_-)-voorraadmodel met noodinkoop, levertijd nul en een continu verdeelde vraag.

Wij veronderstellen nu dat de voorraadbeheerder een noodinkoop verricht, zodra de vraag de aanwezige voorraad te boven gaat. Deze noodinkoop dient uitsluitend om het tekort op te heffen en elke noodinkoop brengt ex ra kosten met zich mee. Indien in een periode de vraag de voorraad overschrijdt, is de voorraad aan het einde van die periode nul. De voorraadbeheerder kan alleen aan het begin van een periode bestellingen plaatsen; de bestellingen worden zonder levertijd geleverd. De inkoopkosten zijn een gegeven funktie $\phi(q)$ van de bestelgrootte q. De verwachting van de voorraad- en noodinkoopkosten, op te lopen in een periode, is een gegeven funktie h(S) van de voorraad S, die na een eventuele aanvulling aan het begin van die periode aanwezig is.

Wij zullen aantonen, dat bij afwezigheid van een levertijd het noodinkoopmodel op dezelfde wijze behandeld kan worden als het reeds behandelde naleveringsmodel. Daartoe merken wij eerst op dat de formules (6.152) t/m (6.157) ook voor het hier beschouwde noodinkoopmodel gelden, mits wij in deze formules $\phi(q)$ vervangen door *)

$$(6.199) \qquad \phi^*(q) \stackrel{\mathrm{def}}{=} \begin{cases} \phi(q) & \text{voor } q < S^+, \\ \\ \phi(S^+) & \text{voor } q \geq S^+. \end{cases}$$

Indien

$$(6.200) \qquad \phi(q) = \begin{cases} cq + K & \text{voor } q > 0, \\ \\ 0 & \text{voor } q = 0, \end{cases}$$

waarin $K \geq 0$ is, dan geldt voor $S \geq s_-$ en $0 \leq u \leq S - s_-$ [vgl. (6.156)]

$$\alpha \int_{S-s_--u}^{\infty} \phi^*(S^+-S+u+v)w(v)dv =$$

$$(6.201) \qquad = \alpha \int_{S-s_--u}^{S-u} \left[c(S^+-S+u+v)+K\right]w(v)dv + \alpha \int_{S-u}^{\infty} (cS^++K)w(v)dv =$$

$$= \alpha \int_{S-s_--u}^{\infty} \left[c(S^+-S+u+v)+K\right]w(v)dv - \alpha c \int_{S-u}^{\infty} (v-S+u)w(v)dv.$$

Een nadere analyse van (6.156) en (6.157) leert ons dat de formules (6.162) t/m (6.175) ook voor het noodinkoopmodel gelden, indien wij $h(S)$ vervangen door

$$(6.202) \qquad h^*(S) \stackrel{\mathrm{def}}{=} h(S) - \alpha c \int_{S}^{\infty} (v-S)w(v)dv.$$

Wij kunnen dus concluderen dat, wanneer er geen levertijd is en we met een lineaire inkoopfunktie te doen hebben, het noodinkoopmodel kan worden opge-

---

*) Bedenk dat $S \geq 0$ is voor het noodinkoopmodel.

vat als naleveringsmodel, als wij h(S) vervangen door $h^*(S)$.

Opmerking 6.3 (S^+,s_-)-voorraadmodel met nalevering, een vaste levertijd *) en een continu verdeelde vraag.

Beschouw wederom het voorraadmodel met nalevering. Uitsluitend aan het begin van elke periode kan een bestelling geplaatst worden. Een bestelling geplaatst aan het begin van periode n wordt afgeleverd aan het begin van periode n+L, waarbij L een vast, niet-negatief geheel getal is. De vraag vindt aan het eind van elke periode plaats en heeft in elke periode dezelfde kansverdeling met dichtheid w(u). De inkoopkosten bedragen $\phi(q)$ als de bestelgrootte q is en worden bij aflevering van de geplaatste bestelling verrekend (door een eventuele verdiscontering kunnen wij er altijd voor zorgen dat deze veronderstelling vervuld is). Wij nemen wederom aan dat de verwachte voorraad- en boetekosten in een periode gelijk zijn aan h(y) als de voorraad aan het begin van die periode na een eventuele binnenkomst van een order gelijk aan y is.

Wij voeren nu een belangrijk begrip uit de voorraadtheorie in en wel de economische voorraad, die gedefinieerd wordt als de som van de voorraad (negatieve voorraad komt overeen met achterstand) en de in bestelling zijnde hoeveelheden. Wij vermelden zonder bewijs dat voor het naleveringsmodel elke optimale strategie een funktie van de economische voorraad is (vgl. het in de voetnoot op blz. 42 vermelde artikel van SCARF). Vandaar dat wij de toestand van het systeem definiëren als de grootte van de economische voorraad. Wij identificeren elke beslissing X met de grootte van de economische voorraad vlak na beslissen; dus beslissing X in toestand S betekent een bestelling ter grootte van X-S.

Als de levertijd $L \geq 1$ is, kunnen wij stellen dat, welke beslissing de voorraadbeheerder ook neemt aan het begin van een periode, de kosten in die periode en de L-1 daarop volgende perioden niet meer beïnvloed kunnen worden. De beslissing heeft alleen invloed op de L perioden later te maken

---

*) Het noodinkoopmodel vereist een geheel andere aanpak dan het naleveringsmodel ingeval de levertijd positief is.

kosten. Verder constateren wij dat, als aan het begin van een periode n de beslissing X genomen wordt en de cumulatieve vraag in de perioden $n,\ldots,n+L-1$ gelijk aan u is, de voorraad aan het begin van periode n+L gelijk aan X-u is.

Laten wij de kostenfunktie $h_L(X)$ definiëren als de verwachte voorraad- en boetekosten in periode n+L, gegeven dat aan het begin van periode n de economische voorraad vlak <u>na</u> beslissen gelijk aan X is. Uit het bovenstaande zal duidelijk zijn dat

$$(6.203) \qquad h_L(X) = \begin{cases} \displaystyle\int_0^\infty h(X-u)w_L(u)du & \text{als } L \geqq 1, \\ h(X) & \text{als } L = 0, \end{cases}$$

waarin $w_L(u)$ de kansdichtheid van de totale vraag in L perioden voorstelt. De verwachte kosten in periode n+L, gegeven dat aan het begin van periode n in toestand S de beslissing X genomen wordt, zijn derhalve gelijk aan*)

$$(6.204) \qquad h_L(S,X) \overset{\text{def}}{=} \begin{cases} h_L(S) & \text{als } X = S, \\ \phi(X-S) + h_L(X) & \text{als } X > S. \end{cases}$$

Veronderstel dat de voorraadbeheerder de strategie $z = (S^+,s_-)$ toepast; d.w.z. "plaats aan het begin van een periode alleen dan een bestelling als de economische voorraad S kleiner dan s_- is en bestel in dat geval een hoeveelheid S^+-S."

Omdat de kosten in de eerste L perioden niet beïnvloed kunnen worden, definiëren wij $f(S;z)$ als de verwachting van de totale verdisconteerde kosten in de perioden $L+1,L+2,\ldots,$ als aan het begin van periode 1 de economische voorraad vlak voor beslissen gelijk aan S is (vgl. fig. 6.5). Wij verdisconteren de kosten naar het begin van periode L+1 i.p.v. periode 1 (dit maakt slechts een factor α^L uit).

---

*) Bedenk dat de bestelkosten bij aflevering verrekend worden.

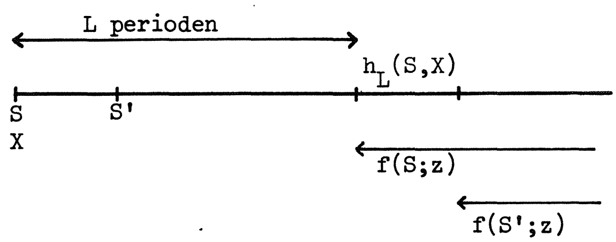

fig. 6.5

De instrumenten voor het voorraadmodel met een levertijd $L \geq 1$.

Men gaat gemakkelijk na dat de resultaten behorende bij het naleveringsmodel met een levertijd van L perioden verkregen worden door in de formules (6.152) t/m (6.175) de funktie h(S) te vervangen door $h_L(S)$.

Voor het beschouwde voorraadprobleem bestaat een optimale strategie die tot de klasse van de (S^+,s_-)-strategieën behoort, als $h_L(S)$ convex is, $\phi(0) = 0$ en $\phi(q) = cq + K$ voor $q > 0$, waarin $K \geq 0$ is; voor $K = 0$ zijn de optimale s_- en S^+ aan elkaar gelijk (vgl. IGLEHART). Met behulp van (6.162) volgt dat voor $K = 0$ de optimale $s_- = S^+$ voldoet aan

$$h_L'(s_-) + c(1-\alpha) = 0. \tag{6.205}$$

Deze formule kunnen wij nog verder vereenvoudigen als de voorraad- resp. boetekosten in een periode gelijk zijn aan $c_1 \max(y,0)$ resp. $-c_2 \min(y,0)$, waarin y de voorraad aan het eind van de beschouwde periode is, nadat de vraag heeft plaatsgevonden. Voor de funktie $h_L(S)$ (die gedefinieerd is als de verwachte voorraad- en boetekosten in periode n+L, gegeven dat aan het begin van periode n de economische voorraad vlak na beslissen S is) vinden wij dan

$$h_L(S) = \begin{cases} c_1 \int_0^S (S-u)w_{L+1}(u)du + c_2 \int_S^\infty (u-S)w_{L+1}(u)du & \text{voor } S \geq 0, \\ -c_2(S-\mu(L+1)) & \text{voor } S < 0, \end{cases} \tag{6.206}$$

waarin μ de verwachte vraag per periode is. Nemen wij aan dat $c_2 > c(1-\alpha)$ is, dan vinden wij uit (6.205) en (6.206) dat voor $K = 0$ de optimale $s_- = S^+$ voldoet aan

$$W_{L+1}(s_-) = \frac{c_2 - c(1-\alpha)}{c_1 + c_2}, \tag{6.207}$$

waarin $W_k(u)$ de verdelingsfunktie is van de totale vraag in k perioden.

Opmerking 6.4 (S^+,s_-)-voorraadmodellen met een discreet verdeelde vraag.

Wij veronderstellen nu dat de behoeften in de opeenvolgende perioden onderling onafhankelijke stochastische variabelen zijn, die eenzelfde discrete kansverdeling bezitten. De kans op een behoefte van i eenheden in een periode wordt gegeven door $p(i)$. Vanzelfsprekend moet voldaan zijn aan

$$\sum_{i=0}^{\infty} p(i) = 1. \tag{6.208}$$

Wij definiëren $p(i;\alpha)$ door (6.145). Uit (6.144) en (6.145) volgt

$$p(i;\alpha) = \alpha p(i) + \alpha \sum_{r=0}^{i} p(i-r;\alpha)p(r) \qquad \text{voor } i \geqq 0. \tag{6.209}$$

Wij nemen aan dat de inkoopfunktie $\phi(i)$ gegeven wordt door

$$\phi(i) = \begin{cases} ci + K & \text{voor } i = 1,2,\ldots, \\ 0 & \text{voor } i = 0, \end{cases} \tag{6.210}$$

waarin $K \geqq 0$ is. Wij nemen wederom aan dat de bestelkosten bij aflevering worden verrekend. De funktie $h(S)$ heeft dezelfde betekenis als in de voorgaande modellen. Beschouw nu het naleveringsmodel met een vaste levertijd van L perioden, waarbij L geheel en niet-negatief is. Definiëren wij $h_L(S)$ als de verwachte voorraad- en boetekosten in periode n+L, gegeven dat aan het begin van periode n de economische voorraad vlak na beslissen gelijk aan S is, dan geldt

$$(6.211) \qquad h_L(S) = \begin{cases} h(S) & \text{voor } L = 0, \\ \sum_{k=0}^{\infty} h(S-k)p_L(k) & \text{voor } L \geqq 1. \end{cases}$$

waarbij $p_L(k)$ de kans is op een totale behoefte van k eenheden in L perioden [vgl. (6.144)]. Op dezelfde wijze als in voorgaande voorraadmodellen kunnen wij aantonen dat voor begintoestand S en strategie $z = (S^+,s_-)$ de verwachting van de totale verdisconteerde kosten in de perioden $L+1, L+2, \ldots$, gegeven wordt door

$$(6.212) \qquad f(S;z) = \begin{cases} \dfrac{h_\alpha(S^+) + \sum_{k=0}^{S^+-s_-} h_\alpha(S^+-k)p(k;\alpha) + K}{(1-\alpha)(1+\sum_{k=0}^{S^+-s_-} p(k;\alpha))} + \dfrac{\alpha\mu c}{1-\alpha} - cS & \text{voor } S < s_-, \\[2ex] \dfrac{\{\alpha-(1-\alpha)\sum_{k=0}^{S-s_-} p(k;\alpha)\}\{h_\alpha(S^+) + \sum_{k=0}^{S^+-s_-} h_\alpha(S^+-k)p(k;\alpha) + K\}}{(1-\alpha)(1+\sum_{k=0}^{S^+-s_-} p(k;\alpha))} + h_\alpha(S) + \\ \quad + \sum_{k=0}^{S-s_-} h_\alpha(S-k)p(k;\alpha) + \dfrac{\alpha\mu c}{1-\alpha} - cS & \text{voor } S \geqq s_-, \end{cases}$$

waarin

$$(6.213) \qquad h_\alpha(S) \stackrel{\text{def}}{=} h_L(S) + c(1-\alpha)S.$$

De grootheid $p(k;\alpha)$ wordt door (6.145) gedefinieerd en μ is de verwachte vraag per periode.

Voor het beschouwde voorraadprobleem bestaat een optimale strategie die tot de klasse van de (S^+,s_-)-strategieën behoort, als er een geheel getal S_0 bestaat zodat de funktie $h_\alpha(S)$ monotoon dalend is voor $S \leqq S_0$ en monotoon stijgend voor $S \geqq S_0$; voor $K = 0$ zijn de optimale s_- en S^+ beide gelijk aan S_0.*) Een getal S_0 met bovengenoemde eigenschap bestaat als de voorraad- resp. boetekosten in een periode gelijk zijn aan $c_1 \max(y,0)$ resp. $-c_2 \min(y,0)$, waarin y de voorraad aan het eind van de beschouwde periode is, nadat de vraag heeft plaatsgevonden. Wij nemen aan dat $c_2 > c(1-\alpha)$ is. Op dezelfde wijze als (6.207) afgeleid is, kan worden aangetoond dat S_0 voldoet aan

$$\sum_{k=0}^{S_0-1} p_{L+1}(k) < \frac{c_2 - c(1-\alpha)}{c_1 + c_2} \leq \sum_{k=0}^{S_0} p_{L+1}(k). \tag{6.214}$$

Voor het noodinkoopmodel met levertijd <u>nul</u> kunnen wij op overeenkomstige wijze te werk gaan. De bijbehorende resultaten verkrijgen wij door in (6.212) en (6.213) de funktie h(S) te vervangen door [vgl. (6.202)]

$$h^*(S) \stackrel{\text{def}}{=} h(S) - \alpha c \sum_{k=S}^{\infty} (k-S)p(k). \tag{6.215}$$

Dit betekent voor formule (6.214) dat c_2 daarin vervangen moet worden door $c_2 - \alpha c$, als de noodinkoopkosten per eenheid gelijk zijn aan c_2.

Vervolgens geven wij een toepassing van de oplossingsmethode geschetst onder d) [vgl. blz. 18 e.v.; iteratiemethode voor strategieverbetering].

<u>Voorbeeld 6.3</u> <u>Een reparateursprobleem (methode d).</u>

Eén van de servicewagens, van de "Tweede Nederlandse Fabriek van Rekenautomaten N.V.", gevestigd te Eindhoven, bezoekt wekelijks 5 steden. Maandag is de wagen in Groningen (1), dinsdag vertoeft de reparateur in Amsterdam (2), woensdag en donderdag behartigt hij de belangen van de firma in resp. Rotterdam (3) en Antwerpen (4), vrijdag tenslotte doet hij Aken (5) aan. De reparateur kan op zijn rondreis ten hoogste 5 reserve-onderdelen U507A

*) Vgl. A.F. VEINOTT jr. and H.M. WAGNER, Computing optimal (s,S) inventory policies, Management Science, <u>11</u>, (1965), p. 525-552 of E.L. JOHNSON, On (s,S) policies, Management Science, <u>15</u>, (1968), p. 80-101.

meenemen. Indien hij aan het eind van een dag zijn voorraad ontoereikend vindt, kan hij de avond benutten om naar Eindhoven te gaan voor het aanvullen van zijn voorraad. Als vanuit stad i een dergelijke rit naar Eindhoven wordt ondernomen, bedragen de kosten a_i, ongeacht de omvang van de aanvulling (zie tabel 6.1). De kans op een behoefte aan r reserveonderdelen in stad i is $p_r^{(i)}$ (zie tabel 6.1). De maximale behoefte in iedere stad is twee onderdelen U507A.

Wanneer de reparateur onderweg tekort komt, wordt de volgende dag vanuit Eindhoven een speciale reparateur gezonden met de ontbrekende onderdelen; de kosten hiervan bedragen f 200,-, ongeacht de omvang van het tekort en de stad, waar de reparatie moet plaatsvinden. De rondreizende reparateur wacht daar echter niet op, maar vervolgt zijn reis al of niet via Eindhoven; de speciale reparateur kan de benodigde reparaties zelfstandig verrichten.

In dit voorbeeld wordt voor de verdisconteringsfactor α de waarde 0,97 gekozen.

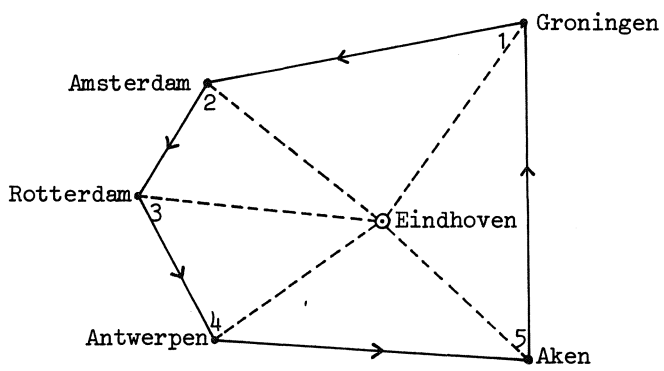

fig. 6.6

Route van de servicewagen

$i \backslash r$	0	1	2	a_i
1	0,5	0,3	0,2	60
2	0,25	0,5	0,25	30
3	0,375	0,375	0,25	50
4	0,3	0,5	0,2	25
5	0,5	0,25	0,25	100

Tabel 6.1

De kansen $p_r^{(i)}$ en de kosten a_i

Oplossing

De toestand van het systeem wordt bepaald door de voorraad aan het eind van de dag en door de stad waar de reparateur zich bevindt. Wij kunnen de toestandsgrootheid S geven door de twee toestandscomponenten i ($\equiv$ stad) en j ($\equiv$ omvang van de voorraad). De component i neemt de waarden 1,...,5 aan en j de waarden 0,...,5. De op het eind van de dag te nemen beslissing X identificeren wij met de omvang van de voorraad na een eventuele aanvulling. De verzameling $\mathfrak{X}(S)$ van toegelaten beslissingen in toestand $S = (i,j)$ wordt derhalve gegeven door

(6.216) $$\mathfrak{X}(S) = \{X \mid j \leq X \leq 5\}.$$

Wij definiëren

(6.217) $$a_{ij}(X) = \begin{cases} a_i & \text{als } X > j. \\ 0 & \text{anders.} \end{cases}$$

De "opbrengst"-funktie h(S,X) heeft betrekking op de kosten van een eventuele omweg van stad i via Eindhoven naar stad i+1 en de mogelijke komst van een speciale reparateur naar stad i+1. De funktie h(S,X) wordt derhalve gegeven door

$$(6.218)\qquad h((i,j),X) = \begin{cases} -a_{ij}(X) & \text{als } X \geqq 2, \\ -a_{ij}(X) - 200 \sum_{k=X+1}^{2} p_k^{(i+1)} & \text{als } X < 2, \end{cases}$$

waarin

$$(6.219)\qquad p_k^{(6)} \overset{\text{def}}{=} p_k^{(1)} \qquad \text{voor } k = 0,1,2.$$

Voor de kansverdeling van de toestand $\underline{S}'$ op het volgende beslissingstijdstip geldt

$$(6.220)\qquad p((i+1,m)|(i,j),X) = \begin{cases} p_{X-m}^{(i+1)} & \text{als } X \geqq 2 \text{ en } X-2 \leqq m \leqq X, \\ p_0^{(i+1)} & \text{als } X = m = 1, \\ \sum_{k=X}^{2} p_k^{(i+1)} & \text{als } 0 \leqq X \leqq 1 \text{ en } m = 0, \\ 0 & \text{anders}. \end{cases}$$

De funktionaalvergelijking (6.69) gaat over in

$$(6.221)\qquad f((i,j);z) = h((i,j),z((i,j))) + \alpha \sum_{k=0}^{2} p_k^{(i+1)} f((i+1,z((i,j))-k);z)$$
$$\text{voor } i = 1,\dots,5\ ;\ j = 0,\dots,5,$$

waarin

$$(6.222)\qquad f((i,j);z) \overset{\text{def}}{=} f((i,0);z) \qquad \text{als } j < 0.$$

De funktie $f((i,j);z)$ vinden wij dus door een stelsel van 30 lineaire vergelijkingen in 30 onbekenden op te lossen. Op blz. 21 hebben wij aangetoond dat de oplossing van dit stelsel uniek is. Wanneer wij voor een strategie z_m de funktie $f((i,j);z_m)$ berekend hebben, bepalen wij een nieuwe strategie z_{m+1}. Wij kiezen voor iedere toestand (i,j) een beslissing X, waarvoor

Tabel 6.2

Berekening van een optimale strategie

	Iteratie 1			Iteratie 2			Iteratie 3			Iteratie 4			Iteratie 5		
S=(i,j)	z_1	$f(S;z_1)$	$f(S;(z_2)z_1)$	z_2	$f(S;z_2)$	$f(S;(z_3)z_2)$	z_3	$f(S;z_3)$	$f(S;(z_4)z_3)$	z_4	$f(S;z_4)$	$f(S;(z_5)z_4)$	z_5	$f(S;z_5)$	$f(S;(z_6)z_5)$
(1,0)	2	-1017,22	-936,28	5	-518,24	-518,24	5	-337,12	-337,12	5	-306,13	-306,13	5	-303,69	-303,69
(1,1)	1	-1014,50	-936,28	5	-518,24	-515,52	1	-346,39	-337,12	5	-306,13	-306,13	5	-303,69	-303,69
(1,2)	2	- 957,22	-936,28	5	-518,24	-465,52	2	-296,39	-296,39	2	-268,93	-268,93	2	-266,08	-266,08
(1,3)	3	- 936,12	-936,12	3	-465,52	-465,52	3	-296,39	-296,39	3	-265,40	-265,40	3	-262,16	-262,16
(1,4)	4	- 907,42	-907,42	4	-465,52	-465,52	4	-290,39	-290,39	4	-255,88	-255,88	4	-253,04	-253,04
(1,5)	5	- 876,28	-876,28	5	-458,24	-458,24	5	-277,12	-277,12	5	-246,13	-246,13	5	-243,69	-243,69
(2,0)	2	- 994,33	-900,28	5	-479,91	-479,91	5	-305,55	-305,55	5	-277,25	-277,25	5	-274,72	-274,72
(2,1)	2	- 994,33	-900,28	5	-479,91	-479,91	5	-305,55	-305,55	5	-277,25	-277,25	5	-274,72	-274,72
(2,2)	2	- 964,33	-900,28	5	-479,91	-479,91	5	-305,55	-305,55	5	-277,25	-275,62	2	-273,09	-273,09
(2,3)	3	- 937,32	-900,28	5	-479,91	-479,91	5	-305,55	-292,10	3	-262,70	-262,70	3	-260,18	-260,18
(2,4)	4	- 902,97	-900,28	5	-479,91	-468,10	4	-280,83	-280,83	4	-252,52	-252,52	4	-250,00	-250,00
(2,5)	5	- 870,28	-870,28	5	-449,91	-449,91	5	-275,55	-275,55	5	-247,25	-247,25	5	-244,72	-244,72
(3,0)	2	-1013,93	-915,67	5	-498,60	-498,60	5	-326,64	-326,64	5	-297,46	-297,46	5	-294,85	-294,85
(3,1)	1	-1011,20	-915,67	5	-498,60	-495,87	1	-331,13	-326,64	5	-297,46	-297.46	5	-294,85	-294,85
(3,2)	2	- 963,93	-915,67	5	-498,60	-455,87	2	-291,13	-291,13	2	-261,95	-261,95	2	-259,36	-259,36
(3,3)	3	- 938,76	-915,67	5	-498,60	-455,87	3	-291,13	-291,13	3	-261,95	-261,95	3	-259,36	-259,36
(3,4)	4	- 901,01	-901,01	4	-455,87	-455,87	4	-286,80	-286,80	4	-257,62	-257,62	4	-255,02	-255,02
(3,5)	5	- 865,67	-865,67	5	-448,60	-448,60	5	-276,64	-276,64	5	-247,46	-247,46	5	-244,85	-244,85
(4,0)	2	-1001,24	-890,46	5	-469,97	-469,97	5	-300,14	-300,14	5	-270,05	-270,05	5	-267,38	-267,38
(4,1)	2	-1001,24	-890,46	5	-469,97	-469,97	5	-300,14	-300,14	5	-270,05	-270,05	5	-267,38	-267,38
(4,2)	2	- 976,24	-890,46	5	-469,97	-469,97	5	-300,14	-300,14	5	-270,05	-270,05	5	-267,38	-267,38
(4,3)	3	- 931,44	-890,46	5	-469,97	-469,97	5	-300,14	-300,14	5	-270,05	-270,05	5	-267,38	-267,38
(4,4)	4	- 893,04	-890,46	5	-469,97	-461,57	4	-285,26	-285,26	4	-255,18	-255,18	4	-252,47	-252,47
(4,5)	5	- 865,46	-865,46	5	-444,97	-444,97	5	-275,14	-275,14	5	-245,05	-245,05	5	-242,38	-242,38
(5,0)	0	-1086,71	-970,66	5	-548,02	-548,02	5	-376,41	-376,41	5	-345,32	-345,32	5	-342,68	-342,68
(5,1)	1	-1025,39	-970,66	5	-548,02	-542,69	1	-371,50	-371,50	1	-336,95	-336,95	1	-334,58	-334,58
(5,2)	2	- 956,81	-956,81	2	-502,69	-502,69	2	-309,94	-309,94	2	-278,91	-278,91	2	-276,34	-276,34
(5,3)	3	- 929,39	-929,39	3	-477,12	-477,12	3	-297,20	-297,20	3	-266,37	-266,37	3	-263,50	-263,50
(5,4)	4	- 898,21	-898,21	4	-461,78	-461,78	4	-284,59	-284,59	4	-253,51	-253,51	4	-250,63	-250,63
(5,5)	5	- 870,66	-870,66	5	-448,02	-448,02	5	-276,41	-276,41	5	-245,32	-245,32	5	-242,68	-242,68

$$(6.223)\qquad h((i,j),X) + \alpha \sum_{k=0}^{2} p_k^{(i+1)} f((i+1,X-k);z_m)$$

maximaal is. Ter wille van de convergentie spreken wij af dat wij $X = z_m((i,j))$ kiezen als deze beslissing (6.223) maximaliseert. Op deze wijze verkrijgen wij een strategie z_{m+1}. Het maximum van (6.223) is gelijk aan $f((i,j);(z_{m+1})z_m)$. Na een eindig aantal iteratiestappen, stel M, vinden wij dat $z_M = z_{M+1}$; deze strategie z_M is dan optimaal.

Wij beginnen de iteratiemethode met een strategie z_1, waarbij wij voor een $z_1(S)$ een beslissing X nemen, die $h(S,X)$ maximaliseert. Met de rekenautomaat hebben wij een optimale strategie bepaald, die na M = 5 stappen werd gevonden. De vijf iteratiestappen zijn in tabel 6.2 samengevat. De strategie z_5 is optimaal.

Tenslotte geven wij een toepassing van de oplossingsmethode geschetst onder e) (vgl. blz. 22 e.v.).

Voorbeeld 6.4 Een vervangingsprobleem (methode e).

Een bedrijfsleider staat aan het begin van elke maand voor de vraag of hij een machine van een bepaald type zal houden of inruilen tegen een machine van hetzelfde type, maar eventueel van een andere leeftijd. Indien hij tot het laatste overgaat, moet hij tevens beslissen wat de leeftijd van de vervangende machine zal moeten zijn. De leeftijd van een machine wordt in maanden uitgedrukt. De bedrijfsleider vervangt de machine altijd als deze de leeftijd n heeft, waarbij n een gegeven getal is.

Van een i maanden oude machine is de inruilwaarde c_i en de aanschafwaarde d_i. Als aan het begin van een maand na een eventuele inruil een machine van de leeftijd i aanwezig is, bedragen de verwachte onkosten (onderhoud e.d.) aan die machine in de betreffende maand e_i. Een i maanden oude machine is met kans p_i aan het eind van de maand in een zodanig slechte staat, dat de bedrijfsleider wel tot vervanging moet overgaan. Wij stellen de leeftijd van de machine dan op n en de inruilwaarde op c_n; m.a.w. wanneer de leeftijd van de machine aan het begin van de maand i is met $0 \leq i < n-1$, dan is de leeftijd aan het begin van de volgende maand i+1 met kans $1-p_i$ en gelijk aan n met kans p_i; een machine van n-1 maanden is aan het begin van de volgende maand n maanden oud. Wij kunnen dus veronderstellen dat $p_{n-1} = 1$ is.

Oplossing

De toestand van het systeem (de machine) wordt gegeven door de leeftijd van de aan het begin van de maand aanwezige machine. De toestandsgrootheid i kan de waarden 0,1,...,n aannemen. De beslissing om de machine te behouden geven we aan met X = H en de beslissing om de machine te vervangen door een a maanden oude machine met X = a. De verzameling $\mathcal{X}(i)$ van toegelaten beslissingen in toestand i wordt gegeven door

$$\mathcal{X}(i) = \begin{cases} \{0,1,\ldots,n-1,H\} & \text{voor } 0 \leq i \leq n-1, \\ \{0,1,\ldots,n-1\} & \text{voor } i = n. \end{cases} \tag{6.224}$$

De opbrengstfunktie h(i,X) voldoet aan

$$h(i,X) = \begin{cases} c_i - d_a - e_a & \text{als } X = a;\ 0 \leq a \leq n-1;\ 0 \leq i \leq n, \\ -e_i & \text{als } X = H;\ 0 \leq i < n. \end{cases} \tag{6.225}$$

Voor de overgangskansen $p_{ij}(X)$ geldt

$$p_{ij}(X) = \begin{cases} 1-p_a & \text{als } X = a \text{ en } j = a+1;\ 0 \leq a < n-1;\ 0 \leq i \leq n, \\ p_a & \text{als } X = a \text{ en } j = n\ ;\ 0 \leq a < n-1;\ 0 \leq i \leq n, \\ 1 & \text{als } X = n-1 \text{ en } j = n;\ 0 \leq i \leq n, \\ 1-p_i & \text{als } X = H \text{ en } j = i+1;\ 0 \leq i < n-1, \\ p_i & \text{als } X = H \text{ en } j = n\ ;\ 0 \leq i < n-1, \\ 1 & \text{als } X = H,\ \ j = n \text{ en } i = n-1, \\ 0 & \text{anders.} \end{cases} \tag{6.226}$$

Wij kunnen een optimale strategie aflezen uit een optimale basisoplossing van het lineaire programmeringsprobleem (vgl. (6.85) op blz. 24): "maximaliseer

$$\sum_{i=0}^{n} \sum_{a=0}^{n-1} h(i,a)v_{ia} + \sum_{i=0}^{n-1} h(i,H)v_{iH} \tag{6.227}$$

onder de voorwaarden

$$\left\{\begin{array}{ll} \sum_{a=0}^{n-1} v_{ja} + v_{jH} - \alpha \sum_{i=0}^{n} \sum_{a=0}^{n-1} p_{ij}(a)v_{ia} - \alpha \sum_{i=0}^{n-1} p_{ij}(H)v_{iH} = 1 & \\ & \text{voor } j = 0,\ldots,n, \\ v_{jX} \geqq 0 & \text{voor } X \in \mathfrak{X}(j) \quad \text{en} \quad j = 0,\ldots,n." \end{array}\right. \tag{6.228}$$

Bovenstaand lineair programmeringsprobleem bevat $n(n+1) + n = n^2 + 2n$ variabelen en n+1 bijvoorwaarden. Wij zullen nu aantonen, dat een gereduceerd lineair programmeringsprobleem (vgl. (6.95) op blz. 27) geformuleerd kan worden met aanzienlijk minder variabelen; hierin komen namelijk 3n+1 variabelen en n+2 bijvoorwaarden voor.

Om aan de voorwaarden (6.88) t/m (6.90) te voldoen, kiezen wij

$$\mathcal{A}(i) = \{0,1,\ldots,n-1\} \qquad \text{voor } 0 \leqq i \leqq n, \tag{6.229}$$

$$c(i) = c_i \qquad \text{voor } 0 \leqq i \leqq n, \tag{6.230}$$

$$d(X) = -d_X - e_X \qquad \text{voor } X = a \text{ en } 0 \leqq a \leqq n-1, \tag{6.231}$$

en

$$p(j,X) = \left\{\begin{array}{ll} 1 - p_a & \text{voor } X = a,\ j = a+1 \text{ en } 0 \leqq a < n-1, \\ p_a & \text{voor } X = a,\ j = n \quad \text{en } 0 \leqq a < n-1, \\ 1 & \text{voor } j = n \text{ en } X = n-1, \\ 0 & \text{anders.} \end{array}\right. \tag{6.232}$$

Aan de voorwaarden (6.88) t/m (6.90) is door bovenstaande keuze voldaan. De op blz. 26 gedefinieerde grootheden voldoen aan

$$(6.233) \qquad \mathfrak{B}(1) = \begin{cases} \{H\} & \text{voor } 0 \leqq i \leqq n-1, \\ \emptyset & \text{voor } i = n, \end{cases}$$

$$(6.234) \qquad k = n+1 \;;\; m = 1 \;;\; j(i) = 1 \qquad \text{voor } 0 \leqq i \leqq n,$$

en

$$(6.235) \qquad \mathfrak{C}(1) = \mathfrak{D}(1) = \mathfrak{A}(1).$$

Het gereduceerde lineaire programmeringsprobleem (6.95) kunnen wij met (6.229) t/m (6.235) schrijven als:

"maximaliseer

$$(6.236) \qquad -\sum_{i=0}^{n-1} e_i\, v_{iH} + \sum_{i=0}^{n} c_i r_i - \sum_{a=0}^{n-1} (d_a + e_a) w_a$$

onder de voorwaarden

$$(6.237) \quad \begin{cases} v_{0H} + r_0 = 1 \\ v_{iH} - \alpha(1-p_{i-1})v_{i-1,H} + r_i - \alpha(1-p_{i-1})w_{i-1} = 1 \quad \text{voor } i = 1,\dots,n-1 \\ -\alpha \sum_{j=0}^{n-1} p_j\, v_{jH} + r_n - \alpha \sum_{a=0}^{n-1} p_a\, w_a = 1 \\ -\sum_{j=0}^{n} r_j + \sum_{a=0}^{n-1} w_a = 1 \\ r_j \geqq 0,\ v_{iH} \geqq 0,\ w_i \geqq 0 \quad \text{voor } j = 0,\dots,n \text{ en } i = 0,\dots,n-1." \end{cases}$$

Voor een numerieke toepassing nemen wij $n = 10$ en $\alpha = 0{,}97$. De overige benodigde getallen staan vermeld in tabel 6.3.

i	0	1	2	3	4	5	6	7	8	9	10
c_i	850	725	625	540	460	400	330	270	200	125	25
d_i	1000	875	750	650	570	500	400	340	280	220	-
e_i	5	30	35	40	45	50	55	65	75	85	-
p_i	0	0,01	0,02	0,04	0,08	0,14	0,22	0,32	0,5	1	-

Tabel 6.3

Getallen voor een numerieke toepassing van voorbeeld 6.4

Met behulp van de rekenautomaat hebben wij een optimale basisoplossing van (6.236) berekend. Van de gevonden oplossing worden de basisvariabelen (die alle positief zijn) gegeven door

(6.238) $\quad v_{0H}, v_{1H}, v_{2H}, v_{3H}, v_{4H}, v_{5H}, v_{6H}, r_7, r_8, r_9, r_{10}, w_0.$

Bijgevolg is de strategie (vgl. blz. 28): "vervang de machine alleen als de leeftijd 7 maanden of hoger is en koop dan een machine van de leeftijd 0 (dus een nieuwe machine)" optimaal. De funktiewaarden van f(i;z), waarbij z een optimale strategie is, worden gegeven in onderstaande tabel (vgl. de voetnoot op blz. 28).

i	f(i;z)	i	f(i;z)
0	-5013,71	6	-5678,68
1	-5163,62	7	-5743,71
2	-5285,36	8	-5813,71
3	-5400,99	9	-5888,71
4	-5507,54	10	-5988,71
5	-5600,42		

Tabel 6.4

Funktiewaarden behorend bij de optimale strategie

7. Het ∞-stapsbeslissingsprobleem met als kriterium de gemiddelde opbrengst per tijdseenheid; de iteratiemethode van HOWARD voor een eindige toestandsruimte.

In deze paragraaf beginnen wij met het stochastische ∞-stapsbeslissingsprobleem. Er wordt weer een beslissingssituatie beschouwd, waarin slechts op equidistante tijdstippen een beslissing genomen kan worden. De opbrengsten worden niet verdisconteerd (dus $\alpha = 1$).

Wij veronderstellen dat het systeem slechts N verschillende toestanden kan aannemen, die wij nummeren met $i = 1,\ldots,N$. Voor iedere toestand i is een eindige verzameling $\mathfrak{X}(i)$ van toegelaten beslissingen gegeven, waaruit de beslisser een keuze moet doen. Als de beslisser in toestand i de toegelaten beslissing X neemt, wordt een eindige opbrengst $h(i,X)$ verkregen en is j met kans $p_{ij}(X)$ de toestand op het volgende beslissingstijdstip. De kansen $p_{ij}(X)$ en de opbrengsten $h(i,X)$ worden bekend verondersteld. Vanzelfsprekend moet voor elke $X \in \mathfrak{X}(i)$ gelden

$$(7.1) \qquad p_{ij}(X) \geq 0 \quad \text{en} \quad \sum_{j=1}^{N} p_{ij}(X) = 1 \qquad \text{voor } i,j = 1,\ldots,N.$$

Wij beperken ons tot de klasse $\mathfrak{Z}$ van dié strategieën, die aan iedere toestand steeds eenzelfde toegelaten beslissing toevoegen. Voor elke strategie $z \in \mathfrak{Z}$ kunnen wij overgangskansen $p_{ij}[z]$ definiëren door

$$(7.2) \qquad p_{ij}[z] = p_{ij}(z(i)),$$

waarbij $z(i)$ de beslissing is, die strategie z in toestand i voorschrijft. Als dus een strategie $z \in \mathfrak{Z}$ wordt toegepast, kan het waar te nemen beslissingsproces beschreven worden door middel van een matrix $\mathfrak{P}(z)$ van overgangskansen $p_{ij}[z]$. Een dergelijk proces is een eindige Markov-keten.

Voor wij verder zullen gaan met de behandeling van het zojuist ingevoerde ∞-stapsbeslissingsprobleem, bespreken wij eerst enkele eigenschappen van eindige Markov-ketens. *)

*) Het hierna volgende komt grotendeels overeen met het behandelde in de paragrafen 3, 4 en 5 van deel 4 (blz. 9 e.v.).

Wij beschouwen daartoe een systeem, dat door een waarnemer slechts op discrete tijdstippen wordt geobserveerd en zich dan in één van de toestanden 1,...,N bevindt. Stel dat het systeem nu in toestand i waargenomen wordt. Als de kans dat het systeem op het volgende tijdstip in toestand j verkeert uitsluitend van i en j afhangt (dus niet van het beschouwde tijdstip en de vroeger door het systeem doorlopen toestanden), dan spreken wij van een eindige Markov-keten met stationaire overgangswaarschijnlijkheden. Wij kunnen dan stellen dat het systeem geen "geheugen" heeft; d.w.z. alleen de toestand, waarin het systeem op een gegeven tijdstip verkeert, is bepalend voor de toestand op het volgende tijdstip. Als p_{ij} de overgangswaarschijnlijkheid is van toestand i naar toestand j, dan moet gelden

$$(7.3) \qquad p_{ij} \geqq 0 \quad \text{en} \quad \sum_{j=1}^{N} p_{ij} = 1 \qquad \text{voor } i,j = 1,\ldots,N.$$

De mogelijke ontwikkelingen van het systeem kunnen volledig worden beschreven als de begintoestand en de overgangswaarschijnlijkheden bekend zijn.

Wij definiëren voor alle $i,j = 1,\ldots,N$

$$(7.4) \qquad p_{ij}^{(0)} = \begin{cases} 1 & \text{als } j = i, \\ 0 & \text{anders} \end{cases}$$

en

$$(7.5) \qquad p_{ij}^{(1)} = p_{ij}.$$

Definiëren wij $p_{ij}^{(k)}$ als de kans dat het systeem uitgaande van toestand i k tijdstippen later in toestand j is, dan volgt uit de eigenschappen van een Markov-keten

$$(7.6) \qquad p_{ij}^{(k)} = \sum_{h=1}^{N} p_{ih}^{(k-1)} p_{hj} \qquad \text{voor } k = 1,2,\ldots;\ i,j = 1,\ldots,N.$$

Formule (7.6) kunnen wij generaliseren tot

$$(7.7) \qquad p_{ij}^{(m+n)} = \sum_{h=1}^{N} p_{ih}^{(m)} p_{hj}^{(n)} \qquad \text{voor } m,n = 0,1,2,\ldots;\ i,j = 1,\ldots,N.$$

Als wij stellen $\mathfrak{P} \overset{\text{def}}{=} (p_{ij})$, dan zijn de k-stapsovergangskansen $p_{ij}^{(k)}$ de elementen van het matrixprodukt $\mathfrak{P}^k$.

Wij zeggen dat toestand j <u>bereikbaar</u> is vanuit toestand i dan en slechts dan als er een positief getal k bestaat waarvoor $p_{ij}^{(k)} > 0$ is. Dus wanneer j niet bereikbaar is vanuit i, geldt

$$p_{ij}^{(k)} = 0 \qquad \text{voor } k = 1,2,\dots \; . \tag{7.8}$$

Wij noemen toestand i een <u>doorgangstoestand</u> als er een vanuit i bereikbare toestand j bestaat, zodanig dat i niet bereikbaar is vanuit toestand j. Een toestand, die geen doorgangstoestand is, heet een <u>terugkeertoestand</u>. Als k bereikbaar is vanuit i én j is bereikbaar vanuit k, dan geldt dat j bereikbaar is vanuit i. Deze eigenschap en de gegeven definities impliceren dat een terugkeertoestand vanuit zichzelf bereikbaar is en dat een doorgangstoestand niet bereikt kan worden vanuit een terugkeertoestand.

Een niet-lege verzameling van toestanden heet een <u>fuik</u> als het onmogelijk is om vanuit een toestand in die verzameling een toestand buiten die verzameling te bereiken (de verzameling van alle toestanden is dus een fuik). Een fuik die geen twee of meer disjuncte (deel-)fuiken bevat, heet een <u>priemfuik</u>. Een fuik die geen kleinere fuik bevat, heet een <u>kernfuik</u>.

Een <u>eindige</u> Markov-keten bezit de volgende eigenschappen:

a) een fuik bevat niet uitsluitend doorgangstoestanden.
b) een kernfuik is een fuik waarin alle toestanden onderling bereikbaar zijn en omgekeerd is elke fuik met deze eigenschap een kernfuik (een kernfuik bevat dus uitsluitend terugkeertoestanden die alle onderling bereikbaar zijn); elke terugkeertoestand behoort tot een kernfuik.
c) elke kernfuik is een priemfuik, daarentegen is alleen een priemfuik zonder doorgangstoestanden een kernfuik.

Een splitsing van de toestandsruimte in één of meer disjuncte priemfuiken en een (mogelijk lege) verzameling van doorgangstoestanden behoeft niet eenduidig te zijn; wel is in elke splitsing het aantal priemfuiken even groot. Een splitsing van de toestandsruimte in één of meer kernfuiken en een verzameling van doorgangstoestanden (d.i. de verzameling van alle doorgangstoestanden!) is echter ondubbelzinnig.

Wij vermelden zonder bewijs dat voor alle i geldt *)

$$\lim_{k\to\infty} p_{ij}^{(k)} = 0 \tag{7.9}$$

als j een doorgangstoestand is. Het systeem bevindt zich dus met kans 1 slechts een eindig aantal stappen in de verzameling van alle doorgangstoestanden. M.a.w. het systeem komt ongeacht de begintoestand met kans 1 na een eindig aantal stappen in een terugkeertoestand en blijft dan voor altijd in de kernfuik, waartoe die terugkeertoestand behoort. Verder vermelden wij zonder bewijs, dat vanuit elke begintoestand in een priemfuik elke terugkeertoestand uit die priemfuik met kans 1 na een eindig aantal stappen wordt aangenomen.

Tenslotte definiëren wij een <u>kringfuik</u> met periode M (waarbij M geheel en groter dan 1 is) als een fuik, die gesplitst kan worden in deelverzamelingen $A_1, A_2, \dots, A_M$, zodanig dat voor $i = 1,\dots,M$ vanuit een toestand behorend tot A_i alleen een directe overgang mogelijk is naar een toestand die tot A_{i+1} behoort, waarbij wij A_{M+1} met A_1 identificeren.

<u>Voorbeeld 7.1</u> <u>De structuur van een Markov-keten.</u>

Stel dat de matrix $\mathfrak{P}$ van overgangskansen er als volgt uitziet:

$$\mathfrak{P} = \begin{array}{c c}
\begin{array}{cccccccc} (1) & (2) & (3) & (4) & (5) & (6) & (7) & (8) \end{array} & \text{naar/van} \\
\begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\
0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2} & \frac{1}{8} & \frac{3}{8} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\
0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix} & \begin{array}{c} (1) \\ (2) \\ (3) \\ (4) \\ (5) \\ (6) \\ (7) \\ (8) \end{array}
\end{array} \tag{7.10}$$

---

*) Voor bewijzen wordt de lezer verwezen naar J.L. DOOB, Stochastic processes, ch. 5, §2, (1953), Wiley, New York, of W. FELLER, An introduction to probability theory and its applications, vol. I, ch. 15, (3e druk, 1968), Wiley, New York.

De toestanden 1, 2, 3, 4, 5 en 8 zijn terugkeertoestanden en 6 en 7 zijn doorgangstoestanden. De verzameling der terugkeertoestanden valt in drie kernfuiken uiteen, n.l. $\{1,2\}$, $\{3,4,5\}$ en $\{8\}$. Behalve deze verzamelingen is ook de verzameling $\{3,4,5,6\}$ een priemfuik. Deze vier verzamelingen zijn de enige priemfuiken. De toestandsruimte is op twee manieren te splitsen in een verzameling van doorgangstoestanden en een aantal priemfuiken. De ene splitsing wordt gegeven door $\{6,7\}$, $\{1,2\}$, $\{3,4,5\}$ en $\{8\}$, de andere door $\{7\}$, $\{1,2\}$, $\{3,4,5,6\}$ en $\{8\}$. De kernfuik $\{1,2\}$ is een kringfuik met periode 2, waarbij $A_1 = \{1\}$ en $A_2 = \{2\}$. Wij merken nog op dat de kernfuik $\{8\}$ uit één toestand bestaat. Als een fuik uit slechts één toestand bestaat, heet deze toestand <u>absorberend</u>.

Om meerdere redenen is het van belang om na te gaan of de rij van getallen $p_{ij}^{(m)}$ een limiet heeft voor iedere i en j als $m \to \infty$ gaat. Wij zullen zien dat deze limiet in vele gevallen bestaat en "min of meer onafhankelijk" is van begintoestand i. Zonder bewijs vermelden wij dat, wanneer de eindige Markov-keten <u>geen</u> kringfuiken bevat, voor alle i en j

$$\lim_{m\to\infty} p_{ij}^{(m)} = q_{ij} \tag{7.11}$$

bestaat, waarbij voor alle i geldt

$$\sum_{j=1}^{N} q_{ij} = 1. \tag{7.12}$$

Indien de Markov-keten wel kringfuiken bevat, bestaat $\lim\limits_{m\to\infty} p_{ij}^{(m)}$ niet voor alle i en j. Wel kan aangetoond worden dat de zgn. Cesarolimiet *)

$$\lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} p_{ij}^{(k)} \tag{7.13}$$

altijd bestaat voor alle i en j en bovendien gelijk is aan $\lim\limits_{m\to\infty} p_{ij}^{(m)}$, als deze laatste limiet bestaat (vgl. J.L. DOOB; voetnoot op blz. 72).

*) De Cesarolimiet van een rij kan bestaan zonder dat de gewone limiet bestaat. Maar als de gewone limiet bestaat, dan bestaat de Cesarolimiet eveneens en zijn de beide limieten aan elkaar gelijk.

De getallen q_{ij} worden vanaf nu gedefinieerd door

$$(7.14) \qquad q_{ij} = \lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} p_{ij}^{(k)} \qquad \text{voor } i,j = 1,\ldots,N.$$

Als van een convergente rij de termen tussen 0 en 1 liggen, dan ligt ook de limiet tussen 0 en 1 en dus geldt

$$(7.15) \qquad 0 \leqq q_{ij} \leqq 1 \qquad \text{voor } i,j = 1,\ldots,N.$$

Voor een eindig aantal convergente rijen geldt dat de som van de limieten gelijk is aan de limiet van de som der rijen. Bijgevolg geldt

$$(7.16) \qquad \sum_{j=1}^{n} q_{ij} = 1 \qquad \text{voor } i = 1,\ldots,N.$$

Uit (7.6) volgt dat voor alle i en j geldt

$$(7.17) \qquad \begin{aligned} \frac{1}{m}\sum_{k=1}^{m} p_{ij}^{(k)} &= \frac{p_{ij}}{m} + \frac{1}{m}\sum_{k=2}^{m}\sum_{h=1}^{N} p_{ih}^{(k-1)} p_{hj} = \\ &= \frac{p_{ij}}{m} + \frac{m-1}{m}\sum_{h=1}^{N}\left\{\frac{1}{m-1}\sum_{k=1}^{m-1} p_{ih}^{(k)}\right\} p_{hj} \qquad \text{voor } m = 1,2,\ldots\ . \end{aligned}$$

Laten wij m naar oneindig gaan, dan volgt uit (7.14) en (7.17)

$$(7.18) \qquad q_{ij} = \sum_{h=1}^{N} q_{ih}\, p_{hj} \qquad \text{voor } i,j = 1,\ldots,N.$$

Op dezelfde manier bewijzen wij met behulp van (7.7), waarin we $m = 1$ en $n = k-1$ nemen, dat geldt

$$(7.19) \qquad q_{ij} = \sum_{h=1}^{N} p_{ih}\, q_{hj} \qquad \text{voor } i,j = 1,\ldots,N.$$

Als wij relatie (7.19) herhaald toepassen, volgt dat voor alle i en j geldt

$$(7.20) \qquad q_{ij} = \sum_{h=1}^{N} p_{ih}^{(k)} q_{hj} \qquad \text{voor } k = 1,2,\ldots\ .$$

Bijgevolg geldt

$$(7.21) \qquad q_{ij} = \frac{1}{m} \sum_{k=1}^{m} \sum_{h=1}^{N} p_{ih}^{(k)} q_{hj} \qquad \text{voor } m = 1,2,\ldots .$$

Laten wij m naar oneindig gaan, dan volgt uit (7.14) en (7.21)

$$(7.22) \qquad q_{ij} = \sum_{h=1}^{N} q_{ih}\, q_{hj} \qquad \text{voor } i,j = 1,\ldots,N.$$

Als wij de $(N \times N)$-matrix $\mathbf{Q}$ als volgt definiëren

$$(7.23) \qquad \mathbf{Q} = (q_{ij})_{i,j = 1,\ldots,N} ,$$

dan kunnen wij (7.14), (7.18), (7.19) en (7.22) in matrix-notatie samenvatten tot

$$(7.24) \qquad \mathbf{Q} = \lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} \mathcal{P}^{k} \;;\; \mathbf{Q} = \mathbf{Q}\mathcal{P} = \mathcal{P}\mathbf{Q} = \mathbf{Q}\mathbf{Q}.$$

De getallen q_{ij} worden vaak de invariante kansen genoemd. Bewezen kan worden dat q_{ij} gelijk is aan de verwachte fractie van het aantal tijdstippen waarop het systeem in toestand j verkeert, als de begintoestand i is en het aantal te beschouwen tijdstippen oneindig groot is. Als i tot een priemfuik behoort, geldt zelfs dat de fractie van het aantal tijdstippen waarop het systeeem in toestand j verkeert met kans 1 gelijk is aan q_{ij}.

Wij kunnen de kansen q_{ij} ook als volgt interpreteren. Veronderstel dat wij het systeem gedurende een <u>zeer lange tijd</u> niet hebben gadegeslagen. Wij weten niet hoeveel toestándswijzigingen zich in deze tijd hebben voltrokken, maar wel dat dit aantal zeer groot is. Als ons tevens bekend is, dat i de begintoestand van het systeem was, zal de kans dat het systeem zich nu in toestand j bevindt bij benadering gelijk zijn aan q_{ij}.

Uit (7.8), (7.9) en (7.14) volgt

$$(7.25) \qquad q_{ij} = 0 ,$$

als j doorgangstoestand is of als j niet bereikbaar is vanuit i. Zonder bewijs vermelden wij dat voor elke terugkeertoestand j geldt

$$q_{jj} = \frac{1}{\mu_{jj}} > 0 \ , \tag{7.26}$$

waarbij μ_{jj} gedefinieerd is als de verwachting van het aantal stappen dat het systeem nodig heeft om uitgaande van toestand j in deze toestand terug te keren. Tevens vermelden wij zonder bewijs dat voor toestanden s en t uit eenzelfde priemfuik geldt

$$q_{sj} = q_{tj} \qquad \text{voor } j = 1,\dots,N. \tag{7.27}$$

Veronderstel dat de Markov-keten geen disjuncte fuiken bevat; m.a.w. er is precies één kernfuik. De verzameling van alle toestanden vormt dan een priemfuik. Bijgevolg bestaan er getallen $q_1,\dots,q_N$, waarvoor geldt

$$q_j = q_{ij} \qquad \text{voor } i,j = 1,\dots,N. \tag{7.28}$$

De relaties (7.16) en (7.18) gaan dan over in

$$\sum_{j=1}^{N} q_j = 1 \tag{7.29}$$

en

$$q_j = \sum_{h=1}^{N} q_h \, p_{hj} \qquad \text{voor } j = 1,\dots,N. \tag{7.30}$$

De relaties (7.19) en (7.22) worden hierdoor gereduceerd tot $q_j = q_j$. De invariante kansen $q_1,\dots,q_N$ voldoen dus aan het stelsel vergelijkingen

$$\begin{cases} x_j = \displaystyle\sum_{h=1}^{N} x_h \, p_{hj} \qquad \text{voor } j = 1,\dots,N, & (7.31a) \\ \displaystyle\sum_{j=1}^{N} x_j = 1 \ . & (7.31b) \end{cases}$$

Het stelsel (7.31) heeft slechts één oplossing. Het bewijs hiervoor is eenvoudig. Door een herhaalde substitutie op (7.31a) toe te passen verkrijgen wij voor k = 1,2,...

$$x_j = \sum_{h=1}^{N} x_h \, p_{hj}^{(k)} \qquad \text{voor } j = 1,\dots,N. \tag{7.32}$$

Nemen wij de Cesarolimiet, dan vinden wij

$$(7.33) \qquad x_j = \sum_{h=1}^{N} x_h\, q_j = q_j \sum_{h=1}^{N} x_h = q_j \qquad \text{voor } j = 1,\dots,N.$$

Als aan N-1 vergelijkingen uit (7.31a) voldaan is en aan (7.31b), dan is automatisch de resterende vergelijking uit (7.31a) vervuld, omdat

$$\sum_{j=1}^{N} x_j = \sum_{j=1}^{N} \sum_{h=1}^{N} x_h\, p_{hj} \ .$$

Als dus in (7.31) één vergelijking uit (7.31a) weggelaten wordt, heeft het resulterende stelsel van N vergelijkingen in N onbekenden precies één oplossing en dit is tevens de oplossing $(q_1,\dots,q_N)$ van (7.31).

Hiermee wordt de bespreking van eindige Markov-ketens besloten en keren wij weer terug naar het ∞-stapsbeslissingsprobleem.

Zoals wij reeds hebben vastgesteld correspondeert met elke strategie $z \in \mathcal{Z}$ een eindige Markov-keten, waarvan $\mathcal{P}(z)$ de matrix der overgangskansen is. De bijbehorende invariante kansen geven wij aan met $q_{ij}[z]$.

Veronderstel dat i de begintoestand is en dat strategie z wordt toegepast. Definiëren wij de stochastische variabele $\underline{h}_k$ als de opbrengst op het k^{de} beslissingstijdstip, dan geldt

$$(7.34) \qquad \mathcal{E}\,\underline{h}_k = \sum_{j=1}^{N} p_{ij}^{(k-1)}[z]\ h(j,z(j)) \qquad \text{voor } k = 1,2,\dots \ ,$$

waarbij

$$(7.35) \qquad p_{ij}^{(0)}[z] = \begin{cases} 1 & \text{als } i = j \ , \\ 0 & \text{als } i \neq j \ . \end{cases}$$

Bijgevolg geldt (vgl. (7.14))

$$(7.36) \qquad \lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} \mathcal{E}\,\underline{h}_k = \sum_{j=1}^{N} q_{ij}[z]\ h(j,z(j)) \ .$$

Wij definiëren nu

(7.37) $$y_i(z) = \sum_{j=1}^{N} q_{ij}[z]\, h(j,z(j)) \qquad \text{voor } i = 1,\dots,N \ .$$

Kiezen wij het interval tussen twee opeenvolgende beslissingstijdstippen als tijdseenheid, dan is $y_i(z)$ gelijk aan de gemiddelde verwachte opbrengst per tijdseenheid als i de begintoestand is en strategie z gedurende een oneindig lange tijd wordt toegepast. De funktie $h(j,z(j))$ is begrensd. Dit heeft tot gevolg dat voor elke begintoestand i geldt (vgl. J.L. DOOB (p.220); voetnoot op blz.72):

(7.38) $$\lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} \underline{h}_k$$

bestaat met kans 1 *) en

(7.39) $$\mathcal{E}\left(\lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} \underline{h}_k\right) = \lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} \mathcal{E}\,\underline{h}_k \ .$$

Wij zien dus dat voor begintoestand i ook de verwachting van de gemiddelde opbrengst per tijdseenheid gelijk is aan $y_i(z)$. Wij kunnen zelfs een nog sterkere uitspraak doen omtrent $y_i(z)$. Als de begintoestand i tot een priemfuik behoort, geldt (vgl. J.L. DOOB (p.220); voetnoot op blz. 72)

(7.40) $$P\{\lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} \underline{h}_k = y_i(z)\} = 1 \ .$$

Dus als strategie z gedurende een oneindig lange tijd wordt toegepast en de begintoestand i tot een priemfuik behoort, dan is de <u>werkelijke</u> gemiddelde opbrengst per tijdseenheid met kans 1 gelijk aan $y_i(z)$. Tegenover de onzekerheid met betrekking tot het toekomstige verloop van het beslissingsproces staat de "bijna" absolute zekerheid van de werkelijke gemiddelde opbrengst per tijdseenheid!

---

*) D.w.z. voor elke mogelijke realisering van het beslissingsproces heeft de gemiddelde opbrengst per tijdseenheid genomen over de eerste m tijdseenheden een eindige limiet voor $m \to \infty$.

Als de begintoestand i niet tot een priemfuik behoort (m.a.w. als i een doorgangstoestand is, waaruit twee of meer disjuncte fuiken bereikt kunnen worden), is de werkelijke gemiddelde opbrengst per tijdseenheid een stochastische variabele met verwachting $y_i(z)$. Het behoeft geen betoog dat de funktie $y_i(z)$, gedefinieerd door (7.37), een uitstekend <u>optimaliteits-kriterium</u> is. Strategie z heet <u>optimaal</u> als $y_i(z)$ maximaal is voor alle i.

Uit (7.27) en (7.37) volgt dat voor toestanden s en t uit eenzelfde priemfuik geldt

$$y_s(z) = y_t(z). \tag{7.41}$$

Dus voor toestanden die tot eenzelfde priemfuik behoren, neemt de gemiddelde opbrengst per tijdseenheid met kans 1 dezelfde waarde aan. De verzameling van alle toestanden is een priemfuik, als de matrix $\mathcal{P}(z)$ geen disjuncte fuiken bevat (in de praktijk is dit vrijwel altijd het geval). De gemiddelde opbrengst per tijdseenheid is dan, ongeacht de begintoestand, met kans 1 gelijk aan y(z), waarbij

$$y(z) \overset{\text{def}}{=} y_1(z) = \ldots = y_N(z). \tag{7.42}$$

Indien wij van een klein aantal strategieën de beste willen weten, dan kunnen wij voor elke strategie afzonderlijk de kriteriumfunktie berekenen. Deze werkwijze lichten wij toe aan de hand van een voorbeeld.

<u>Voorbeeld 7.2</u> <u>Een brouwerijprobleem.</u>

De stichter van de bierbrouwerij "De Hop" is in zijn woelige studententijd één van de oprichters geweest van de bekende studentenvereniging "Euclides". Deze vereniging heeft het mede hieraan te danken dat de brouwerij haar elke vrijdag gratis een hoeveelheid bier levert, waaronder minstens 1 en hoogstens 3 vaten met het befaamde Scheldebier. Afhankelijk van het aantal vaten Scheldebier dat de wekelijkse leverantie omvat, krijgen de dorstige studenten bovendien een aantal vaten extra met een andere biersoort. De kosten voor deze extra bierleverantie bedragen voor de brouwerij $a_1 = 9$, $a_2 = 7$ of $a_3 = 3$, al naar gelang het aantal vaten Scheldebier in de leverantie 1, 2 of 3 is. Het Scheldebier is niet in de handel verkrijgbaar. Het wordt nog in kleine hoeveelheden volgens het oorspronkelijke procédé van de stichter gebrouwen. Elke woensdag zijn in de lagerkelders 1 of 2 vaten Scheldebier voldoende gerijpt voor consumptie; voor elke woensdag is de kans dat 1 vat met dit bier gerijpt is, gelijk aan $p_1 = 0{,}5$ en de kans op 2 vaten gerijpt bier is $p_2 = 1 - p_1 = 0{,}5$. Om redenen van technische aard moeten de vaten met gerijpt Scheldebier dan uit de lagerkelders worden overgebracht naar het "taphuis", waar ten hoogste 3 vaten Scheldebier opgeslagen kunnen worden (geen voorraadkosten). Een vat, waarvoor geen plaats is, wordt in de personeelskantine gezet, waar het personeel tegen een geringe vergoeding een glas bier kan drinken. De opbrengst van een vat bier in de personeelskantine is gelijk aan $c_1 = 0{,}75$.

's Maandags vindt in de brouwerij de wekelijkse rondleiding van bezoekers plaats. Na afloop van de rondleiding worden de bezoekers door de brouwerij diverse consumpties aangeboden. Als de samenstelling van de bezoekende groep zodanig is dat één vat Scheldebier vereist is voor de groep (de kans hierop is elke maandag $r_1 = 0{,}6$) en als dit bier dan niet voorradig is in het taphuis, brengt dit voor de brouwerij onkosten $c_2 = 5$ met zich mee. Elke maandag is de kans dat geen Scheldebier nodig is voor de bezoekers $r_0 = 1 - r_1 = 0{,}4$.

Gevraagd wordt met behulp van deze gegevens voor de brouwerij een optimaal leverantieschema aan de studentenvereniging "Euclides" op te stellen.

Oplossing

De toestand van het systeem wordt bepaald door het aantal vaten Scheldebier dat op vrijdag vlak <u>voor</u> de leverantie aanwezig is in het taphuis. De toestandsgrootheid i kan de waarden 1, 2 en 3 aannemen. De beslissing X zullen wij identificeren met het aantal vaten Scheldebier, dat vlak <u>na</u> de leverantie in het taphuis aanwezig is; dus als in toestand i beslissing X genomen wordt, omvat de leverantie aan de studenten i-X vaten Scheldebier. De leverantie moet tenminste 1 vat Scheldebier bevatten. De verzameling van toegelaten beslissingen in toestand i wordt dus gegeven door

(7.43) $\mathcal{X}(i) = \{0,\dots,i-1\}$ voor $i = 1,2,3$.

De overgangskansen $p_{ij}(X)$ en de opbrengstfunktie $h(i,X)$ voldoen aan

$$(7.44)\qquad p_{ij}(X) = \begin{cases} p_1 & \text{voor } X = 0;\ j = 1;\ i = 1,2,3, \\ p_2 & \text{voor } X = 0;\ j = 2;\ i = 1,2,3, \\ r_1p_1 & \text{voor } X = 1;\ j = 1;\ i = 2,3, \\ r_0p_1 + r_1p_2 & \text{voor } X = 1;\ j = 2;\ i = 2,3, \\ r_0p_2 & \text{voor } X = 1;\ j = 3;\ i = 2,3, \\ r_1p_1 & \text{voor } X = 2;\ j = 2;\ i = 3, \\ r_0 + r_1p_2 & \text{voor } X = 2;\ j = 3;\ i = 3, \\ 0 & \text{anders} \end{cases}$$

resp.

$$(7.45)\qquad h(i,X) = \begin{cases} -a_i - r_1c_2 & \text{voor } X = 0;\ i = 1,2,3, \\ -a_{i-1} & \text{voor } X = 1;\ i = 2,3, \\ -a_1 + r_0p_2c_1 & \text{voor } X = 2;\ i = 3. \end{cases}$$

De numerieke waarden van $p_{ij}(X)$ en $h(i,X)$ worden in tabel 7.1 gegeven.

toestand	beslissing	overgangskansen			opbrengst-funktie
i	X	$p_{i1}(X)$	$p_{i2}(X)$	$p_{i3}(X)$	$h(i,X)$
1	0	0,5	0,5	0	-12
2	0	0,5	0,5	0	-10
	1	0,3	0,5	0,2	- 9
3	0	0,5	0,5	0	- 6
	1	0,3	0,5	0,2	- 7
	2	0	0,3	0,7	- 8,85

Tabel 7.1

Numerieke waarden van $p_{ij}(X)$ en $h(i,X)$ uit het brouwerijprobleem

Het aantal mogelijke strategieën is $1 \times 2 \times 3 = 6$. Wij zullen voor elke strategie afzonderlijk de waarden van de kriteriumfunktie $y_i(z)$ berekenen.

Strategie $z^{(1)} = (0,1,2)$. De matrix $\mathcal{P}(z^{(1)})$ van overgangskansen wordt gegeven door

$$
(7.46) \qquad \mathcal{P}(z^{(1)}) = \begin{array}{c} \\ \\ \\ \\ \end{array}
\begin{array}{c}
\begin{array}{ccc} (1) & (2) & (3) \end{array} \\
\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{3}{10} & \frac{1}{2} & \frac{2}{10} \\ 0 & \frac{3}{10} & \frac{7}{10} \end{pmatrix}
\end{array}
\begin{array}{c} \\ (1) \\ (2) \\ (3). \end{array}
$$

Uit deze matrix volgt dat de toestanden 1, 2 en 3 onder strategie $z^{(1)}$ terugkeertoestanden zijn, die alle onderling bereikbaar zijn. De matrix $\mathcal{P}(z^{(1)})$ bevat dus geen disjuncte fuiken. De invariante kansen q_1, q_2 en q_3 vormen de oplossing van het volgende stelsel vergelijkingen (vgl.relaties (7.29) en (7.30) op blz. 76):

$$
(7.47) \qquad \begin{cases} q_1 = \frac{1}{2} q_1 + \frac{3}{10} q_2 , \\ q_2 = \frac{1}{2} q_1 + \frac{1}{2} q_2 + \frac{3}{10} q_3 , \\ q_3 = \frac{2}{10} q_2 + \frac{7}{10} q_3 , \\ q_1 + q_2 + q_3 = 1. \end{cases}
$$

De oplossing wordt gegeven door

$$
(7.48) \qquad q_1 = \frac{9}{34} , \quad q_2 = \frac{15}{34} , \quad q_3 = \frac{10}{34} .
$$

Zowel voor begintoestand 1, 2 als 3 wordt de gemiddelde opbrengst per tijdseenheid met kans 1 gegeven door

$$
(7.49) \qquad \begin{cases} y(z^{(1)}) = \sum_{j=1}^{3} q_j \, h(j, z^{(1)}(j)) = -\frac{9}{34} \cdot 12 - \frac{15}{34} \cdot 9 - \frac{10}{34} \cdot 8{,}85 = \\ \qquad = -9{,}75. \end{cases}
$$

Strategie $z^{(2)} = (0,1,1)$. Bij deze strategie behoort de volgende matrix van overgangskansen:

$$\begin{array}{cc} & \begin{array}{ccc} (1) & (2) & (3) \end{array} & \\ (7.50) \quad \mathcal{P}(z^{(2)}) = & \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{3}{10} & \frac{1}{2} & \frac{2}{10} \\ \frac{3}{10} & \frac{1}{2} & \frac{2}{10} \end{pmatrix} & \begin{array}{c} (1) \\ (2) \\ (3). \end{array} \end{array}$$

Onder strategie $z^{(2)}$ zijn de toestanden 1, 2 en 3 terugkeertoestanden, die alle onderling bereikbaar zijn. De matrix $\mathcal{P}(z^{(2)})$ bevat dus geen disjuncte fuiken. De invariante kansen q_1, q_2 en q_3 vormen de oplossing van

$$(7.51) \qquad \begin{cases} q_1 = \frac{1}{2} q_1 + \frac{3}{10} q_2 + \frac{3}{10} q_3 \ , \\ q_2 = \frac{1}{2} q_1 + \frac{1}{2} q_2 + \frac{1}{2} q_3 \ , \\ q_3 = \qquad\quad \frac{2}{10} q_2 + \frac{2}{10} q_3 \ , \\ q_1 + q_2 + q_3 = 1. \end{cases}$$

De oplossing luidt

$$(7.52) \qquad q_1 = \frac{3}{8} \ , \quad q_2 = \frac{1}{2} \ , \quad q_3 = \frac{1}{8} \ .$$

De gemiddelde opbrengst per tijdseenheid wordt ongeacht de begintoestand met kans 1 gegeven door

$$(7.53) \qquad y(z^{(2)}) = \sum_{j=1}^{3} q_j \ h(j,z^{(2)}(j)) = -\frac{3}{8} \cdot 12 - \frac{1}{2} \cdot 9 - \frac{1}{8} \cdot 7 = -9{,}875.$$

Strategie $z^{(3)} = (0,1,0)$. De matrix $\mathcal{P}(z^{(3)})$ van overgangskansen ziet er als volgt uit:

$$(7.54)\qquad \mathfrak{P}(z^{(3)}) = \begin{array}{c c} \begin{array}{ccc} (1) & (2) & (3) \end{array} & \\ \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{3}{10} & \frac{1}{2} & \frac{2}{10} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} & \begin{array}{c} (1) \\ (2) \\ (3). \end{array} \end{array}$$

De toestanden 1, 2 en 3 zijn onder $z^{(3)}$ terugkeertoestanden, die alle onderling bereikbaar zijn. De matrix $\mathfrak{P}(z^{(3)})$ bevat dus geen disjuncte fuiken. Voor de invariante kansen vinden wij

$$(7.55)\qquad q_1 = \frac{2}{5}\,,\quad q_2 = \frac{1}{2}\,,\quad q_3 = \frac{1}{10}\,.$$

De gemiddelde opbrengst per tijdseenheid wordt ongeacht de begintoestand met kans 1 gegeven door

$$(7.56)\qquad y(z^{(3)}) = -\frac{2}{5}\cdot 12 - \frac{1}{2}\cdot 9 - \frac{1}{10}\cdot 6 = -9{,}9.$$

<u>Strategie $z^{(4)} = (0,0,2)$.</u> De matrix $\mathfrak{P}(z^{(4)})$ wordt gegeven door

$$(7.57)\qquad \mathfrak{P}(z^{(4)}) = \begin{array}{c c} \begin{array}{ccc} (1) & (2) & (3) \end{array} & \\ \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0, & \frac{3}{10} & \frac{7}{10} \end{pmatrix} & \begin{array}{c} (1) \\ (2) \\ (3). \end{array} \end{array}$$

De toestanden 1 en 2 zijn onder strategie $z^{(4)}$ terugkeertoestanden, die onderling bereikbaar zijn. Toestand 3 is onder $z^{(4)}$ een doorgangstoestand. De matrix $\mathfrak{P}(z^{(4)})$ bevat dus geen disjuncte fuiken. Voor de invariante kansen vinden wij

$$(7.58)\qquad q_1 = q_2 = \frac{1}{2}\,,\quad q_3 = 0.$$

De gemiddelde opbrengst per tijdseenheid wordt ongeacht de begintoestand met kans 1 gegeven door

(7.59) $$y(z^{(4)}) = -\frac{1}{2}.12 - \frac{1}{2}.10 = -11.$$

Op dezelfde wijze vinden wij zowel voor <u>strategie $z^{(5)} = (0,0,1)$</u> als voor <u>strategie $z^{(6)} = (0,0,0)$</u> de invariante kansen

(7.60) $$q_1 = q_2 = \frac{1}{2}, \quad q_3 = 0.$$

De gemiddelde opbrengst per tijdseenheid wordt voor deze strategieën ongeacht de begintoestand met kans 1 gegeven door

(7.61) $$y(z^{(5)}) = y(z^{(6)}) = -11.$$

Vergelijken wij de zes strategieën, dan zien wij dat strategie $z^{(1)} = (0,1,2)$ optimaal is. De leverantie aan de studenten omvat dus altijd precies één vat Scheldebier.

In dit voorbeeld hebben wij voor elke strategie afzonderlijk de waarde van de kriteriumfunktie berekend. Indien het aantal strategieën groot is, kunnen wij veelal niet op deze wijze te werk gaan. R.A. HOWARD heeft echter een iteratieprocedure ontwikkeld, die altijd tot een optimale strategie leidt.

Alvorens tot de formulering van deze iteratieprocedure over te gaan zullen wij een stelsel lineaire vergelijkingen analyseren, dat hierin een belangrijke rol speelt. Na de iteratieprocedure te hebben gegeven, zullen wij bewijzen dat deze convergeert naar een optimale strategie. De lezer die voornamelijk geïnteresseërd is in praktische resultaten kan zich zonder bezwaar beperken tot de iteratieprocedure op blz. 94 en 95, daarna tot de toepassingen op blz. 106 e.v.

Wij beschouwen een eindige Markov-keten met $\mathcal{P} = (p_{ij})_{i,j = 1,\dots,N}$ als matrix van overgangskansen.

<u>Stelling 7.1</u>

Als de verzameling W van toestanden geen fuik bevat en de matrix $\mathcal{Q}$ gedefinieerd is door

(7.62) $$\mathcal{W} = (p_{ij})_{i,j \in W} ,$$

dan bestaat $(\mathcal{I} - \mathcal{W})^{-1}$ en wordt deze inverse gegeven door $\mathcal{I} + \sum_{k=1}^{\infty} \mathcal{W}^k$.

Bewijs:

Het element $w_{ij}^{(k)}$ van de matrix $\mathcal{W}^k$ is niets anders dan de kans dat het systeem uitgaande van toestand $i \in W$ over k stappen in toestand $j \in W$ is zonder dat het systeem tussentijds een toestand buiten W aangenomen heeft. Omdat de verzameling W geen fuik bevat, geldt dat het systeem, ongeacht de uitgangstoestand, na eindig veel stappen met kans 1 een toestand buiten W aanneemt; dat wil zeggen

(7.63) $$\lim_{k \to \infty} w_{ij}^{(k)} = 0 \qquad \text{voor alle } i,j \in W.$$

De stelling volgt nu met behulp van relatie (6.71) op blz. 21.

Stelling 7.2

Elke oplossing $X = (x_1, \ldots, x_N)$ van het stelsel lineaire vergelijkingen

(7.64) $$x_i = \sum_{j=1}^{N} p_{ij}\, x_j \qquad \text{voor } i = 1, \ldots, N$$

heeft de eigenschap dat voor toestanden s en t uit eenzelfde priemfuik geldt

(7.65) $$x_s = x_t \; .$$

Bewijs:

Het behoeft geen betoog, dat (7.64) een oplossing bezit (bijv. X = 0). Door (7.64) herhaald toe te passen vinden wij met behulp van relatie (7.6), dat

(7.66) $$x_i = \sum_{j=1}^{N} p_{ij}^{(k)} x_j \qquad \text{voor } i = 1, \ldots, N;\ k \geqq 1.$$

Nemen wij de Cesarolimiet, dan krijgen wij (vgl. (7.14))

$$x_i = \sum_{j=1}^{N} q_{ij}\, x_j \qquad \text{voor } i = 1,\dots,N. \tag{7.67}$$

De stelling volgt nu uit (7.67) en (7.27).

Stelling 7.3

Als de Markov-keten geen disjuncte fuiken bevat en $h_1,\dots,h_N$ constanten zijn, dan geldt:

a) het stelsel lineaire vergelijkingen

$$y + v_i - \sum_{j=1}^{N} p_{ij}\, v_j = h_i \qquad \text{voor } i = 1,\dots,N \tag{7.68}$$

is oplosbaar en elke oplossing (y,V) heeft de eigenschap dat

$$y = \sum_{i=1}^{N} q_i\, h_i \;; \tag{7.69}$$

b) als (y,U) en (y,W) oplossingen van (7.68) zijn, dan hebben alle componenten van $U-W$ dezelfde waarde;

c) als i_1 een willekeurige toestand is en wij aan het stelsel (7.68) de voorwaarde

$$v_{i_1} = 0 \tag{7.70}$$

toevoegen, dan heeft het resulterende stelsel een unieke oplossing.

Bewijs:

a) De verzameling van alle toestanden bevat tenminste één terugkeertoestand. Veronderstel dat toestand N een terugkeertoestand is. Definieer de $(N-1) \times (N-1)$ matrix $\mathcal{R}$ door

$$\mathcal{R} = (p_{ij})_{i,j=1,\dots,N-1}. \tag{7.71}$$

De verzameling $\{1,\ldots,N-1\}$ bevat geen disjuncte fuiken omdat toestand N vanuit elke toestand bereikbaar is. Uit stelling 7.1 volgt nu dat de matrix $\mathfrak{I} - \mathfrak{R}$ een inverse bezit. Derhalve heeft het stelsel lineaire vergelijkingen

$$(7.72) \qquad k_i = h_i + \sum_{j=1}^{N-1} p_{ij} k_j \qquad \text{voor } i=1,\ldots,N-1$$

een unieke oplossing $(k_1,\ldots,k_{N-1})$, en het stelsel lineaire vergelijkingen

$$(7.73) \qquad t_i = 1 + \sum_{j=1}^{N-1} p_{ij} t_j \qquad \text{voor } i=1,\ldots,N-1$$

een unieke oplossing $(t_1,\ldots,t_{N-1})$. De grootheid t_i kan geïnterpreteerd worden als de verwachte tijd om toestand N te bereiken uitgaande van toestand i, terwijl k_i geïnterpreteerd kan worden als de verwachte opbrengst gedurende dit tijdsinterval. Definieer nu

$$(7.74) \qquad y = \frac{h_N + \sum_{j=1}^{N-1} p_{Nj} k_j}{1 + \sum_{j=1}^{N-1} p_{Nj} t_j}$$

en

$$(7.75) \qquad v_i = \begin{cases} k_i - y t_i & \text{voor } i=1,\ldots,N-1, \\ 0 & \text{voor } i=N. \end{cases}$$

Het rechterlid van (7.74) kan geïnterpreteerd worden als het quotient van de verwachte opbrengst in het tijdsinterval tussen twee opeenvolgende tijdstippen waarop het systeem in toestand N is en de verwachte duur van dit tijdsinterval. Wij zullen nu aantonen dat $(y,v_1,\ldots,v_N)$ aan (7.68) voldoet. Uit (7.72),(7.73) en (7.75) volgt dat voor $i=1,\ldots,N-1$ geldt

$$(7.76) \qquad y + v_i = y + h_i + \sum_{j=1}^{N-1} p_{ij} k_j - y\,\{1 + \sum_{j=1}^{N-1} p_{ij} t_j\} = \\ = h_i + \sum_{j=1}^{N-1} p_{ij}\{k_j - y t_j\} = h_i + \sum_{j=1}^{N} p_{ij} v_j ,$$

terwijl uit (7.74) en (7.75) volgt dat

$$(7.77) \qquad y + v_n = h_n + \sum_{j=1}^{N-1} p_{Nj}k_j - y \sum_{j=1}^{N-1} p_{Nj}t_j =$$

$$= h_N + \sum_{j=1}^{N} p_{Nj} \quad k_j - yt_j \quad = h_N + \sum_{j=1}^{N} p_{Nj}v_j .$$

Hiermee is aangetoond dat (7.68) een oplossing bezit.

Vermenigvuldigen wij beide leden van (7.68) met de invariante kans q_i en sommeren wij over i, dan vinden wij

$$(7.78) \qquad y \sum_{i=1}^{N} q_i + \sum_{i=1}^{N} q_iv_i - \sum_{i=1}^{N} q_i \sum_{j=1}^{N} p_{ij}v_j = \sum_{i=1}^{N} q_ih_i .$$

Verwisselen wij in de laatste term uit het linkerlid van (7.78) de sommatievolgorde en passen wij (7.29) en (7.30) toe, dan vinden wij

$$(7.79) \qquad y + \sum_{i=1}^{N} q_iv_i - \sum_{j=1}^{N} q_jv_j = \sum_{i=1}^{N} q_ih_i ,$$

waarmee (7.69) bewezen is.

b) Uit (7.68 en (7.69) volgt door aftrekken dat

$$(7.80) \qquad u_i - w_i = \sum_{j=1}^{N} p_{ij} \; (u_j - w_j) \qquad \text{voor } i=1,\dots,N.$$

De verzameling van alle toestanden is een priemfuik, omdat de matrix P geen disjuncte fuiken heeft. Uit stelling 7.2 volgt nu dat alle componenten van U - W dezelfde waarde hebben

c) Wij zullen eerst aantonen dat (7.68) onder (7.70) een oplossing bezit. Daartoe merken wij eerst op dat voor een willekeurige constante c geldt dat $(y,v_1+c,\dots,v_N+c)$ voldoet aan (7.68) als $(y,v_1,\dots,v_N)$ aan (7.68) voldoet. Dit volgt direct uit het feit dat $p_{i1} + \dots + p_{iN} = 1$ voor alle $i = 1,\dots,N$. Uit bewering a) van de stelling volgt nu dat het stelsel vergelijkingen (7.68) onder de voorwaarde (7.70) een oplossing bezit. Vervolgens tonen wij aan dat deze oplossing uniek is. Stel dat zowel (y,U) als (y,V) voldoet aan het stelsel gevormd door (7.68) en (7.70). Uit bewering b) van de stelling volgt dan dat $u_i - v_i = c$ voor alle $i = 1,\dots,N$. Aangezien $u_{i_1} = v_{i_1} = 0$, volgt dan dat $c = 0$. Hiermee is bewering c) bewezen.

Stelling 7.4

a) Het stelsel lineaire vergelijkingen

$$\left\{\begin{array}{llr} (7.81a) & y_i - \sum_{j=1}^{N} p_{ij}\, y_j = 0 & \text{voor } i = 1,\dots,N, \\ (7.81b) & y_i + v_i - \sum_{j=1}^{N} p_{ij}\, v_j = h_i & \text{voor } i = 1,\dots,N \end{array}\right.$$

is oplosbaar en voor elke oplossing (Y,V) geldt

$$y_i = \sum_{j=1}^{N} q_{ij}\, h_j \qquad \text{voor } i = 1,\dots,N. \tag{7.82}$$

b) Veronderstel dat een splitsing van de toestandsruimte in disjuncte verzamelingen E_0, $F_1,\dots,F_m$ gegeven is, waarbij E_0 een (mogelijk lege) verzameling van doorgangstoestanden is en $F_1,\dots,F_m$ priemfuiken zijn Kiezen wij in elke priemfuik F_k een willekeurige toestand i_k en voegen wij aan het stelsel (7.81) de voorwaarden

$$v_{i_k} = 0 \qquad \text{voor } k = 1,\dots,m \tag{7.83}$$

toe, dan heeft het resulterende stelsel een unieke oplossing.

Bewijs:

Voor toestanden i en j, die niet tot eenzelfde (priem)fuik behoren, geldt

$$p_{ij} = 0. \tag{7.84}$$

Deze relatie heeft tot gevolg dat de vergelijkingen uit (7.81), behorende bij de toestanden uit een priemfuik F_k, onafhankelijk zijn van de overige vergelijkingen. Uit de stellingen 7.2 en 7.3 volgt nu dat de vergelijkingen uit (7.81), behorende bij de toestanden uit een priemfuik, oplosbaar zijn. De vergelijkingen uit (7.81), behorende bij de toestanden uit E_0 zijn wél afhankelijk van de overige vergelijkingen. De onbekenden y_i, v_i met $i \in E_0$ kunnen echter op eenduidige wijze in de overige onbekenden uitgedrukt worden. Definiëren wij de vektoren V_e en V_f met als componenten v_i met $i \in E_0$ resp. $i \notin E_0$, de vektoren Y_e en Y_f met als componenten y_i, waarbij $i \in E_0$ resp. $i \notin E_0$, de vektor H_e met als componenten h_i $(i \in E_0)$, de matrices $\mathfrak{R}_e$ en $\mathfrak{R}_f$ met als elementen de overgangskansen p_{ij}, waarbij $i,j \in E_0$ resp. $i \in E_0$, $j \notin E_0$, dan geldt

$$(7.85) \qquad \begin{cases} Y_e = (\mathfrak{I} - \mathfrak{R}_e)^{-1} \ \mathfrak{R}_f \, Y_f \ , \\[2ex] V_e = (\mathfrak{I} - \mathfrak{R}_e)^{-1} \ (H_e - Y_e + \mathfrak{R}_f V_f). \end{cases}$$

Hiermee is aangetoond dat het stelsel (7.81) oplosbaar is. De bewering b) van de stelling volgt direct uit stelling 7.3 en (7.85).

Om relatie (7.82) te bewijzen vermenigvuldigen wij beide leden van (7.81b) met de invariante kansen q_{ki} en sommeren over i. Wij vinden dan voor $k = 1,\ldots,N$

$$(7.86) \quad \sum_{i=1}^{N} q_{ki}\, y_i + \sum_{i=1}^{N} q_{ki}\, v_i - \sum_{i=1}^{N} q_{ki} \sum_{j=1}^{N} p_{ij}\, v_j = \sum_{i=1}^{N} q_{ki}\, h_i \ .$$

Door toepassing van (7.18) gaat (7.86) over in

$$(7.87) \qquad \sum_{k=1}^{N} q_{ki}\, y_i = \sum_{i=1}^{N} q_{ki}\, h_i \qquad\qquad \text{voor } k = 1,\ldots,N.$$

Uit (7.67) en (7.87) volgt nu de betrekking (7.82), waarmee het bewijs van bewering a) van de stelling voltooid is.

Opmerking 7.1

Laat a het aantal toestanden uit E_0 zijn ($a = 0$ als E_0 de lege verzameling is). Stelling 7.2 en relatie (7.84) hebben tot gevolg dat wij het stelsel (7.81) onder de voorwaarden (7.83) kunnen beschouwen als N+a lineaire vergelijkingen in N+a onbekenden. Deze onbekenden zijn de v_i (waarbij $i \neq i_1,\dots,i_m$), $y_{i_1},\dots,y_{i_m}$ en de y_i, waarbij $i \in E_0$. Zoals reeds in het bewijs van stelling 7.4 is opgemerkt, kunnen de vergelijkingen, behorende bij de toestanden uit een priemfuik, onafhankelijk van de overige vergelijkingen worden opgelost.

Wij keren terug naar het beslissingsproces behorende bij een strategie $z \in \mathcal{Z}$. Dit beslissingsproces wordt beschreven door een Markov-keten met $\mathcal{P}(z)$ als matrix van overgangskansen. Beschouwen wij het stelsel lineaire vergelijkingen

$$\left\{\begin{array}{llr} (7.88a) & y_i - \sum_{j=1}^{N} p_{ij}[z]\, y_j = 0 & \text{voor } i = 1,\dots,N, \\ (7.88b) & y_i + v_i - \sum_{j=1}^{N} p_{ij}[z]\, v_j = h(i,z(i)) & \text{voor } i = 1,\dots,N, \end{array}\right.$$

dan leert stelling 7.4 ons dat voor elke oplossing (Y,V) geldt

$$(7.89) \qquad y_i = \sum_{j=1}^{N} q_{ij}[z]\; h(j,z(j)) \qquad \text{voor } i = 1,\dots,N.$$

Het rechterlid van (7.89) is juist $y_i(z)$.*) De kriteriumfunktie $y_i(z)$ kan dus ook bepaald worden door een oplossing te berekenen van (7.88). Op het eerste gezicht lijkt dit geen winstpunt; ten onrechte echter! Wij kunnen namelijk met behulp van de grootheden y_i en v_i een iteratieprocedure opstellen, die in een eindig aantal stappen naar een optimale strategie convergeert.

---

*) Zowel uit de sterke wet van de grote aantallen als uit (7.74) en (7.89) tesamen kan worden afgeleid dat voor een begintoestand, die terugkeertoestand is, de gemiddelde opbrengst per tijdseenheid met kans 1 gelijk aan de verwachte opbrengst in het interval tussen twee opvolgende aannames van die toestand gedeeld door de verwachte lengte van dat tijdsinterval.

Om de betekenis van de grootheden v_i te doorgronden passen wij op (7.88) een k-voudig herhaalde substitutie toe. Wij vinden dan dat voor elke oplossing $(Y(z),V(z))$ van (7.88) geldt (voor $i = 1,\dots,N$ en $k \geqq 1$)

$$y_i(z) = \sum_{j=1}^{N} p_{ij}^{(k)}[z]\, y_j(z) \tag{7.90}$$

en

$$v_i(z) = \sum_{r=0}^{k-1} \sum_{j=1}^{N} p_{ij}^{(r)}[z]\{h(j,z(j)) - y_j(z)\} + \sum_{j=1}^{N} p_{ij}^{(k)}[z] v_j(z). \tag{7.91}$$

Bijgevolg geldt voor $i = 1,\dots,N$ en $k \geqq 1$

$$v_i(z) = \sum_{r=0}^{k-1} \sum_{j=1}^{N} p_{ij}^{(r)}[z] h(j,z(j)) - k y_i(z) + \sum_{j=1}^{N} p_{ij}^{(k)}[z] v_j(z). \tag{7.92}$$

Stel dat de toestanden s en t tot eenzelfde priemfuik behoren; dan volgt met behulp van (7.41) dat

$$\begin{aligned} v_s(z) - v_t(z) = {} & \sum_{r=0}^{k-1} \sum_{j=1}^{N} p_{sj}^{(r)}[z] h(j,z(j)) - \sum_{r=0}^{k-1} \sum_{j=1}^{N} p_{tj}^{(r)}[z] h(j,z(j)) + \\ & + \sum_{j=1}^{N} \{p_{sj}^{(k)}[z] - p_{tj}^{(k)}[z]\}\, v_j(z) \qquad \text{voor } k \geqq 1. \end{aligned} \tag{7.93}$$

Als het beslissingsproces geen enkele kringfuik bevat, geldt [vgl. (7.11) en (7.27)]

$$\lim_{k\to\infty} p_{sj}^{(k)}[z] = \lim_{k\to\infty} p_{tj}^{(k)}[z] \qquad \text{voor } j = 1,\dots,N. \tag{7.94}$$

Hieruit volgt

$$v_s(z) - v_t(z) = \lim_{k\to\infty} \left\{ \sum_{r=0}^{k-1} \sum_{j=1}^{N} p_{sj}^{(r)}[z] h(j,z(j)) - \sum_{r=0}^{k-1} \sum_{j=1}^{N} p_{tj}^{(r)}[z] h(j,z(j)) \right\}. \tag{7.95}$$

De uitdrukking tussen de accolades in (7.95) geeft voor begintoestanden s en t het verschil in totale verwachte opbrengst over de eerste k beslissingstijdstippen aan. Voor toestanden s en t uit eenzelfde priemfuik, die geen kringfuik bevat, geeft $v_s(z) - v_t(z)$ dus de stijging van de totale verwachte opbrengst over een oneindige tijdsperiode aan als het systeem start in toestand s in plaats van in toestand t. Dit maakt ook de reden duidelijk waarom de getallen $v_i(z)$ de "relative values" worden genoemd.

Wij beschouwen nu de door R.A. HOWARD ontwikkelde iteratieprocedure. Wij beginnen de eerste stap met een willekeurige strategie z_1. Hieronder wordt de n^{de} iteratiestap beschreven.

De n^{de} iteratiestap van HOWARD's "Policy-Iteration Method".

I) "Value Determination Procedure".

Aan het einde van de $(n-1)^{ste}$ stap werd de strategie z_n verkregen. Ga nu als volgt te werk:

a) bepaal voor de Markov-keten met $\mathcal{P}(z_n)$ als matrix van overgangskansen de splitsing van de toestandsruimte in de disjuncte verzamelingen $E_0^*, K_1, \ldots, K_m$, waarbij E_0^* de verzameling van alle doorgangstoestanden is en $K_1, \ldots, K_m$ kernfuiken zijn *). Kies in elke kernfuik K_h op ondubbelzinnige wijze een toestand i_h (stel: de toestand met de grootste index);

b) bereken de eenduidig bepaalde oplossing $(Y(z_n), V(z_n))$ van het stelsel lineaire vergelijkingen

$$(7.96)\qquad \begin{cases} y_i - \sum\limits_{j=1}^{N} p_{ij}[z_n]\, y_j = 0 & \text{voor } i = 1,\ldots,N, \\ y_i + v_i - \sum\limits_{j=1}^{N} p_{ij}[z_n]\, v_j = h(i,z_n(i)) & \text{voor } i = 1,\ldots,N, \\ v_{i_h} = 0 & \text{voor } h = 1,\ldots,m. \end{cases}$$

*) Een algoritme om zo'n splitsing te bepalen kan gevonden worden in B.L. FOX en D.M. LANDI, An algorithm for identifying the ergodic subchains and transient states of a stochastic matrix, Communications of the ACM, 11, (1968), p.619-621.

II) "Policy Improvement Operation".

Bepaal voor elke toestand i de verzameling $\mathfrak{D}_n(i)$ bestaande uit dié beslissing(en) $X \in \mathfrak{X}(i)$, waarvoor

$$Y_n(i,X) \overset{\text{def}}{=} \sum_{j=1}^{N} p_{ij}(X)\, y_j(z_n) \tag{7.97}$$

maximaal is. Maximaliseer vervolgens

$$V_n(i,X) \overset{\text{def}}{=} h(i,X) - \sum_{j=1}^{N} p_{ij}(X)\, y_j(z_n) + \sum_{j=1}^{N} p_{ij}(X)\, v_j(z_n) \tag{7.98}$$

naar $X \in \mathfrak{D}_n(i)$ en voeg aan i een maximaliserende beslissing toe. Wij spreken hierbij af, dat wij $X = z_n(i)$ kiezen, als $V_n(i,X)$ onder de voorwaarde $X \in \mathfrak{D}_n(i)$ maximaal is voor $X = z_n(i)$. Op deze wijze verkrijgen wij een strategie z_{n+1}. (Hoewel $Y_n(i,X)$ voor elke $X \in \mathfrak{D}_n(i)$ dezelfde waarde aanneemt, hebben wij ter wille van de notatie in de bewijsvoering deze uitdrukking in $V_n(i,X)$ opgenomen.)

Einde van de n^{de} stap.

Ingeval de matrix $\mathfrak{P}(z_n)$ geen disjuncte fuiken bevat, kan de n^{de} iteratiestap aanzienlijk vereenvoudigd worden (vgl. opmerking 7.3d op blz.105). Bovenstaande iteratieprocedure leidt tot steeds betere strategieën en convergeert na een eindig aantal stappen. Voor de volledigheid zullen wij hiervoor bewijzen geven. HOWARD heeft wel bewezen dat de iteratieprocedure tot steeds betere strategieën leidt, maar niet dat na een eindig aantal stappen convergentie plaatsvindt, m.a.w. dat "cycling" niet kan optreden.

De gemengde strategie $(z')z$ schrijft op het eerste beslissingstijdstip een beslissing voor volgens strategie $z' \in \mathfrak{Z}$ en daarna uitsluitend beslissingen volgens strategie $z \in \mathfrak{Z}$. Laat $(Y(z),V(z))$ een oplossing van (7.88) zijn. Wij definiëren dan voor $i = 1,\dots,N$

$$y_i((z')z) = \sum_{j=1}^{N} p_{ij}[z']\, y_j(z) \tag{7.99}$$

en

$$(7.100)\qquad v_i((z')z) = h(i,z'(i)) - y_i((z')z) + \sum_{j=1}^{N} p_{ij}[z']\ v_j(z).$$

Veronderstel dat de strategieën z_n en z_{n+1} achtereenvolgens verkregen zijn in HOWARD's iteratieprocedure. Dan geldt voor i = 1,...,N

$$(7.101)\qquad y_i((z_{n+1})z_n) = \max_{X\in\mathfrak{X}(i)} \{\sum_{j=1}^{N} p_{ij}(X)\ y_j(z_n)\}$$

en

$$(7.102)\quad v_i((z_{n+1})z_n) = \max_{X\in\mathfrak{D}_n(i)} \{h(i,X) - y_i((z_{n+1})z_n) + \sum_{j=1}^{N} p_{ij}(X)\ v_j(z_n)\}.$$

Verder geldt [vgl. (7.96)] voor i = 1,...,N

$$(7.103)\qquad y_i(z_n) = \sum_{j=1}^{N} p_{ij}[z_n]\ y_j(z_n)$$

en

$$(7.104)\qquad v_i(z_n) = h(i,z_n(i)) - y_i(z_n) + \sum_{j=1}^{N} p_{ij}[z_n]\ v_j(z_n).$$

Uit het bovenstaande volgt nu op eenvoudige wijze

Stelling 7.5

Voor elke i = 1,...,N geldt hetzij

$$(7.105)\qquad y_i((z_{n+1})z_n) > y_i(z_n)\ ,$$

hetzij

$$(7.106)\qquad y_i((z_{n+1})z_n) = y_i(z_n)\ ;\quad v_i((z_{n+1})z_n) \geq v_i(z_n).$$

De beslissing $z_{n+1}(i)$ is dan en slechts dan gelijk aan $z_n(i)$ als in de relatie (7.106) beide gelijktekens van toepassing zijn.

II) "Policy Improvement Operation".

Bepaal voor elke toestand i de verzameling $\mathfrak{D}_n(i)$ bestaande uit dié beslissing(en) $X \in \mathfrak{X}(i)$, waarvoor

$$(7.97) \qquad Y_n(i,X) \overset{\text{def}}{=} \sum_{j=1}^{N} p_{ij}(X)\, y_j(z_n)$$

maximaal is. Maximaliseer vervolgens

$$(7.98) \qquad V_n(i,X) \overset{\text{def}}{=} h(i,X) - \sum_{j=1}^{N} p_{ij}(X)\, y_j(z_n) + \sum_{j=1}^{N} p_{ij}(X)\, v_j(z_n)$$

naar $X \in \mathfrak{D}_n(i)$ en voeg aan i een maximaliserende beslissing toe. Wij spreken hierbij af, dat wij $X = z_n(i)$ kiezen, als $V_n(i,X)$ onder de voorwaarde $X \in \mathfrak{D}_n(i)$ maximaal is voor $X = z_n(i)$. Op deze wijze verkrijgen wij een strategie z_{n+1}. (Hoewel $Y_n(i,X)$ voor elke $X \in \mathfrak{D}_n(i)$ dezelfde waarde aanneemt, hebben wij ter wille van de notatie in de bewijsvoering deze uitdrukking in $V_n(i,X)$ opgenomen.)

Einde van de n^{de} stap.

Ingeval de matrix $\mathfrak{P}(z_n)$ geen disjuncte fuiken bevat, kan de n^{de} iteratiestap aanzienlijk vereenvoudigd worden (vgl. opmerking 7.3d op blz.105). Bovenstaande iteratieprocedure leidt tot steeds betere strategieën en convergeert na een eindig aantal stappen. Voor de volledigheid zullen wij hiervoor bewijzen geven. HOWARD heeft wel bewezen dat de iteratieprocedure tot steeds betere strategieën leidt, maar niet dat na een eindig aantal stappen convergentie plaatsvindt, m.a.w. dat "cycling" niet kan optreden.

De gemengde strategie $(z')z$ schrijft op het eerste beslissingstijdstip een beslissing voor volgens strategie $z' \in \mathfrak{Z}$ en daarna uitsluitend beslissingen volgens strategie $z \in \mathfrak{Z}$. Laat $(Y(z),V(z))$ een oplossing van (7.88) zijn. Wij definiëren dan voor $i = 1,\dots,N$

$$(7.99) \qquad y_i((z')z) = \sum_{j=1}^{N} p_{ij}[z']\, y_j(z)$$

en

$$(7.100) \qquad v_i((z')z) = h(i,z'(i)) - y_i((z')z) + \sum_{j=1}^{N} p_{ij}[z'] \; v_j(z).$$

Veronderstel dat de strategieën z_n en z_{n+1} achtereenvolgens verkregen zijn in HOWARD's iteratieprocedure. Dan geldt voor i = 1,...,N

$$(7.101) \qquad y_i((z_{n+1})z_n) = \max_{X \in \mathfrak{X}(i)} \{ \sum_{j=1}^{N} p_{ij}(X) \; y_j(z_n) \}$$

en

$$(7.102) \quad v_i((z_{n+1})z_n) = \max_{X \in \mathfrak{D}_n(i)} \{ h(i,X) - y_i((z_{n+1})z_n) + \sum_{j=1}^{N} p_{ij}(X) \; v_j(z_n) \}.$$

Verder geldt [vgl. (7.96)] voor i = 1,...,N

$$(7.103) \qquad y_i(z_n) = \sum_{j=1}^{N} p_{ij}[z_n] \; y_j(z_n)$$

en

$$(7.104) \qquad v_i(z_n) = h(i,z_n(i)) - y_i(z_n) + \sum_{j=1}^{N} p_{ij}[z_n] \; v_j(z_n).$$

Uit het bovenstaande volgt nu op eenvoudige wijze

<u>Stelling 7.5</u>

Voor elke i = 1,...,N geldt hetzij

$$(7.105) \qquad y_i((z_{n+1})z_n) > y_i(z_n) ,$$

hetzij

$$(7.106) \qquad y_i((z_{n+1})z_n) = y_i(z_n) \; ; \quad v_i((z_{n+1})z_n) \geq v_i(z_n).$$

De beslissing $z_{n+1}(i)$ is dan en slechts dan gelijk aan $z_n(i)$ als in de relatie (7.106) beide gelijktekens van toepassing zijn.

Stelling 7.6

Laten z en z' twee strategieën uit $\mathcal{Z}$ zijn en zij (Y(z),V(z)) een oplossing van (7.88). Als voor elke i = 1,...,N hetzij

(7.107) $$y_i((z')z) > y_i(z)$$

hetzij

(7.108) $$y_i((z')z) = y_i(z) \ ; \quad v_i((z')z) \geqq v_i(z)$$

van toepassing is, dan kunnen de volgende uitspraken gedaan worden:

a) als s een terugkeertoestand onder z' is, dan geldt

(7.109) $$y_s((z')z) = y_s(z) \ ;$$

b) voor elke i = 1,...,N geldt

(7.110) $$y_i(z') \geqq y_i((z')z) \geqq y_i(z) \ ;$$

c) als s een terugkeertoestand onder z' is en als bovendien voldaan is aan

(7.111) $$v_s((z')z) > v_s(z) \ ,$$

dan geldt

(7.112) $$y_s(z') > y_s(z) \ .$$

Deze stelling blijft van kracht als we het ">" teken vervangen door het "<" teken.

Bewijs:

Voor de strategieën z,z' $\in \mathcal{Z}$ geldt

(7.113) $$y_i((z')z) \geqq y_i(z) \qquad \text{voor } i = 1,\dots,N,$$

ofwel [zie definitie (7.99)]

(7.114) $$\sum_{j=1}^{N} p_{ij}[z'] \ y_j(z) \geqq y_i(z) \qquad \text{voor } i = 1,\dots,N.$$

Door herhaalde toepassing van deze relatie vinden wij met behulp van (7.6)

$$(7.115) \qquad \sum_{j=1}^{N} p_{ij}^{(k)}[z'] \; y_j(z) \geqq y_i((z')z) \geqq y_i(z) \qquad \text{voor } i = 1,\ldots,N;\ k \geqq 1.$$

Uit relatie (7.6) en definitie (7.99) volgt voor $i = 1,\ldots,N$ en $k \geqq 1$

$$(7.116) \qquad \sum_{j=1}^{N} p_{ij}^{(k)}[z'] \; y_j((z')z) = \sum_{j=1}^{N} p_{ij}^{(k+1)}[z'] \; y_j(z) \; .$$

Bijgevolg geldt voor $i = 1,\ldots,N$ en $k \geqq 1$

$$(7.117) \qquad \sum_{j=1}^{N} p_{ij}^{(k)}[z'] \; y_j((z')z) \geqq y_i((z')z) \geqq y_i(z).$$

Als wij van het linkerlid van (7.117) de Cesarolimiet nemen, dan vinden wij

$$(7.118) \qquad \sum_{j=1}^{N} q_{ij}[z'] \; y_j((z')z) \geqq y_i((z')z) \geqq y_i(z) \qquad \text{voor } i = 1,\ldots,N.$$

De relaties (7.99) en (7.18) impliceren achtereenvolgens

$$(7.119) \qquad \left\{ \begin{aligned} \sum_{j=1}^{N} q_{ij}[z'] \; y_j((z')z) &= \sum_{j=1}^{N} q_{ij}[z'] \sum_{h=1}^{N} p_{jh}[z'] \; y_h(z) = \\ &= \sum_{h=1}^{N} q_{ih}[z'] \; y_h(z) \qquad \text{voor } i = 1,\ldots,N. \end{aligned} \right.$$

Als s een terugkeertoestand onder z' is, geldt [vgl. (7.26)]

$$(7.120) \qquad q_{ss}[z'] > 0.$$

De relatie (7.119) geldt in het bijzonder voor $i = s$. Bewering a) van de stelling volgt nu met behulp van (7.113) en (7.120). *)

Uit (7.108) en (7.109) volgt

$$(7.121) \qquad v_j((z')z) \geqq v_j(z),$$

*) Met behulp van stelling 7.2 en definitie (7.99) kunnen wij uit bewering a) afleiden dat $y_s(z) = y_t(z)$ is, als s en t tot eenzelfde priemfuik van $\mathcal{P}(z')$ behoren.

als j een terugkeertoestand onder z' is. Met gebruikmaking van (7.25) geldt bijgevolg

$$(7.122)\qquad \sum_{j=1}^{N} q_{ij}[z']\; v_j((z')z) \geqq \sum_{j=1}^{N} q_{ij}[z']\; v_j(z) \qquad \text{voor } i = 1,\dots,N.$$

Als wij het linkerlid van (7.122) met behulp van (7.100) uitwerken, dan vinden wij voor i = 1,...,N

$$(7.123)\quad \begin{cases} \sum_{j=1}^{N} q_{ij}[z']\; \{h(j,z'(j)) - y_j((z')z)\} + \\ + \sum_{j=1}^{N} q_{ij}[z'] \sum_{h=1}^{N} p_{jh}[z']\; v_h(z) \geqq \sum_{j=1}^{N} q_{ij}[z']\; v_j(z)\;. \end{cases}$$

Door toepassing van (7.18) gaat (7.123) over in

$$(7.124)\qquad \sum_{j=1}^{N} q_{ij}[z']\; \{h(j,z'(j)) - y_j((z')z)\} \geqq 0 \qquad \text{voor } i = 1,\dots,N.$$

Bijgevolg geldt voor i = 1,...,N

$$(7.125)\qquad y_i(z') \overset{\text{def}}{=} \sum_{j=1}^{N} q_{ij}[z']\; h(j,z'(j)) \geqq \sum_{j=1}^{N} q_{ij}[z']\; y_j((z')z).$$

De relaties (7.118) en (7.125) impliceren bewering b) van de stelling.

Als s een terugkeertoestand onder z' is, waarvoor relatie (7.111) geldt, dan volgt met behulp van (7.120) dat voor de relaties (7.122) t/m (7.125) het ">" teken van toepassing is als i = s. Bewering c) van de stelling volgt nu direct met behulp van (7.118).

De stellingen 7.5 en 7.6 hebben tot gevolg:

Stelling 7.7

Als z_n en z_{n+1} twee strategieën zijn, die achtereenvolgens in de iteratieprocedure verkregen zijn, dan geldt

$$(7.126)\qquad y_i(z_{n+1}) \geqq y_i(z_n) \qquad \text{voor } i = 1,\dots,N.$$

Dus de iteratieprocedure leidt tot steeds betere strategieën. Om aan te tonen, dat in de iteratieprocedure geen "cycling" kan optreden, bewijzen wij eerst de volgende stelling:

Stelling 7.8

Als

(7.127) $$y_i(z_{n+1}) = y_i(z_n) \qquad \text{voor } i = 1,\dots,N,$$

dan is

(7.128) $$v_i(z_{n+1}) \geqq v_i(z_n) \qquad \text{voor } i = 1,\dots,N,$$

waarbij het ongelijkheidsteken geldt als $z_{n+1}(i) \neq z_n(i)$. Op de terugkeertoestanden van z_{n+1} schrijven de strategieën z_n en z_{n+1} dezelfde beslissingen voor en voor deze toestanden geldt in (7.128) het gelijkheidsteken.

Bewijs:

De relaties (7.99) en (7.127) hebben tot gevolg dat

(7.129) $$y_i((z_{n+1})z_n) = y_i(z_{n+1}) = y_i(z_n) \qquad \text{voor } i = 1,\dots,N.$$

Uit stelling 7.5 en (7.129) volgt dat

(7.130) $$v_i((z_{n+1})z_n) \geqq v_i(z_n) \qquad \text{voor } i = 1,\dots,N,$$

waarbij het ongelijkheidsteken uitsluitend geldt als $z_{n+1}(i) \neq z_n(i)$. Bewering c) van stelling 7.6 en (7.127) impliceren dat in (7.130) het gelijkheidsteken geldt, als i een terugkeertoestand onder z_{n+1} is. Bijgevolg geldt

(7.131) $$z_{n+1}(i) = z_n(i) \qquad \text{als i terugkeertoestand onder } z_{n+1} \text{ is.}$$

Uit deze relatie volgt dat de terugkeertoestanden van z_{n+1} ook terugkeertoestanden van z_n zijn, waarbij de overgangskansen tussen deze toestanden voor beide strategieën dezelfde zijn. Elke kernfuik van $\mathcal{P}(z_{n+1})$ is dus ook een kernfuik van $\mathcal{P}(z_n)$. De vergelijkingen uit (7.88), behorende bij toestanden uit een kernfuik, kunnen onafhankelijk van de overige vergelijkingen opgelost worden (vgl. opmerking 7.2, blz. 91). In de iteratieprocedure lossen wij deze vergelijkingen op door uit de kernfuik de toestand met de grootste index te nemen en daarvan de v-variabele nul te stellen. Wij verkrijgen dan een eenduidig bepaalde oplossing. Bijgevolg impliceert (7.131) dat

$$(7.132)\qquad v_i(z_{n+1}) = v_i(z_n) \qquad \text{als } i \text{ terugkeertoestand onder } z_{n+1} \text{ is.}$$

Met behulp van (7.100) en (7.129) gaat (7.130) voor $i = 1,\dots,N$ over in

$$(7.133)\qquad h(i,z_{n+1}(i)) - y_i(z_{n+1}) + \sum_{j=1}^{N} p_{ij}[z_{n+1}]\; v_j(z_n) \geqq v_i(z_n).$$

Door herhaalde toepassing van (7.133) vinden wij met behulp van (7.6)

$$(7.134)\quad \begin{cases} \sum\limits_{r=0}^{k-1} \sum\limits_{j=1}^{N} p_{ij}^{(r)}[z_{n+1}]\; \{h(j,z_{n+1}(j)) - y_j(z_{n+1})\} + \\[2ex] + \sum\limits_{j=1}^{N} p_{ij}^{(k)}[z_{n+1}]\; v_j(z_n) \geqq v_i((z_{n+1})z_n) \quad \text{voor } i = 1,\dots,N;\ k \geqq 1. \end{cases}$$

Als wij van (7.134) de Cesarolimiet nemen, verkrijgen wij

$$(7.135)\quad \begin{cases} \lim\limits_{m\to\infty} \frac{1}{m} \sum\limits_{k=1}^{m} \sum\limits_{r=0}^{k-1} \sum\limits_{j=1}^{N} p_{ij}^{(r)}[z_{n+1}]\; \{h(j,z_{n+1}(j)) - y_j(z_{n+1})\} + \\[2ex] + \sum\limits_{j=1}^{N} q_{ij}[z_{n+1}]\; v_j(z_n) \geqq v_i((z_{n+1})z_n) \qquad \text{voor } i = 1,\dots,N. \end{cases}$$

Door de Cesarolimiet van (7.91) te nemen, vinden wij dat $v_i(z_{n+1})$ voldoet aan

$$(7.136)\quad \begin{cases} v_i(z_{n+1}) = \lim\limits_{m\to\infty} \frac{1}{m} \sum\limits_{k=1}^{m} \sum\limits_{r=0}^{k-1} \sum\limits_{j=1}^{N} p_{ij}^{(r)}[z_{n+1}]\{h(j,z_{n+1}(j)) - y_j(z_{n+1})\} + \\[2ex] \qquad + \sum\limits_{j=1}^{N} q_{ij}[z_{n+1}]\; v_j(z_{n+1}) \qquad \text{voor } i = 1,\dots,N. \end{cases}$$

Uit (7.25) en (7.132) volgt dat het linkerlid van (7.135) gelijk is aan $v_i(z_{n+1})$. Bijgevolg geldt [vgl. relatie (7.130)]

(7.137) $$v_i(z_{n+1}) \geqq v_i((z_{n+1})z_n) \geqq v_i(z_n) \qquad \text{voor } i = 1,\ldots,N,$$

waarmee de stelling bewezen is.

Stelling 7.9

De iteratieprocedure leidt na een eindig aantal stappen n tot $z_{n+1} = z_n$. De strategie $z^* = z_n$ is dan optimaal.

Bewijs:

Omdat de klasse $\mathcal{Z}$ van strategieën uit eindig veel elementen bestaat, volgt uit de stellingen 7.7 en 7.8 direct dat de iteratieprocedure na een eindig aantal stappen convergeert. Laat z een willekeurige strategie uit $\mathcal{Z}$ zijn. De iteratieprocedure is geconvergeerd naar z^*, zodat voor elke $i = 1,\ldots,N$ geldt hetzij

(7.138) $$y_i((z)z^*) < y_i(z^*)$$

hetzij

(7.139) $$y_i((z)z^*) = y_i(z^*) \;;\quad v_i((z)z^*) \leqq v_i(z^*) \;.$$

Uit stelling 7.6 volgt nu

(7.140) $$y_i(z) \leqq y_i(z^*) \qquad \text{voor } i = 1,\ldots,N.$$

Omdat strategie z willekeurig uit $\mathcal{Z}$ gekozen is, is strategie z^* dus optimaal.

Opmerking 7.3

a) Veronderstel dat HOWARD's iteratieprocedure convergeert naar een strategie z^*. Wij definiëren $\mathrm{E}(z^*)$ als de verzameling van alle strategieën z, waarvoor (7.139) met de gelijkheidstekens van toepassing is; dus $\mathbf{E}(z^*)$ bevat z^*. Door de definities (7.99) en (7.100) toe te passen, zien wij dat $(Y(z^*),V(z^*))$ een oplossing is van (7.88) als $z \in \mathbf{E}(z^*)$. Bijgevolg geldt $Y(z) = Y(z^*)$, waarmee is aangetoond, dat elke strategie uit $\mathbf{E}(z^*)$ optimaal is.

b) Wij definiëren de verzameling

$$(7.141) \quad \mathcal{Z}' = \{z' \mid z' \in \mathcal{Z},\ y_i(z') \geqq y_i(z) \quad \text{voor } i = 1,\ldots,N \text{ en alle } z \in \mathcal{Z}\}.$$

Dus $\mathcal{Z}'$ is de verzameling van alle optimale strategieën, d.w.z. van alle strategieën met maximale gemiddelde opbrengst per tijdseenheid. Als $\mathcal{Z}'$ uit slechts één strategie bestaat, dan is dit zonder meer de beste strategie die te bedenken valt. Als $\mathcal{Z}'$ echter meer dan één strategie bevat, kan men zich afvragen of sommige strategieën van $\mathcal{Z}'$ niet de voorkeur verdienen boven de overige op grond van een meer genuanceerde opbrengstenbeschouwing. De gemiddelde opbrengst per tijdseenheid wordt immers alleen bepaald door de opbrengsten in de "long run"; de opbrengsten in de beginfase van het beslissingsproces beïnvloeden de gemiddelde opbrengst per tijdseenheid niet. Wij kunnen echter een kriterium formuleren, dat niet alleen een maat is voor de gemiddelde opbrengst per tijdseenheid, maar dat ook rekening houdt met de opbrengsten in de beginfase.

Daartoe definiëren wij $T_i(k;z)$ als de totale verwachte opbrengst over de eerste k beslissingstijdstippen, als de begintoestand i is en de strategie $z \in \mathcal{Z}$ wordt toegepast. Uit (7.92) volgt dat voor elke oplossing $(Y(z),V(z))$ van (7.88) geldt (voor $i = 1,\ldots,N$; $k \geqq 1$)

$$(7.142) \qquad v_i(z) = T_i(k;z) - k\, y_i(z) + \sum_{j=1}^{N} p_{ij}^{(k)}[z]\; v_j(z)\ .$$

Door de Cesarolimiet te nemen vinden wij dat elke oplossing $(Y(z),V(z))$ van (7.88) voor $i = 1,\ldots,N$ voldoet aan

$$(7.143) \qquad v_i(z) = \lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} \{T_i(k;z) - k\, y_i(z)\} + \sum_{j=1}^{N} q_{ij}[z]\; v_j(z)\ .$$

Als $(Y(z),V_1)$ en $(Y(z),V_2)$ voldoen aan (7.88) dan volgt door aftrekken dat $V_1 - V_2 = \mathfrak{P}(z)(V_1 - V_2)$. Door deze gelijkheid herhaald toe te passen en daarna de Cesarolimiet te nemen, vinden wij dat $V_1 - V_2 = \mathbf{Q}(z)(V_1 - V_2)$, waarbij $\mathbf{Q}(z) = ((q_{ij}[z]))$. Dus onder de voorwaarden

$$\sum_{j=1}^{N} q_{ij}[z]\; v_j(z) = 0 \qquad \text{voor } i = 1,\dots,N \tag{7.144}$$

heeft (7.88) een unieke oplossing. Geven wij deze oplossing aan met $(Y(z), W(z))$, dan geldt [vgl. (7.143)]

$$w_i(z) = \lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} \{T_i(k;z) - k\, y_i(z)\} \qquad \text{voor } i = 1,\dots,N. \tag{7.145}$$

Uit (7.145) kunnen wij eenvoudig afleiden dat voor een strategie $z'' \in \mathcal{Z}$ dan en slechts dan voor alle $z \in \mathcal{Z}$ geldt

$$\lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} \{T_i(k;z'') - T_i(k;z)\} \geqq 0 \qquad \text{voor } i = 1,\dots,N, \tag{7.146}$$

als

$$z'' \in \mathcal{Z}' \tag{7.147a}$$

en

$$w_i(z'') \geqq w_i(z') \qquad \text{voor } i = 1,\dots,N \text{ en alle } z' \in \mathcal{Z}'. \tag{7.147b}$$

Een strategie $z'' \in \mathcal{Z}$ wordt nu <u>(Y,W)-optimaal</u> genoemd als aan (7.147) is voldaan. BLACKWELL *) heeft aangetoond dat de klasse $\mathcal{Z}$ een (Y,W)-optimale strategie bevat. De kriteria (7.146) en (7.147) zijn equivalent. Een (Y,W)-optimale strategie heeft dus niet alleen de eigenschap dat de gemiddelde opbrengst per tijdseenheid maximaal is, maar is tevens "optimaal" te noemen met betrekking tot de opbrengsten in de beginfase van het beslissingsproces. HOWARD's iteratiemethode convergeert wel naar een strategie waarvoor de gemiddelde opbrengst per tijdseenheid maximaal

---

*) D. BLACKWELL, Discrete dynamic programming, The Annals of Mathematical Statistics, <u>33</u>, (1962), p. 719 - 726.

is, maar deze strategie behoeft niet (Y,W)-optimaal te zijn. HOWARD's iteratiemethode kan worden uitgebreid om een (Y,W)-optimale strategie te berekenen. *) Wij zullen ons echter beperken tot HOWARD's iteratiemethode.

Veronderstel dat, zoals in de praktijk vaak gebeurt, HOWARD's iteratiemethode convergeert naar een strategie z^* waarvoor $\mathbf{E}(z^*)$ alleen uit z^* bestaat (vgl. opmerking 7.3a). BLACKWELL heeft niet alleen bewezen dat z^* dan (Y,W)-optimaal is, maar hij heeft tevens aangetoond dat, als de opbrengsten verdisconteerd worden met een factor α, die voldoende dicht bij 1 ligt, de totale verwachte verdisconteerde opbrengst maximaal is voor strategie z^*.

c) Veronderstel dat $\mathfrak{Z}$ een strategie z bevat, zodat $\mathfrak{P}(z)$ geen disjuncte fuiken heeft, en dat onder z alle toestanden terugkeertoestanden zijn. HOWARD's iteratieprocedure convergeert dan naar een strategie z^*, waarvoor geldt

$$y_1(z^*) = \ldots = y_N(z^*). \tag{7.148}$$

Uit bewering a) van stelling 7.6 volgt immers dat voor $i = 1,\ldots,N$ geldt $\sum_{j=1}^{N} p_{ij}[z]\ y_j(z^*) = y_i(z^*)$; de relatie (7.148) volgt nu direct uit stelling 7.2.

d) In de praktijk komt het zelden voor dat de bij een strategie behorende Markov-keten disjuncte fuiken bevat. Zelfs als de klasse $\mathfrak{Z}$ dergelijke strategieën bevat, heeft vrijwel altijd de Markov-keten van een optimale strategie geen disjuncte fuiken. Als de matrix $\mathfrak{P}(z_n)$ geen disjuncte fuiken bevat, kan de n^{de} stap uit HOWARD's iteratieprocedure als volgt vereenvoudigd worden (vgl. de stellingen 7.2 t/m 7.4):

<u>De n^{de} iteratiestap als $\mathfrak{P}(z_n)$ geen disjuncte fuiken bevat.</u>

a) Kies een willekeurige toestand h en bereken de unieke oplossing

---

*) A.F. VEINOTT jr., On finding optimal policies in discrete dynamic programming with no discounting, The Annals of Mathematical Statistics, <u>37</u>, (1966), p. 1284 - 1294.

$(y(z_n), v_1(z_n), \ldots, v_N(z_n))$ van het stelsel lineaire vergelijkingen

$$(7.149)\qquad \begin{cases} y + v_i - \sum_{j=1}^{N} p_{ij}[z_n]\; v_j = h(i, z_n(i)) & \text{voor } i = 1,\ldots,N, \\ v_h = 0. \end{cases}$$

b) Voeg aan elke toestand i een beslissing $X \in \mathcal{X}(i)$ toe, welke

$$(7.150)\qquad V_n(i,X) \overset{\text{def}}{=} h(i,X) - y(z_n) + \sum_{j=1}^{N} p_{ij}(X)\; v_j(z_n)$$

maximaliseert. Ter wille van de convergentie kiezen wij $X = z_n(i)$ als deze beslissing $V_n(i,X)$ maximaliseert. Op deze wijze verkrijgen wij een strategie z_{n+1}.

<u>Einde n^de stap.</u>

Wij zullen HOWARD's iteratieprocedure op een aantal voorbeelden toepassen.

<u>Voorbeeld 7.3</u> <u>Het brouwerijprobleem, opgelost met HOWARD's iteratiemethode.</u>

De tekst van dit voorbeeld is opgenomen onder voorbeeld 7.2 (vgl. blz. 79). Wij berekenen nu de optimale strategie met de iteratieprocedure. De eerste stap van deze procedure beginnen wij met strategie z_1, waarbij $z_1(i)$ dié beslissing X is, welke $h(i,X)$ maximaliseert. Wij vinden

$$(7.151)\qquad z_1 = (0,1,0).$$

<u>Eerste stap</u>

a) Wij constateren eerst dat de matrix $\mathcal{P}(z_1)$ geen disjuncte fuiken bevat [vgl. (7.54)]. Vervolgens bepalen wij de oplossing van [vgl. (7.149)]

$$(7.152)\quad \begin{cases} y + v_1 - \frac{1}{2}v_1 - \frac{1}{2}v_2 & = -12 \,, \\ y + v_2 - \frac{3}{10}v_1 - \frac{1}{2}v_2 - \frac{2}{10}v_3 & = -9 \,, \\ y + v_3 - \frac{1}{2}v_1 - \frac{1}{2}v_2 & = -6 \,, \\ v_3 & = 0 \,. \end{cases}$$

De oplossing wordt gegeven door

$$(7.153)\quad y(z_1) = -9{,}9 \;;\; v_1(z_1) = -6 \;;\; v_2(z_1) = -1{,}8 \;;\; v_3(z_1) = 0 \,.$$

b) Om de strategie z_2 te bepalen, moeten wij de testgrootheid $V_1(i,X)$ bij vaste i naar X maximaliseren. In tabel 7.2 worden de numerieke waarden van deze testgrootheid gegeven.

i	X	$V_1(i,X)$	
1	0	$-12 + 9{,}9 - \frac{1}{2}.6 - \frac{1}{2}.1{,}8 = -6$	←
2	0	$-10 + 9{,}9 - \frac{1}{2}.6 - \frac{1}{2}.1{,}8 = -4$	
	1	$-9 + 9{,}9 - \frac{3}{10}.6 - \frac{5}{10}.1{,}8 = -1{,}8$	←
3	0	$-6 + 9{,}9 - \frac{1}{2}.6 - \frac{1}{2}.1{,}8 = 0$	
	1	$-7 + 9{,}9 - \frac{3}{10}.6 - \frac{5}{10}.1{,}8 = 0{,}2$	
	2	$-8{,}85 + 9{,}9 - \frac{3}{10}.1{,}8 = 0{,}51$	←

Tabel 7.2

De numerieke waarden van $V_1(i,X)$

Bijgevolg geldt

$$(7.154)\quad z_2 = (0,1,2).$$

Einde eerste stap.

Tweede stap

a) De matrix $\mathcal{P}(z_2)$ bevat geen disjuncte fuiken. De oplossing van het stelsel

$$(7.155)\qquad \left\{\begin{array}{llll} y + v_1 - \frac{1}{2} v_1 - \frac{1}{2} v_2 & & = -12 & , \\ y + v_2 - \frac{3}{10} v_1 - \frac{1}{2} v_2 - \frac{1}{2} v_3 & & = -9 & , \\ y + v_3 \qquad - \frac{3}{10} v_2 - \frac{7}{10} v_3 & & = -8{,}85 & , \\ & v_3 & = 0 & \end{array}\right.$$

wordt gegeven door

$$(7.156)\qquad y(z_2) = -9{,}75 \;;\; v_1(z_2) = -7{,}5 \;;\; v_2(z_2) = -3 \;;\; v_3(z_2) = 0 \,.$$

b) Vervolgens berekenen wij de testgrootheid $V_2(i,X)$, waarvan de waarden in tabel 7.3 gegeven worden.

i	X	$V_2(i,X)$	
1	0	$-12 + 9{,}75 - \frac{1}{2}\cdot 7{,}5 - \frac{1}{2}\cdot 3 = -7{,}5$	←
2	0	$-10 + 9{,}75 - \frac{1}{2}\cdot 7{,}5 - \frac{1}{2}\cdot 3 = -5{,}5$	
	1	$-9 + 9{,}75 - \frac{3}{10}\cdot 7{,}5 - \frac{1}{2}\cdot 3 = -3$	←
3	0	$-6 + 9{,}75 - \frac{1}{2}\cdot 7{,}5 - \frac{1}{2}\cdot 3 = -1{,}5$	
	1	$-7 + 9{,}75 - \frac{3}{10}\cdot 7{,}5 - \frac{1}{2}\cdot 3 = -1$	
	2	$-8{,}85 + 9{,}75 - \quad - \frac{3}{10}\cdot 3 = 0$	←

Tabel 7.3

De numerieke waarden van $V_2(i,X)$

Bijgevolg geldt

(7.157) $$z_3 = (0,1,2).$$

Einde tweede stap.

De strategieën z_2 en z_3 zijn aan elkaar gelijk en dus is $z^* = (0,1,2)$ een optimale strategie.

In voorbeeld 7.4 beginnen wij HOWARD's iteratiemethode met een strategie waarvan de matrix van overgangskansen disjuncte fuiken bevat.

Voorbeeld 7.4 Een produktieprobleem.

Een fabrikant vraagt zich elke ochtend af hoeveel eenheden van een bepaald artikel hij die dag zal produceren. Ten hoogste 2 eenheden kunnen geproduceerd worden en een produktie is voor de middag beëindigd. De produktiekosten bedragen 6 resp. 10 voor een produktie van 1 resp. 2 eenheden. Het omschakelen van de ene op de andere dag van produceren naar niet produceren brengt kosten 2 met zich mee, terwijl met de omschakeling van niet produceren naar produceren kosten 4 gepaard gaan.

De vraag naar het produkt treedt 's middags op; elke middag is de vraag met kans 0,25 gelijk aan 1 en met kans 0,75 gelijk aan 2. Als de voorraad ontoereikend is om aan de vraag te voldoen, verricht de fabrikant een noodinkoop om het tekort op te heffen; de noodinkoopkosten zijn 12 per eenheid. Maximaal 2 eenheden kunnen 's avonds in voorraad gehouden worden. Voor elke eenheid die 's avonds in voorraad is, zijn de voorraadkosten 4.

Gevraagd wordt een optimaal produktieschema op te stellen.

Oplossing

De toestand van het systeem wordt bepaald door de 's ochtends aanwezige voorraad en het feit of de vorige dag al dan niet een produktie heeft plaatsgevonden. De toestandsgrootheid bestaat uit twee componenten, aan te geven met S_1 en S_2. De component S_1 is gelijk aan de 's ochtends aanwezige voorraad en kan uitsluitend de waarden 0, 1, 2 aannemen; de component S_2 kan alleen 0 of 1 zijn, n.l. 0 als de vorige dag geen produktie heeft plaatsgevonden en 1 anders.

De beslissing X geeft de grootte van de produktie aan; X = 0 betekent dat niet geproduceerd wordt. De verzameling van toegelaten beslissingen in toestand (S_1,S_2) wordt gegeven door

$$(7.158) \qquad \mathcal{X}(S_1,S_2) = \begin{cases} \{0,\ldots,3-S_1\} & \text{als } S_1 \neq 0, \\ \{0,1,2\} & \text{als } S_1 = 0. \end{cases}$$

Elke middag wordt tenminste 1 eenheid gevraagd en de maximale produktiegrootte is 2, zodat de toestand (2,0) niet kan optreden. Wij laten deze toestand derhalve buiten beschouwing. Ter wille van de notatie nummeren wij de toestanden (S_1,S_2) = (0,0), (1,0), (0,1), (1,1) en (2,1) als i = 1, 2, 3, 4 resp. 5.

In tabel 7.4 op blz. 111 staan de nùmerieke waarden van de overgangskansen $p_{ij}(X)$ en de opbrengstfunktie $h(i,X)$.

Wij beginnen HOWARD's iteratiemethode met strategie

$$(7.159) \qquad z_1 = (1,1,1,2,1).$$

Eerste stap

a) De matrix $\mathcal{P}(z_1)$ van overgangskansen wordt gegeven door

$$(7.160) \qquad \mathcal{P}(z_1) = \begin{array}{c} \begin{array}{ccccc} (1) & (2) & (3) & (4) & (5) \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{3}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{3}{4} & \frac{1}{4} \\ 0 & 0 & 0 & \frac{3}{4} & \frac{1}{4} \end{pmatrix} \end{array} \begin{array}{c} (1) \\ (2) \\ (3) \\ (4) \\ (5). \end{array}$$

Toestand i	Beslissing X	Overgangskansen $p_{ij}(X)$ $j = 1\ 2\ 3\ 4\ 5$	Opbrengstfunktie $h(i,X)$
$1 \equiv (0,0)$	0	1 0 0 0 0	$-12-\frac{3}{4}.12 = -21$
	1	0 0 1 0 0	$-4-6-\frac{3}{4}.12 = -19$
	2	0 0 $\frac{3}{4}$ $\frac{1}{4}$ 0	$-4-10-\frac{1}{4}.4 = -15$
$2 \equiv (1,0)$	0	1 0 0 0 0	$-\frac{3}{4}.12 = -9$
	1	0 0 $\frac{3}{4}$ $\frac{1}{4}$ 0	$-4-6-\frac{1}{4}.4 = -11$
	2	0 0 0 $\frac{3}{4}$ $\frac{1}{4}$	$-4-10-4-\frac{1}{4}.4 = -19$
$3 \equiv (0,1)$	0	1 0 0 0 0	$-2-12-\frac{3}{4}.12 = -23$
	1	0 0 1 0 0	$-6-\frac{3}{4}.12 = -15$
	2	0 0 $\frac{3}{4}$ $\frac{1}{4}$ 0	$-10-\frac{1}{4}.4 = -11$
$4 \equiv (1,1)$	0	1 0 0 0 0	$-2-\frac{3}{4}.12 = -11$
	1	0 0 $\frac{3}{4}$ $\frac{1}{4}$ 0	$-6-\frac{1}{4}.4 = -7$
	2	0 0 0 $\frac{3}{4}$ $\frac{1}{4}$	$-10-4-\frac{1}{4}.4 = -15$
$5 \equiv (2,1)$	0	$\frac{3}{4}$ $\frac{1}{4}$ 0 0 0	$-2-\frac{1}{4}.4 = -3$
	1	0 0 0 $\frac{3}{4}$ $\frac{1}{4}$	$-6-4-\frac{1}{4}.4 = -11$

Tabel 7.4

De numerieke waarden van $p_{ij}(X)$ en $h(i,X)$.

Deze matrix bevat twee kernfuiken. Zowel de verzameling {3} als de verzameling {4,5} is een kernfuik. De toestanden 1 en 2 zijn doorgangstoestanden. Vanuit toestand 1 is uitsluitend de kernfuik {3} bereikbaar, terwijl vanuit toestand 2 beide kernfuiken bereikbaar zijn. De verzameling {1,3} is dus een priemfuik. Uit deze analyse volgt

$$(7.161) \qquad y_1(z_1) = y_3(z_1) \; ; \; y_4(z_1) = y_5(z_1).$$

Het stelsel vergelijkingen (7.96) wordt dus voor strategie z_1

$$(7.162) \qquad \begin{cases} y_2 - \frac{3}{4} y_3 - \frac{1}{4} y_5 & = 0 \, , \\ y_3 + v_1 - v_3 & = -19 \, , \\ y_2 + v_2 - \frac{3}{4} v_3 - \frac{1}{4} v_4 & = -11 \, , \\ y_3 + v_3 - v_3 & = -15 \, , \\ y_5 + v_4 - \frac{3}{4} v_4 - \frac{1}{4} v_5 & = -15 \, , \\ y_5 + v_5 - \frac{3}{4} v_4 - \frac{1}{4} v_5 & = -11 \, , \\ v_3 = v_5 & = 0 \, . \end{cases}$$

Deze zes vergelijkingen in zes onbekenden kunnen eenvoudig met de hand opgelost worden (vgl. opmerking 7.2 op blz. 91). De oplossing luidt

$$(7.163) \begin{cases} y_2(z_1) = -14{,}75 \; ; \; y_3(z_1) = -15 \; ; \; y_5(z_1) = -14 \; ; \; v_1(z_1) = -4 \; ; \\ v_2(z_1) = 2{,}75 \; ; \; v_3(z_1) = 0 \; ; \; v_4(z_1) = -4 \; ; \; v_5(z_1) = 0 \, . \end{cases}$$

b) Vervolgens passen wij de "Policy Improvement Operation" (zie blz. 95) toe om strategie z_2 te bepalen. Daartoe berekenen wij eerst de testgrootheid $Y_1(i,X)$, waarvan de numerieke waarden gegeven worden in tabel 7.5 op blz. 113.

Voor elke toestand i geldt dat $Y_1(i,X)$ voor precies één $X \in \mathcal{X}(i)$ maximaal is; bijgevolg behoeven wij de tweede testgrootheid $V_1(i,X)$ niet te beschouwen. Uit tabel 7.5 op blz. 113 volgt

$$(7.164) \qquad z_2 = (2,2,2,2,1).$$

<u>Einde eerste stap.</u>

Toestand i	Beslissing $X \in \mathfrak{X}(i)$	Testgrootheid $Y_1(i,X) = \sum_{j=1}^{5} p_{ij}(x)\, y_j(z_1)$	
1	0	$-1.15 = -15$	
	1	$-1.15 = -15$	
	2	$-\frac{3}{4}.15-\frac{1}{4}.14 = -14\frac{3}{4}$	←
2	0	$-1.15 = -15$	
	1	$-\frac{3}{4}.15-\frac{1}{4}.14 = -14\frac{3}{4}$	
	2	$-\frac{3}{4}.14-\frac{1}{4}.14 = -14$	←
3	0	$-1.15 = -15$	
	1	$-1.15 = -15$	
	2	$-\frac{3}{4}.15-\frac{1}{4}.14 = -14\frac{3}{4}$	←
4	0	$-1.15 = -15$	
	1	$-\frac{3}{4}.15-\frac{1}{4}.14 = -14\frac{3}{4}$	
	2	$-\frac{3}{4}.14-\frac{1}{4}.14 = -14$	←
5	0	$-\frac{3}{4}.15-\frac{1}{4}.14\frac{3}{4} = -14\frac{15}{16}$	
	1	$-\frac{3}{4}.15-\frac{1}{4}.14 = -14$	←

Tabel 7.5

De numerieke waarden van de testgrootheid $Y_1(i,X)$

Tweede stap

a) Schrijven wij de matrix $\mathcal{P}(z_2)$ op, dan blijkt deze matrix geen disjuncte fuiken te bevatten, zodat $y_i(z_2) = y(z_2)$ voor alle i. Wij kunnen ons dus beperken tot het stelsel vergelijkingen [kies bijv. h = 5 in (7.149)]

$$(7.165)\quad \begin{cases} y + v_1 - \frac{3}{4} v_3 - \frac{1}{4} v_4 & = -15\ , \\ y + v_2 \qquad\quad - \frac{3}{4} v_4 - \frac{1}{4} v_5 & = -19\ , \\ y + v_3 - \frac{3}{4} v_3 - \frac{1}{4} v_4 & = -11\ , \\ y + v_4 \qquad\quad - \frac{3}{4} v_4 - \frac{1}{4} v_5 & = -15\ , \\ y + v_5 \qquad\quad - \frac{3}{4} v_4 - \frac{1}{4} v_5 & = -11\ , \\ v_5 & = \ 0\ . \end{cases}$$

De oplossing wordt gegeven door

$$(7.166)\quad \begin{cases} y(z_2) = -14\ ;\ v_1(z_2) = \ 4\ ;\ v_2(z_2) = -8\ ; \\ v_3(z_2) = \ 8\ ;\ v_4(z_2) = -4\ ;\ v_5(z_2) = \ 0\ . \end{cases}$$

b) Omdat $\mathcal{P}(z_2)$ geen disjuncte fuiken heeft, vinden wij strategie z_3 door de testgrootheid $V_2(i,X)$ naar $X \in \mathfrak{X}(i)$ te maximaliseren. Als wij dit uitvoeren, verkrijgen wij

$$(7.167)\qquad z_3 = (2,0,2,1,0).$$

Einde tweede stap.

Derde stap

a) De matrix $\mathcal{P}(z_3)$ van overgangskansen heeft geen disjuncte fuiken. De oplossing van het met strategie z_3 corresponderende stelsel vergelijkingen (7.149) wordt gegeven door (wij hebben $h = 4$ genomen)

$$(7.168)\quad \begin{cases} y(z_3) = -10\ ;\ v_1(z_3) = -8\ ;\ v_2(z_3) = -7\ ; \\ v_3(z_3) = -4\ ;\ v_4(z_3) = \ 0\ ;\ v_5(z_3) = -0{,}75. \end{cases}$$

b) De maximalisatie van $V_3(i,X)$ naar $X \in \mathfrak{X}(i)$ leidt tot

$$(7.169)\qquad z_4 = (2,1,2,1,0).$$

Einde derde stap.

Vierde stap

a) De matrix $\mathcal{P}(z_4)$ bevat geen disjuncte fuiken. De oplossing van het met strategie z_4 corresponderende stelsel vergelijkingen (7.149) wordt gegeven door (wij hebben $h = 4$ genomen)

$$(7.170) \qquad \begin{cases} y(z_4) = -10 \; ; \; v_1(z_4) = -8 \; ; \; v_2(z_4) = -4 \; ; \\ v_3(z_4) = -4 \; ; \; v_4(z_4) = 0 \; ; \; v_5(z_4) = 0 \; . \end{cases}$$

b) De maximalisatie van $V_4(i,X)$ naar $X \in \mathfrak{X}(i)$ leidt tot

$$(7.171) \qquad z_5 = (2,1,2,1,0).$$

Einde vierde stap.

De strategieën z_4 en z_5 zijn hetzelfde en dus is $z^* = (2,1,2,1,0)$ optimaal. Wij merken op dat in dit voorbeeld de relaties (7.168) en (7.170) een toepassing vertegenwoordigen van stelling 7.8.

De iteratiemethode van HOWARD kan aanzienlijk worden vereenvoudigd als de toestandsveranderingen deterministisch zijn.

Opmerking 7.4 Het deterministische ∞-stapsbeslissingsprobleem.

Wij nemen nu aan dat de overgangskansen voldoen aan

$$(7.172) \qquad p_{ij}(X) = 1 \text{ of } 0 \qquad \text{voor alle } X \in \mathfrak{X}(i) \; ; \; i,j = 1,\ldots,N.$$

In dit geval spreken wij van een deterministisch ∞-stapsbeslissingsprobleem. De berekeningen in HOWARD's iteratiemethode kunnen nu aanzienlijk verminderd worden. De grootheden y_i en v_i kunnen n.l. successievelijk berekend worden zonder een stelsel lineaire vergelijkingen op te lossen. Om dit nader toe te lichten, definiëren wij voor een terugkeertoestand van een Markov-keten een cyclus als het interval tussen twee opéénvolgende tijdstippen waarop het systeem in die terugkeertoestand is. Stel dat toestand t onder strategie z een terugkeertoestand is. Neem aan dat in de cyclus behorende bij toestand t achtereenvolgens de toestanden (bedenk dat het beslissingsproces deterministisch is)

$$t = i_1 \to \ldots \to i_j \to \ldots i_M \to i_{M+1} = i_1 = t \tag{7.173}$$

aangenomen worden. Uit eigenschap b) op blz. 71 volgt dat $\{i_1,\ldots,i_M\}$ een kernfuik is. Dus $y_{i_k}(z) = y_t(z)$ voor $1 \leq k \leq M$. Uit stelling 7.4 volgt dat getallen v_{i_k} $(1 \leq k \leq M)$ bestaan met $y_t(z) + v_{i_k} - v_{i_{k+1}} = h(i_k,z(i_k))$ voor $1 \leq k \leq M$. Tellen wij deze M gelijkheden op, dan vinden wij (vgl. ook de voetnoot op blz. 92)

$$y_t(z) = \frac{1}{M}\sum_{k=1}^{M} h(i_k,z(i_k)). \tag{7.174}$$

Het beslissingsproces is deterministisch; dus het kan niet voorkomen dat vanuit een toestand twee of meer disjuncte fuiken bereikbaar zijn. Bijgevolg behoort elke toestand tot een priemfuik. Voor toestanden uit eenzelfde priemfuik heeft de kriteriumfunktie $y_i(z)$ dezelfde waarde. Om deze waarde te berekenen, kiezen wij in de priemfuik een terugkeertoestand t en berekenen $y_t(z)$ met behulp van (7.174).

Als eenmaal de $y_i(z)$ bepaald zijn, kunnen wij successievelijk waarden voor de grootheden $v_i(z)$ berekenen en wel op de volgende wijze. Stel in elke kernfuik van $\mathcal{P}(z)$ de v-waarde van een of andere toestand uit die kernfuik gelijk aan nul. Bij elke toestand a behoort één toestand b met $p_{ab}[z] = 1$ en omgekeerd behoort bij elke toestand b tenminste één toestand a met $p_{ab}[z] = 1$. Als nu $p_{ab}[z] = 1$ (dus $p_{aj}[z] = 0$ als $j \neq b$), dan geldt

$$y_a(z) + v_a(z) - v_b(z) = h(a,z(a)). \tag{7.175}$$

Bijgevolg kunnen wij successievelijk waarden voor de grootheden $v_i(z)$ berekenen. Wij geven een toepassing aan de hand van voorbeeld 7.5, waarin sprake is van een seizoensafhankelijke vraag.

Voorbeeld 7.5 Een voorraadprobleem met "seizoensvraag".

De beheerder van een magazijn beschikt over de volgende gegevens. De vraag naar een bepaald artikel is seizoensafhankelijk en wordt, voor een onbegrensd aantal jaren, gegeven in tabel 7.6.

maand (j)	1	2	3	4	5	6	7	8	9	10	11	12
vraag (d_j)	1	2	3	4	5	6	7	8	9	1	2	3

Tabel 7.6

Uit deze tabel blijkt bijvoorbeeld, dat de vraag in de maand april van elk jaar 4 eenheden bedraagt. De vraag vindt steeds plaats op de eerste dag van de betrokken maand. Indien de aanwezige voorraad kleiner is dan de vraag, mag het ontbrekende worden nageleverd, met dien verstande echter dat de achterstand in de levering nooit meer dan 10 eenheden mag bedragen. Voor iedere nageleverde eenheid is de winstderving $c_2 = 3$ voor elke maand dat te laat wordt geleverd. De voorraad kan eens per maand worden aangevuld door het plaatsen van een bestelling. Een bestelling, geplaatst in maand j, wordt afgeleverd in de ochtend van de eerste dag van de maand $j + 1$. De bestelkosten van q eenheden ($q > 0$) worden gegeven door $\phi(q) = 5q + 20$. In het magazijn kunnen ten hoogste 20 eenheden worden opgeslagen. Voor elke eenheid bedragen de voorraadkosten $c_1 = 1$ per maand. Goederen, welke worden verkocht op de dag waarop ze zijn binnengekomen, worden niet eerst in het magazijn opgeslagen. Dus als op de tweede dag van de maand j de voorraad i eenheden bedraagt, kunnen in die maand maximaal $20 - i + d_{j+1}$ eenheden besteld worden.

De beheerder vraagt zich af voor welke bestelstrategie de gemiddelde kosten per maand, over een onbegrensd aantal jaren, minimaal zijn.

Oplossing

De toestand van het systeem wordt beschreven door de voorraad i op de tweede dag van de betrokken maand j. Wij nemen als toestandsruimte

(7.176) $\mathcal{S} = \{(i,j) \mid i,j \text{ geheel}; -10 \leq i \leq 20; 1 \leq j \leq 12\}$.

De beslissing X geeft de grootte van de bestelling aan. De verzameling van toegelaten beslissingen in toestand (i,j) is gelijk aan

$$(7.177) \quad \begin{cases} \mathcal{X}((i,j)) = \{X \mid 0 \leqq X \leqq 20 - i + d_{j+1}\} \\ \qquad\qquad\qquad\qquad \text{voor } -10 \leq i \leq 20;\ 1 \leq j \leq 12, \end{cases}$$

waarbij $d_{13} = d_1$. Als in toestand (i,j) beslissing X genomen wordt, is

$$(7.178) \qquad (i + X - d_{j+1} \text{ , } j + 1)$$

de toestand op het volgende beslissingstijdstip. Wij identificeren toestand $(i + X - d_{13}\text{ , } 13)$ met toestand $(i + X - d_1\text{ , } 1)$.

De opbrengstfunktie wordt gegeven door

$$(7.179) \quad h((i,j),X) = \begin{cases} -i - \phi(X) & \text{voor } \quad 0 \leq i \leq 20;\ 1 \leq j \leq 12, \\ 3i - \phi(X) & \text{voor } -10 \leq i < \ 0;\ 1 \leq j \leq 12, \end{cases}$$

waarbij $\phi(0) = 0$ en $\phi(X) = 5X + 20$ als $X > 0$.

Wij beginnen HOWARD's iteratiemethode met de volgende strategie:

$$(7.180) \quad z_1((i,j)) = 20 - i \qquad \text{voor } -10 \leq i \leq 20;\ 1 \leq j \leq 12.$$

<u>Eerste stap</u>

a) Eerst gaan wij na of de matrix $\mathcal{P}(z_1)$ disjuncte fuiken bevat of niet. Het feit dat het beslissingsproces deterministisch is, vereenvoudigt dit onderzoek. Ongeacht de uitgangstoestand komt het systeem na een eindig aantal stappen in één van de toestanden (i,1) terecht. Omdat $z_1((i,1)) = 20 - i$ en de vraag in de tweede maand gelijk aan 2 is, neemt het systeem na elk van de toestanden (i,1) als volgende toestand (18,2) aan. Bijgevolg komt het systeem ongeacht de uitgangstoestand na een eindig aantal stappen in toestand (18,2) terecht. Dus de matrix $\mathcal{P}(z_1)$ bevat geen disjuncte fuiken en toestand (18,2) is een terugkeertoestand. De cyclus behorende bij deze toestand ziet er als volgt uit:

$$(7.181)\quad \begin{cases} (18,2) \to (17,3) \to (16,4) \to (15,5) \to (14,6) \to (13,7) \to (12,8) \to \\ \to (11,9) \to (19,10) \to (18,11) \to (17,12) \to (19,1) \to (18,2). \end{cases}$$

Deze cyclus bestaat uit 12 maanden. Uit (7.174) en (7.179) volgt

$$(7.182)\quad y(z_1) = \frac{1}{12}\{-48-52-56-60-64-68-72-76-44-48-52-44\} = -57.$$

Om de waarden voor de $v_{(i,j)}(z_1)$ te berekenen stellen wij

$$(7.183)\quad v_{(18,2)}(z_1) = 0.$$

De overige v-waarden kunnen wij nu successievelijk uitrekenen. Zo wordt na toestand (i,1) als volgende toestand (18,2) aangenomen, zodat voor alle i geldt

$$(7.184)\quad y(z_1) + v_{(i,1)}(z_1) - v_{(18,2)}(z_1) = h((i,1),z_1((i,1))).$$

De volgende toestand na toestand (i,12) is (19,1), zodat voor alle i geldt

$$(7.185)\quad y(z_1) + v_{(i,12)}(z_1) - v_{(19,1)}(z_1) = h((i,12),z_1((i,12))).$$

Op deze wijze voortgaand vinden wij alle v-waarden.

b) De maximalisatie van $V_1((i,j),X)$ naar $X \in \mathcal{X}((i,j))$ leidt voor alle i en j (de berekeningen zijn uitgevoerd op een rekenautomaat) tot

$$(7.186)\quad z_2((i,j)) = \begin{cases} S^{(j)} - i & \text{als } i < s^{(j)}, \\ 0 & \text{als } i \geqq s^{(j)} \end{cases}$$

waarbij de getallen $s^{(j)}$ en $S^{(j)}$ gegeven worden in tabel 7.7 op blz. 120.

j	$(s^{(j)},S^{(j)})$	j	$(s^{(j)},S^{(j)})$	j	$(s^{(j)},S^{(j)})$
1	(-1,9)	5	(2,13)	9	(-1,6)
2	(0,12)	6	(3,15)	10	(1,6)
3	(1,15)	7	(3,23)	11	(0,4)
4	(1,11)	8	(7,15)	12	(-5,1)

Tabel 7.7

De strategie z_2

Einde eerste stap.

Tweede stap

a) Uit de analyse van de matrix $\mathcal{P}(z_2)$ blijkt dat vanuit elke toestand $(i,1)$ het systeem na een eindig aantal stappen terecht komt in toestand $(6,9)$. Bijgevolg wordt toestand $(6,9)$ vanuit elke toestand bereikt. Dit betekent dat $\mathcal{P}(z_2)$ geen disjuncte fuiken heeft en dat toestand $(6,9)$ een terugkeertoestand is. De cyclus behorende bij deze toestand ziet er als volgt uit:

$$(7.187)\quad \begin{cases} (6,9) \to (5,10) \to (3,11) \to (0,12) \to (-1,1) \to (-3,2) \to (9,3) \to \\ \to (5,4) \to (0,5) \to (7,6) \to (0,7) \to (15,8) \to (6,9). \end{cases}$$

Deze cyclus bestaat uit 12 maanden. Uit (7.174) en (7.179) volgt

$$(7.188)\qquad y(z_2) = \frac{1}{12}\{-6-5-3-0-3-104-9-5-85-7-135-15\} = -31\frac{5}{12}\ .$$

Door $v_{(6,9)}(z_2) = 0$ te stellen kunnen vervolgens de overige v-waarden successievelijk berekend worden.

b) De maximalisatie van $V_2((i,j),X)$ naar $X \in \mathfrak{X}((i,j))$ leidt voor alle i en j tot

$$(7.189)\qquad z_3((i,j)) = \begin{cases} S^{(j)} - i & \text{als } i < s^{(j)}\ , \\ 0 & \text{als } i \geq s^{(j)}\ , \end{cases}$$

waarbij de getallen $s^{(j)}$ en $S^{(j)}$ gegeven worden in tabel 7.8.

j	$(s^{(j)},S^{(j)})$	j	$(s^{(j)},S^{(j)})$	j	$(s^{(j)},S^{(j)})$
1	(0,9)	5	(1,13)	9	(-1,7)
2	(-1,12)	6	(3,15)	10	(0,8)
3	(1,15)	7	(3,18)	11	(1,6)
4	(2,11)	8	(7,16)	12	(-1,6)

Tabel 7.8

De strategie z_3

Einde tweede stap.

Derde stap

a) Uit de analyse van de matrix $\mathcal{P}(z_3)$ blijkt, dat toestand (8,7) vanuit elke toestand bereikt wordt. Bijgevolg bevat $\mathcal{P}(z_3)$ geen disjuncte fuiken en is (8,7) een terugkeertoestand. De cyclus behorende bij deze toestand ziet er als volgt uit:

$$(7.190)\quad \begin{cases} (8,7) \to (0,8) \to (7,9) \to (6,10) \to (4,11) \to (1,12) \to (0,1) \to \\ \to (-2,2) \to (9,3) \to (5,4) \to (0,5) \to (7,6) \to (0,7) \to \\ \to (10,8) \to (1,9) \to (0,10) \to (-2,11) \to (3,12) \to (2,1) \to \\ \to (0,2) \to (-3,3) \to (11,4) \to (6,5) \to (0,6) \to (8,7). \end{cases}$$

Deze cyclus bestaat uit 24 maanden. Uit (7.174) en (7.179) volgt

$$(7.191)\quad \begin{cases} y(z_3) = \frac{1}{24}\{-8-100-7-6-4-1-0-96-9-5-85-7 + \\ \qquad -110-10-1-0-66-3-2-0-119-11-6-95\} = -31\frac{7}{24}\,. \end{cases}$$

Door $v_{(8,7)}(z_3) = 0$ te stellen kunnen vervolgens de overige v-waarden successievelijk berekend worden.

b) De maximalisatie van $V_3((i,j),X)$ naar $X\in\mathfrak{X}((i,j))$ leidt tot

(7.192) $$z_4 = z_3 .$$

Einde derde stap.

Strategie z_3 is optimaal, omdat $z_3 = z_4$.

De in deze paragraaf beschouwde ∞-stapsbeslissingsproblemen kunnen als lineaire programmeringsproblemen geformuleerd worden.

Opmerking 7.5 Formulering als lineair programmeringsprobleem.

Het lineaire programmeringsprobleem, dat wij zullen geven, is niet het meest algemene, doch in de praktijk kan er vrijwel altijd mee worden volstaan. Wij gaan uit van de volgende:

Veronderstelling: De klasse $\mathcal{Z}$ bevat een optimale strategie z_0 waarvan de gemiddelde opbrengst per tijdseenheid niet afhangt van de begintoestand.

Als y_0 de gemiddelde opbrengst per tijdseenheid is van strategie z_0, dan geldt voor elke optimale strategie $z' \in \mathcal{Z}$

(7.193) $$y_i(z') = y_0 \qquad \text{voor } 1 \leq i \leq N.$$

Omdat HOWARD's iteratiemethode naar een optimale strategie convergeert, bestaan een optimale strategie z^* en getallen $v_1^*,\ldots,v_N^*$, zodanig dat

(7.194) $$y_0 + v_i^* - \sum_{j=1}^{N} p_{ij}[z^*]\, v_j^* = h(i,z^*(i)) \qquad \text{voor } 1 \leq i \leq N,$$

en

(7.195) $$v_i^* \geq h(i,X) - y_0 + \sum_{j=1}^{N} p_{ij}(X)\, v_j^* \qquad \text{voor } X \in \mathcal{X}(i);\ 1 \leq i \leq N.$$

Beschouw het volgende stelsel ongelijkheden in de vrije variabelen $x,u_1,\ldots,u_N$:

(7.196) $$x + u_i - \sum_{j=1}^{N} p_{ij}(X)\, u_j \geq h(i,X) \qquad \text{voor } X \in \mathcal{X}(i);\ 1 \leq i \leq N.$$

Dit stelsel bezit een oplossing, omdat $(y_0, v_1^*, \dots, v_N^*)$ eraan voldoet.

Stelling 7.10

Elke optimale basisoplossing $(\bar{x}, \bar{U})$ van het lineaire programmeringsprobleem

(7.197) "minimaliseer x onder de bijvoorwaarden (7.196); $x_1, u_1, \dots, u_N$ vrij"

bezit de volgende eigenschappen:

a) (7.198) $$\bar{x} = y_0 \,,$$

b) als strategie $z' \in \mathcal{Z}$ optimaal is en toestand t is onder z' een terugkeertoestand, dan geldt

(7.199) $$\bar{x} + \bar{u}_t - \sum_{j=1}^{N} p_{tj}[z'] \, \bar{u}_j = h(t, z'(t)).$$

Bewijs:

a) Laat (x,U) een willekeurige oplossing van (7.196) zijn. Uit (7.194) en (7.196) volgt dat voor $1 \leq i \leq N$ geldt

(7.200) $$x - y_0 + u_i - v_i^* - \sum_{j=1}^{N} p_{ij}[z^*](u_j - v_j^*) \geq 0.$$

Door beide leden van (7.200) met bijv. de invariante kans $q_{1i}[z^*]$ te vermenigvuldigen en daarna over i te sommeren, vinden wij met behulp van (7.18)

(7.201) $$x - y_0 + \sum_{i=1}^{N} q_{1i}[z^*](u_i - v_i^*) - \sum_{j=1}^{N} q_{1j}[z^*](u_j - v_j^*) \geq 0.$$

Bijgevolg geldt voor elke oplossing (x,U) van (7.196) dat $x \geq y_0$ is. De relatie (7.198) volgt nu uit het feit dat $(y_0, v_1^*, \dots, v_N^*)$ voldoet aan (7.196).

b) Omdat $(\bar{x},\bar{U})$ voldoet aan (7.196), geldt voor $1 \leqq i \leqq N$

$$\bar{x} + \bar{u}_i - \sum_{j=1}^{N} p_{ij}[z'] \, \bar{u}_j \geqq h(i,z'(i)). \tag{7.202}$$

Onder strategie z' is t een terugkeertoestand, dus $q_{tt}[z'] > 0$ [vgl. (7.26)]. Laten wij eens aannemen dat in (7.202) voor i = t het ongelijkheidsteken geldt. Vermenigvuldigen wij beide leden van (7.202) dan met $q_{ti}[z']$ en sommeren wij daarna over i, dan vinden wij met behulp van (7.18) dat geldt

$$\bar{x} > \sum_{i=1}^{N} q_{ti}[z'] \, h(i,z'(i)). \tag{7.203}$$

Het rechterlid van (7.203) is per definitie gelijk aan $y_t(z')$. Strategie z' is optimaal en dus is $y_t(z') = y_0 = \bar{x}$. De veronderstelling dat in (7.202) voor i = t het ongelijkheidsteken geldt, is dus onjuist. Bijgevolg geldt de relatie (7.199).

Stelling 7.11

Als $(\bar{x},\bar{U})$ een optimale basisoplossing van (7.197) is, zodanig dat toegelaten beslissingen $X_1,\ldots,X_N$ bestaan, waarvoor geldt

$$\bar{x} + \bar{u}_i - \sum_{j=1}^{N} p_{ij}(X_i) \, \bar{u}_j = h(i,X_i) \qquad \text{voor } 1 \leqq i \leqq N, \tag{7.204}$$

dan is strategie $z' = (X_1,\ldots,X_N)$ optimaal.

Bewijs:

Uit (7.204) volgt dat $(y_1 = \bar{x},\ldots,y_N = \bar{x}, v_1 = \bar{u}_1,\ldots,v_N = \bar{u}_N)$ een oplossing is van het met strategie z' corresponderende stelsel vergelijkingen (7.88). De relatie (7.89) heeft tot gevolg dat

$$y_i(z') = \bar{x} \qquad \text{voor } 1 \leqq i \leqq N. \tag{7.205}$$

Uit (7.198) volgt dat $\bar{x} = y_0$; dus strategie z' is optimaal.

Bewering b) van stelling 7.10 leert ons dat elke optimale basisoplossing van het lineaire programmeringsprobleem (7.197) aan de voorwaarden van stelling 7.11 voldoet, indien elke toestand onder tenminste één optimale strategie een terugkeertoestand is. Dus als voor een optimale basisoplossing $(\bar{x},\bar{U})$ en een toestand d geldt dat

$$(7.206) \qquad \bar{x} + \bar{u}_d - \sum_{j=1}^{N} p_{dj}(X)\,\bar{u}_j > h(d,X) \qquad \text{voor alle } X \in \mathcal{X}(d),$$

dan is d onder elke optimale strategie een doorgangstoestand.

Indien wij een optimale basisoplossing van (7.197) vinden, zodanig dat niet aan de voorwaarde van stelling 7.11 voldaan is, dan kan uit deze oplossing een optimale strategie bepaald worden m.b.v. de volgende stelling:

Stelling 7.12

Stel dat $(\bar{x},\bar{U})$ een optimale basisoplossing van (7.197) is en neem aan dat voor de toegelaten beslissingen $X_{i_1},\ldots,X_{i_k}$ geldt

$$(7.207) \qquad \bar{x} + \bar{u}_i - \sum_{j=1}^{N} p_{ij}(X_i)\,\bar{u}_j = h(i,X_i) \qquad \text{voor } i = i_1,\ldots,i_k.$$

Als strategie z' in de toestanden $i_1,\ldots,i_k$ resp. de beslissingen $X_{i_1},\ldots,X_{i_k}$ dicteert en in de overige toestanden zodanige beslissingen dicteert dat deze laatstgenoemde toestanden onder z' doorgangstoestanden zijn, dan is strategie z' optimaal.

Het bewijs van deze stelling is eenvoudig en wordt achterwege gelaten.

Het lineaire programmeringsprobleem (7.197) bevat meer bijvoorwaarden dan variabelen; dus uit rekentechnisch oogpunt is het voordeliger om het duale lineaire programmeringsprobleem op te lossen. Het duale probleem luidt als volgt (vgl. deel 6a, blz. 113):

$$(7.208) \qquad \text{"maximaliseer } \sum_{i=1}^{N} \sum_{X \in \mathcal{X}(i)} h(i,X)\, w_{iX}$$

onder de bijvoorwaarden

$$(7.209)\quad \begin{cases} \sum\limits_{X\in\mathfrak{X}(j)} w_{jX} - \sum\limits_{i=1}^{N} \sum\limits_{X\in\mathfrak{X}(i)} p_{ij}(X)\, w_{iX} = 0 & \text{voor } 1 \leqq j \leqq N, \\ \sum\limits_{i=1}^{N} \sum\limits_{X\in\mathfrak{X}(i)} w_{iX} = 1 & \\ \text{alle } w_{iX} \geqq 0.\text{"} & \end{cases}$$

Uit de "complementary slackness" eigenschap (vgl. deel 6a, blz. 117) en de stellingen 7.11 en 7.12 volgt

Stelling 7.13

Stel dat $\bar{W} = (\bar{w}_{iX})$ een optimale basisoplossing is van het lineaire programmeringsprobleem (7.208), dan geldt

a) als

$$(7.210)\qquad \bar{w}_{iX_i} > 0 \qquad \text{voor } 1 \leqq i \leqq N,$$

dan is strategie $z' = (X_1,\ldots,X_N)$ optimaal.

b) als

$$(7.211)\qquad \bar{w}_{iX_i} > 0 \qquad \text{voor } i = i_1,\ldots,i_k$$

en strategie z' is gedefinieerd als in stelling 7.12, dan is deze strategie optimaal.

De lineaire programmeringsproblemen (7.197) en (7.208) kunnen gegeneraliseerd worden als de op blz. 122 gemaakte veronderstelling niet vervuld is. Voorts merken wij op dat als het beslissingsprobleem de op blz. 26 vermelde eigenschappen bezit, een lineair programmeringsprobleem geformuleerd kan worden waarin gebruik wordt gemaakt van de structuur van het beslissingsprobleem. Wij gaan op deze kwesties niet nader in.

Tenslotte geven wij een toepassing van stelling 7.13 aan de hand van het volgende voorbeeld.

Voorbeeld 7.6 Het produktieprobleem opgelost m.b.v. lineaire programmering.

Beschouw het in voorbeeld 7.4 gestelde probleem. Voor dit produktieprobleem bevat het lineaire programmeringsprobleem (7.208) 14 variabelen en 6 bijvoorwaarden. Met behulp van de rekenautomaat hebben wij een optimale basisoplossing berekend. Het maximum van de kriteriumfunktie is -10 en de positieve componenten van de gevonden optimale basisoplossing zijn

$$w_{32} = \frac{3}{4}, \quad w_{41} = \frac{1}{4}. \tag{7.212}$$

Kiezen wij in toestanden 3 en 4 resp. de beslissingen 2 en 1 en in de overige toestanden 1, 2 en 5 resp. de beslissingen 2, 2 en 1, dan zijn de toestanden 1, 2 en 5 onder strategie

$$z' = (2,2,2,1,1) \tag{7.213}$$

doorgangstoestanden. Dus strategie z' is optimaal [zie stelling 7.13]. Wij kunnen op deze wijze uit (7.212) twaalf strategieën afleiden, waarvoor de gemiddelde opbrengst per tijdseenheid maximaal is.

Opmerking 7.6 Een successieve approximatiemethode als alternatief voor de iteratiemethode van HOWARD.

In HOWARD's iteratiemethode wordt in elke stap een stelsel lineaire vergelijkingen opgelost. In het algemeen kan gesteld worden dat in elke iteratiestap de bepaling van de oplossing van het stelsel vergelijkingen verreweg de meeste tijd kost. Het aantal vergelijkingen is van dezelfde orde als het aantal toestanden. Dit is de reden dat beslissingsproblemen met een groot aantal toestanden veelal niet met HOWARD's iteratiemethode aangepakt kunnen worden, omdat de rekenautomaat grenzen stelt aan de omvang van het op te lossen stelsel lineaire vergelijkingen. Vandaar het belang van een oplossingsmethode waarin geen stelsels lineaire vergelijkingen opgelost behoeven te worden. Wij zullen een dergelijke oplossingsmethode geven en deze "successieve approximatiemethode" leidt in vele praktische beslissingsproblemen tot een optimale strategie. Alvorens wij deze oplossingsmethode formuleren, bewijzen wij eerst de volgende stelling:

Stelling 7.14

Stel dat getallen $y^*, v_1^*, \dots, v_N^*$ bestaan zodat

$$(7.214) \qquad y^* + v_i^* = \max_{X \in \mathcal{X}(i)} \{h(i,X) + \sum_{j=1}^{N} p_{ij}(X)\, v_j^*\} \qquad \text{voor } i = 1,\dots,N.$$

Laat strategie $z^* \in \mathcal{Z}$ zodanig zijn dat voor elke $i = 1,\dots,N$ de beslissing $z^*(i)$ het rechterlid van (7.214) maximaliseert. Dan is strategie z^* optimaal en geldt $y_i(z^*) = y^*$ voor $i = 1,\dots,N$.

Bewijs:

Laat z een willekeurige strategie uit $\mathcal{Z}$ zijn, dan geldt

$$(7.215) \qquad y^* + v_i^* \geqq h(i,z(i)) + \sum_{j=1}^{N} p_{ij}[z]\, v_j^* \qquad \text{voor } i = 1,\dots,N.$$

waarbij het gelijkheidsteken geldt voor alle i als $z = z^*$. Vermenigvuldigen wij beide leden van (7.215) met de invariante kans $q_{ki}[z]$ en sommeren wij over i, dan vinden wij met behulp van (7.16) en (7.18) dat

$$(7.216) \qquad y^* \geqq \sum_{i=1}^{N} q_{ki}[z]\, h(i,z(i)) \qquad \text{voor } k = 1,\dots,N,$$

waarbij het gelijkheidsteken geldt voor alle k als $z = z^*$. Uit (7.216) en (7.37) volgt nu dat $y^* \geqq y_k(z)$ voor $k = 1,\dots,N$ en alle $z \in \mathcal{Z}$ waarbij $y^* = y_k(z^*)$ voor alle k.

Laat $f_0(i)$, $i = 1,\dots,N$, een willekeurige funktie zijn. Wij definiëren de funkties $f_1(i)$, $f_2(i),\dots$ door

$$(7.217) \qquad f_n(i) = \max_{X \in \mathcal{X}(i)} \{h(i,X) + \sum_{j=1}^{N} p_{ij}(X) f_{n-1}(j)\} \qquad \text{voor } i = 1,\dots,N.$$

Laat $z_n \in \mathfrak{Z}$ een strategie zijn zodat voor alle i de beslissing $z_n(i)$ het rechterlid van (7.217) maximalissert, $n \geq 1$. Wij definiëren voor alle $n \geq 1$

$$(7.218) \qquad L_n = \min_{1\leq i\leq N} \{f_n(i) - f_{n-1}(i)\} \text{ en } U_n = \max_{1\leq i\leq N} \{f_n(i) - f_{n-1}(i)\}.$$

Definiëer

$$(7.219) \qquad y_i^* = \max_{z\in\mathfrak{Z}} y_i(z) \qquad \text{voor } i = 1,\dots,N.$$

Stelling 7.15

a) Voor alle $n \geq 1$ geldt $L_n \leq y_i(z_n) \leq y_i^* \leq U_n$ voor $i = 1,\dots,N$, waarbij L_n resp. U_n niet-dalend resp. niet-stijgend is in $n \geq 1$.

Als $y_i^* = y^*$ voor $i = 1,\dots,N$ en als voor elke optimale strategie $z \in \mathfrak{Z}$ geldt dat de matrix $\mathcal{P}(z)$ geen kringfuiken bevat, dan geldt

b) $\lim_{n\to\infty} \{f_n(i) - ny^*\}$ bestaat en is eindig voor $i = 1,\dots,N$.

c) L_n en U_n convergeren naar y^* voor $n\to\infty$.

d) Er bestaat een geheel getal n_0 zodat voor alle $n \geq n_0$ geldt $y_i(z_n) = y^*$ voor alle i, m.a.w. voor alle $n \geq n_0$ is de strategie z_n optimaal.

Bewijs:

a) Kies $n \geq 1$ vast. Uit (7.217) volgt dat voor elke strategie $z \in \mathfrak{Z}$ geldt

$$(7.220) \qquad f_n(i) \geq h(i,z(i)) + \sum_{j=1}^{N} p_{ij}[z]\, f_{n-1}(j) \qquad \text{voor } i = 1,\dots,N,$$

waarbij het gelijkheidsteken geldt voor alle i als $z = z_n$. Kies nu een willekeurige strategie $z \in \mathfrak{Z}$. Uit $f_n(i) - f_{n-1}(i) \leq U_n$ voor alle i en uit (7.220) volgt dat voor $i = 1,\dots,N$,

$$(7.221) \qquad U_n + f_{n-1}(i) \geq h(i,z(i)) + \sum_{j=1}^{N} p_{ij}[z]\, f_{n-1}(j)$$

Vermenigvuldigen wij beide leden van (7.221) met $q_{ki}[z]$ en sommeren wij over i, dan vinden wij met behulp van (7.16) en (7.18) dat

$$U_n \geqq \sum_{i=1}^{N} q_{ki}[z]\ h(i,z(i)) \qquad \text{voor } k = 1,\dots,N. \tag{7.222}$$

Uit (7.222) en (7.37) volgt nu dat $U_n \geqq y_i(z)$ voor $i = 1,\dots,N$ en alle $z \in \mathfrak{Z}$, hetgeen impliceert dat $U_n \geqq y_i^*$ voor alle i. Voor $z = z_n$ geldt in (7.220) het gelijkheidsteken voor alle i. Hieruit en uit de ongelijkheid $f_n(i) - f_{n-1}(i) \geqq L_n$ voor alle i volgt dat voor $i = 1,\dots,N$,

$$L_n + f_{n-1}(i) \leqq h(i,z_n(i)) + \sum_{j=1}^{N} p_{ij}[z_n]\ f_{n-1}(j). \tag{7.223}$$

Op dezelfde wijze als hierboven kunnen wij uit (7.223) afleiden dat voor alle i geldt $L_n \leqq y_i(z_n)$. Hiermee is het eerste gedeelte van bewering (a) bewezen. Uit (7.217) volgt na aftrekken dat voor $i = 1,\dots,N$,

$$\begin{aligned} f_{n+1}(i) - f_n(i) &\geqq \sum_{j=1}^{N} p_{ij}[z_n]\ f_n(j) - \sum_{j=1}^{N} p_{ij}[z_n]\ f_{n-1}(j) \geqq \\ &\geqq \sum_{j=1}^{N} p_{ij}[z_n]\ \{f_n(j) - f_{n-1}(j)\} \geqq L_n \end{aligned} \tag{7.224}$$

waaruit volgt dat $L_{n+1} \geqq L_n$. Op dezelfde wijze bewijst men dat $U_{n+1} \leqq U_n$.

b) Een bewijs van deze bewering kan gevonden worden in E. LANERY, Étude asymptotique des systèmes Markoviens à commande, Revue Francaise d'Informatique et de Recherche Operationelle, <u>1</u> (1967), 3-56.

c) Deze bewering volgt direkt uit bewering (b).

d) De klasse $\mathfrak{Z}$ van strategieën is eindig. Derhalve bestaat een geheel getal n_0 met de volgende eigenschap: als voor een strategie z geldt $z_m = z$ voor een m met $m \geqq n_0$ dan is $z_n = z$ voor oneindig veel waarden van n. Laat z^* een strategie zijn zodat $z_n = z^*$ voor een n met $n \geqq n_0$. Kies nu een deelrij $\{n_k, k \geqq 1\}$ zodat $z_{n_k} = z^*$ voor alle k. In (7.220) geldt het gelijkheidsteken voor alle i als $z = z_n$. Derhalve geldt voor $i = 1,\dots,N$ en alle $k \geqq 1$ dat

$$(7.225)\qquad f_{n_k}(i) - n_k y^* = h(i,z^*(i)) + \sum_{j=1}^{N} p_{ij}[z^*]\,\{f_{n_k-1}(j) - (n_k-1)y^* - y^*\}.$$

Stellen wij $v(i) = \lim_{n\to\infty} \{f_n(i) - ny^*\}$, dan volgt uit (7.225) dat

$$(7.226)\qquad v(i) = h(i,z^*(i)) + \sum_{j=1}^{N} p_{ij}\,[z^*]\,v(j) - y^* \qquad \text{voor } i = 1,\dots,N.$$

Vermenigvuldigen wij beide leden van (7.226) met $q_{ki}[z^*]$ en sommeren wij over i, dan vinden wij met behulp van (7.16) en (7.18) dat

$$(7.227)\qquad 0 = \sum_{i=1}^{N} q_{ki}[z^*]\,h(i,z^*(i)) - y^* \qquad \text{voor } k = 1,\dots,N.$$

Uit (7.37) en (7.227) volgt dat $y_k(z^*) = y^*$ voor $k = 1,\dots,N$.

Uit bovenstaande stelling volgt dat wij in elke iteratiestap van de successieve approximatiemethode (7.217) kunnen aangeven hoe goed de gevonden strategie z_n is ten opzichte van een optimale strategie.

Een nadeel van de bepaling van de rij $\{z_n\}$ door middel van (7.217) is dat $f_n(i)$ onbegrensd groot wordt als n toeneemt. Dit bezwaar kan onder zekere voorwaarden ondervangen worden. Daartoe kiezen wij een vaste toestand, stel toestand i_0, en definiëren wij voor $n \geqq 0$, de funktie $v_n(i)$ door

$$(7.228)\qquad v_n(i) = f_n(i) - f_n(i_0) \qquad \text{voor } i = 1,\dots,N.$$

Uit (7.217) en (7.228) volgt direkt dat voor $i = 1,\dots,N$ en $n \geq 1$ geldt

$$(7.229)\qquad v_n(i) = \max_{X\in\mathcal{X}(i)} \{h(i,X) + \sum_{j=1}^{N} p_{ij}(X)\,v_{n-1}(j)\} - y_n,$$

waarbij

$$(7.230)\qquad y_n = \max_{X\in\mathcal{X}(i_0)} \{h(i_0,X) + \sum_{j=1}^{N} p_{i_0 j}(X)\,v_{n-1}(j)\}.$$

Men gaat direkt na dat de rechterleden van (7.217) en (7.229) voor dezelfde beslissingen maximaal zijn, d.w.z. de rij van strategieën $\{z_n\}$ wordt ook door (7.229) voortgebracht. Uit (7.218), (7.228) en (7.230) volgt dat

$$(7.231) \qquad L_n = y_n + \min_{1 \leq i \leq N} \{v_n(i) - v_{n-1}(i)\}, \; U_n = y_n + \max_{1 \leq i \leq N} \{v_n(i) - v_{n-1}(i)\}.$$

Stel nu eens dat voor elke i de rij $\{v_n(i), n \geq 1\}$ convergeert naar eindige limiet v(i), zoals het geval is onder de voorwaarden gegeven in stelling 7.15. Dan volgt uit (7.230) dat y_n convergeert naar een eindig getal y. Vervolgens impliceert (7.229) dat

$$(7.232) \qquad y + v(i) = \max_{X \in \mathfrak{X}(i)} \{h(i,X) + \sum_{j=1}^{N} p_{ij}(X)\, v_j\} \qquad \text{voor } i = 1,\ldots,N.$$

Op grond van stelling 7.14 kunnen wij nu stellen dat de convergentie van de "relatieve opbrengstfunkties" $v_n(i)$ een optimale strategie bepaalt.

Samenvattend kunnen wij de volgende successieve approximatiemethode formuleren. Kies een vaste toestand i_0 en kies een willekeurige funktie $v_0(i)$ met $v_0(i_0) = 0$. De n^{de} stap van de successieve approximatiemethode ziet er als volgt uit:

<u>De n^{de} stap uit de successieve approximatiemethode</u>.

Uitgaande van de in de vorige stap verkregen getallen $v_{n-1}(i)$, $i = 1,\ldots,N$, berekenen wij eerst y_n door middel van (7.230), waarna de getallen $v_n(i)$ worden berekend door middel van (7.229). Hierbij vinden wij een strategie z_n door voor elke i de beslissing $z_n(i)$ gelijk te kiezen aan een beslissing waarvoor het rechterlid van (7.229) maximaal is. Een onder- en bovengrens zowel voor de gemiddelde opbrengst behorende bij strategie z_n als voor de maximale gemiddelde opbrengst kunnen worden berekend uit (7.231).

<u>Voorbeeld 7.7</u> <u>Het brouwerijprobleem opgelost met de methode van successieve approximatie</u>.

In dit voorbeeld passen wij bovenstaande methode toe op het probleem uit voorbeeld 7.2, zie blz. 79. Voor de keuze $i_0 = 1$ geven wij in tabel 7.9 de waarden van y_n en $v_n(i)$ voor de eerste 19 iteratiestappen.

	y_n	$v_n(1)$	$v_n(2)$	$v_n(3)$	$z_n(1)$	$z_n(2)$	$z_n(3)$
n = 1	-12,000	0	3,000	6,000	0	1	0
n = 2	-10,500	0	4,200	6,750	0	1	2
n = 3	- 9,900	0	4,350	7,035	0	1	2
n = 4	- 9,825	0	4,407	7,205	0	1	2
n = 5	- 9,797	0	4,441	7,312	0	1	2
n = 6	- 9,780	0	4,462	7,380	0	1	2
n = 7	- 9,769	0	4,476	7,424	0	1	2
n = 8	- 9,762	0	4,485	7,451	0	1	2
n = 9	- 9,758	0	4,490	7,469	0	1	2
n = 10	- 9,755	0	4,494	7,480	0	1	2
n = 11	- 9,753	0	4,496	7,487	0	1	2
n = 12	- 9,752	0	4,498	7,492	0	1	2
n = 13	- 9,751	0	4,498	7,495	0	1	2
n = 14	- 9,751	0	4,499	7,497	0	1	2
n = 15	- 9,751	0	4,499	7,498	0	1	2
n = 16	- 9,750	0	4,500	7,499	0	1	2
n = 17	- 9,750	0	4,500	7,499	0	1	2
n = 18	- 9,750	0	4,500	7,500	0	1	2
n = 19	- 9,750	0	4,500	7,500	0	1	2

Tabel 7.9

De numerieke waarden van y_n en $v_n(i)$ voor $n \leq 19$

8. De invariante kansverdeling en de gemiddelde opbrengst per tijdseenheid.

Beschouw de situatie waarin een beslisser gedurende een tijdsinterval van onbegrensde lengte een systeem beheert en waarin hij alleen op van tevoren gegeven equidistante tijdstippen een beslissing nemen kan. De lengte van het interval tussen twee opeenvolgende beslissingstijdstippen wordt als tijdseenheid genomen.

De toestandsruimte $\mathcal{S}$ wordt verondersteld de verzameling van de reële getallen of een deelverzameling daarvan te zijn. Stel dat de beslisser een vaste strategie z toepast. Laat een opbrengst [of kosten] h(S,z(S)) worden verkregen als het systeem op een beslissingstijdstip toestand S aanneemt. Wij nemen aan dat het bij strategie z behorende beslissingsproces beschreven kan worden door een Markov-keten. De karakteristieke eigenschap van dit Markov-proces is dat de kansverdeling van de toestand op een equidistant tijdstip uitsluitend bepaald wordt door de op het vorige tijdstip waargenomen toestand, dus de kennis van het "verleden" is irrelevant en alleen het "heden" is bepalend voor de "toekomst". Voorts veronderstellen wij dat de overgangswaarschijnlijkheden van het Markov-proces dezelfde zijn voor elk tijdstip.

Als wij de kansverdeling kunnen bepalen die het gedrag van het systeem op een heel ver weg gelegen tijdstip beschrijft, dan kunnen wij met deze kansverdeling (die uiteraard niet afhangt van de structuur van de opbrengsten!) direct de gemiddelde opbrengst per tijdseenheid berekenen. Voor het geval van een eindige toestandsruimte hebben wij dit reeds laten zien in paragraaf 7 (zie blz. 78).

Als de toestandsruimte aftelbaar is, kunnen wij op analoge wijze te werk gaan. De mogelijke toestanden van het systeem nummeren wij als 1,2,... en de één-staps- en k-stapsovergangskans van toestand i naar toestand j geven wij aan met $p_{ij}[z]$ resp. $p_{ij}^{(k)}[z]$. Voor een aftelbare Markov-keten behoeft een invariante kansverdeling niet te bestaan.*) Wij nemen echter aan

*) In de boeken, die genoemd zijn in de voetnoot op blz. 72 kunnen voorwaarden gevonden worden waaronder een invariante kansverdeling bestaat.

dat voor alle i en j

$$(8.1) \qquad q_j[z] \overset{\mathrm{def}}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)}[z]$$

bestaat en dat deze limiet niet afhangt van de begintoestand i. Voorts veronderstellen wij dat $(q_1[z], q_2[z], \ldots)$ een kansverdeling vormt. Zonder bewijs vermelden wij dat de $q_j[z]$ de unieke oplossing zijn van het stelsel

$$\begin{cases} (8.2) \quad q_j[z] = \sum_{h=1}^{\infty} p_{hj}[z]\, q_h[z] \qquad \text{voor } j = 1,2,\ldots, \\ (8.3) \quad \sum_{j=1}^{\infty} q_j[z] = 1 \,. \end{cases}$$

De kansverdeling $(q_1[z], q_2[z], \ldots)$ wordt een <u>invariante</u> kansverdeling genoemd. Onder algemene netheidsvoorwaarden, waaraan in de praktijk veelal voldaan is, kan worden aangetoond dat ongeacht de begintoestand de gemiddelde opbrengst per tijdseenheid met kans 1 gelijk is aan

$$(8.4) \qquad y(z) = \sum_{j=1}^{\infty} q_j[z]\, h(j, z(j)).$$

Beschouw vervolgens het geval dat de mogelijke toestanden een <u>continuum</u> vormen. Laat $K(S|y,z)$ de kans zijn dat het systeem op het volgende beslissingstijdstip een toestand kleiner dan of gelijk aan S aanneemt, als het systeem nu in toestand y is. De funktie $K(S|y,z)$ is voor vaste y een verdelingsfunktie in S. De verdelingsfunktie van de toestand m tijdstippen later, als de toestand nu y is, geven wij aan met $K^{(m)}(S|y,z)$. Wij nemen aan dat voor alle $y, S \in \mathcal{S}$

$$(8.5) \qquad Q(S,z) \overset{\mathrm{def}}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} K^{(m)}(S|y,z)$$

bestaat en dat deze limiet niet afhangt van de begintoestand y. Voorts veronderstellen wij niet alleen dat $Q(S,z)$ een verdelingsfunktie is, maar ook dat $Q(S,z)$ een kansdichtheid $q(S,z)$ bezit. Zonder bewijs vermelden wij dat $Q(S,z)$ de enige verdelingsfunktie is, die voldoet aan

$$(8.6) \qquad Q(S,z) = \int_{y \in \mathcal{S}} K(S|y,z)\, q(y,z)\, dy \qquad \text{voor } S \in \mathcal{S}.$$

Als $K(S|y,z)$ een kansdichtheid $k(S|y,z)$ heeft, dan geldt tevens dat de kansdichtheid $q(S,z)$ voldoet aan

$$(8.7) \qquad q(S,z) = \int_{y \in \mathcal{S}} k(S|y,z)\, q(y,z)\, dy \qquad \text{voor } S \in \mathcal{S}.$$

De funkties $Q(S,z)$ en $q(S,z)$ worden de invariante kansverdeling resp. de invariante kansdichtheid genoemd. Ruwweg kan gesteld worden dat, ongeacht de toestand waarin het systeem zich nu bevindt, op een oneindig ver weg gelegen beslissingstijdstip het systeem met kans $Q(S,z)$ een toestand kleiner dan of gelijk aan S aanneemt.

In de praktijk zal veelal aan de netheidsvoorwaarden (zie bijv. het op blz. 72 genoemde boek van DOOB) voldaan zijn, die niet alleen garanderen dat een unieke invariante kansverdeling bestaat, maar die tevens tot gevolg hebben dat, ongeacht de begintoestand, de gemiddelde opbrengst per tijdseenheid met kans 1 gelijk is aan

$$(8.8) \qquad y(z) = \int_{S \in \mathcal{S}} h(S,z(S))\, q(S,z)\, dS .$$

Zonder er verder op in te gaan, merken wij op dat het geval waarin $Q(S,z)$ de verdelingsfunktie is van een gemengde kansverdeling, op analoge wijze behandeld kan worden als het discrete en continue geval. Verder merken wij op dat $y(z)$ ook berekend kan worden door een gegeneraliseerde versie van het stelsel vergelijkingen uit HOWARD's iteratiemethode te beschouwen. Op deze aanpak, die tot een iteratieprocedure leidt, komen wij terug in deel 7c van deze syllabusserie.

De berekening van de gemiddelde opbrengst per tijdseenheid door eerst de invariante kansverdeling van de toestand te bepalen, zullen wij illustreren aan de hand van een aantal voorraadproblemen. In deze voorraadproblemen beperken wij ons tot strategieën die van één of twee parameters afhangen. Door de kriteriumfunktie $y(z)$ naar de strategieparameters te optimaliseren, vindt men de beste strategie uit de beschouwde klasse van strategieën.

In de te behandelen voorbeelden zullen wij steeds stilzwijgend aannemen dat de toestand een unieke invariante kansverdeling bezit en dat de gemiddelde opbrengst per tijdseenheid met kans 1 gelijk is aan het rechterlid van (8.8).

Voorbeeld 8.1 Een (s_-,Q)-strategie voor het noodinkoopmodel zonder levertijd.

Een voorraadbeheerder kan zijn voorraad alleen aanvullen op van tevoren gegeven equidistante tijdstippen en wel aan het begin van de perioden 1,2,... . De levertijd van een bestelling is verwaarloosbaar. De inkoopkosten bedragen c per eenheid en per bestelling zijn de vaste kosten K. De voorraadkosten in een periode zijn c_1 maal de aanwezige voorraad aan het eind van die periode. De behoeften in de opeenvolgende perioden zijn onafhankelijk van elkaar en hebben eenzelfde exponentiële verdeling met parameter λ. Als de vraag de voorraad overschrijdt, wordt het tekort direct teniet gedaan door een noodinkoop. De noodinkoopkosten zijn c_2 per eenheid. Wij nemen aan dat $c_2 > c$ is (als $c_2 \leq c$ zou zijn, dan was het optimaal om alleen noodinkopen te verrichten).

De voorraadbeheerder wenst zich te beperken tot een (s_-,Q)-strategie met $Q > s_-$; d.w.z. als aan het begin van een periode de voorraad minder dan s_- is, wordt een hoeveelheid Q besteld, maar als de voorraad s_- of meer is, wordt geen bestelling geplaatst.

Gevraagd wordt de beste (s_-,Q)-strategie te bepalen.

Oplossing

De toestand S van het systeem aan het begin van een periode wordt gedefinieerd als de grootte van de voorraad minus de grootte van de in de vorige periode eventueel verrichte noodinkopen; de toestand wordt steeds vlak voor een eventuele aanvulling gemeten. Dus als $S \geq 0$ is dan is de voorraad gelijk aan S en zijn geen noodinkopen verricht in de vorige periode; als $S < 0$ is, dan is de voorraad nul en is een hoeveelheid $-S$ aan noodinkopen verricht. Bij het toepassen van strategie $z = (s_-, Q)$ kan de toestandsgrootheid alleen waarden kleiner dan s_-+Q aannemen.

De aan een beslissingstijdstip toe te kennen kosten definiëren wij als de som van de bestelkosten behorende bij de te nemen beslissing en de voorraad- en noodinkoopkosten uit de vorige periode, zodat

$$h(S,z(S)) = \begin{cases} c_1 S & \text{voor } S \geqq s_- , \\ c_1 S + cQ + K & \text{voor } 0 \leqq S < s_- , \\ -c_2 S + cQ + K & \text{voor } S < 0 . \end{cases} \tag{8.9}$$

Aangezien $e^{-\lambda u}$ de kans is dat de behoefte in een periode groter dan of gelijk aan u is, wordt de verdelingsfunktie $K(S|y,z)$ gegeven door

$$K(S|y,z) = \begin{cases} e^{-\lambda(Q-S)} & \text{voor } y < 0 \; ; \; S < Q, \\ e^{-\lambda(y+Q-S)} & \text{voor } 0 \leqq y < s_- ; \; S < y+Q, \\ e^{-\lambda(y-S)} & \text{voor } y \geqq s_- ; \; S < y, \\ 1 & \text{anders.} \end{cases} \tag{8.10}$$

Deze verdelingsfunktie heeft als kansdichtheid

$$k(S|y,z) = \begin{cases} \lambda e^{-\lambda(Q-S)} & \text{voor } y < 0 \; ; \; S < Q, \\ \lambda e^{-\lambda(y+Q-S)} & \text{voor } 0 \leqq y < s_- ; \; S < y+Q, \\ \lambda e^{-\lambda(y-S)} & \text{voor } y \geqq s_- ; \; S < y, \\ 0 & \text{anders.} \end{cases} \tag{8.11}$$

Uit (8.7) volgt dat de invariante kansdichtheid $q(S,z)$ voldoet aan (bedenk dat $Q > s_-$ is):

$$\begin{aligned} q(S,z) = & \int_{s_-}^{s_-+Q} \lambda e^{-\lambda(y-S)} \, q(y,z) \, dy + \int_0^{s_-} \lambda e^{-\lambda(y+Q-S)} \, q(y,z) \, dy + \\ & + \int_{-\infty}^{0} \lambda e^{-\lambda(Q-S)} \, q(y,z) \, dy \qquad \text{voor } S < s_- , \end{aligned} \tag{8.12a}$$

$$q(S,z) = \int_{S}^{s_-+Q} \lambda e^{-\lambda(y-S)}\, q(y,z)\, dy + \int_{0}^{s_-} \lambda e^{-\lambda(y+Q-S)}\, q(y,z)\, dy +$$

(8.12b)

$$+ \int_{-\infty}^{0} \lambda e^{-\lambda(Q-S)}\, q(y,z)\, dy \qquad \text{voor } s_- \leqq S < Q,$$

$$q(S,z) = \int_{S}^{s_-+Q} \lambda e^{-\lambda(y-S)}\, q(y,z)\, dy + \int_{S-Q}^{s_-} \lambda e^{-\lambda(y+Q-S)}\, q(y,z)\, dy$$

(8.12c)

$$\text{voor } Q \leqq S < s_-+Q.$$

Om deze integraalvergelijkingen op te lossen, gaan wij als volgt te werk. Wij differentiëren beide leden van (8.12a) t/m (8.12c) naar S, daarbij aannemend dat de betreffende afgeleiden bestaan. Na enig rekenwerk vinden wij

(8.13a) $\quad q'(S,z) = \lambda q(S,z) \qquad$ voor $S < s_-$,

(8.13b) $\quad q'(S,z) = 0 \qquad$ voor $s_- < S < Q$,

(8.13c) $\quad q'(S,z) = -\lambda q(S-Q,z) \qquad$ voor $Q < S < s_-+Q$.

Uit (8.13a) volgt dat voor een nog nader te bepalen constante d geldt

(8.14) $\quad q(S,z) = de^{\lambda S} \qquad$ voor $S < s_-$.

Uit (8.13b) volgt dat $q(S,z)$ constant is voor $s_- < S < Q$. Een onderzoek van (8.12a) en (8.12b) leert ons dat $q(S,z)$ continu is in $S = s_-$. Bijgevolg geldt

(8.15) $\quad q(S,z) = de^{\lambda s_-} \qquad$ voor $s_- \leqq S < Q$.

De relaties (8.13c) en (8.14) impliceren dat $q'(S,z) = -d\lambda e^{\lambda(S-Q)}$ voor $Q < S < s_-+Q$. Uit (8.12c) volgt dat $q(S,z)$ rechtscontinu is in $S = Q$. Bijgevolg geldt voor een nog nader te bepalen constante d_1

(8.16) $\quad q(S,z) = -de^{\lambda(S-Q)} + d_1 \qquad$ voor $Q \leqq S < s_-+Q$.

Uit (8.12b), (8.12c) en (8.15) kunnen wij afleiden dat

(8.17) $$de^{\lambda s_-} = q(Q,z) + \int_{-\infty}^{0} \lambda q(y,z)\, dy \,.$$

Na enig rekenwerk vinden wij met behulp van (8.14), (8.16) en (8.17) dat

(8.18) $$d_1 = d\, e^{\lambda s_-}$$

Nu rest ons alleen nog de constante d te bepalen. De funktie q(S,z) is een kansdichtheid, zodat geldt

(8.19) $$\int_{-\infty}^{s_-+Q} q(S,z)\, dS = 1.$$

De relaties (8.14) t/m (8.16), (8.18) en (8.19) leiden na enig rekenwerk tot

(8.20) $$d = \frac{\lambda}{\lambda Q e^{\lambda s_-} + 1}.$$

Bijgevolg is de invariante kansdichtheid q(S,z) gelijk aan

(8.21) $$q(S,z) = \begin{cases} de^{\lambda s_-} - de^{\lambda(S-Q)} & \text{voor } Q \leqq S < s_-+Q, \\ de^{\lambda s_-} & \text{voor } s_- \leqq S < Q, \\ de^{\lambda S} & \text{voor } S < s_-. \end{cases}$$

waarbij d gegeven wordt door (8.20). Deze kansdichtheid is op nogal formele wijze afgeleid. Achteraf kan men controleren dat de door (8.21) gegeven funktie inderdaad de invariante kansdichtheid is door deze te substitueren in (8.12a) t/m (8.12c).

Als aan het begin van een periode de toestand $S \geqq 0$ is, zijn in de voorafgaande periode geen noodinkopen verricht; als $S < 0$ is, zijn wel noodinkopen verricht. De waarschijnlijkheidsrekening leert ons nu, dat over een oneindig aantal perioden genomen de fractie van het aantal perioden waarin een noodinkoop wordt verricht met kans 1 gelijk is aan

$$(8.22) \qquad P_{tekort} = \int_{-\infty}^{0} q(S,z)\, dS = \frac{1}{\lambda Q e^{\lambda s_-} + 1} .$$

Uit (8.9) en (8.21) volgt na enig rekenwerk voor de kriteriumfunktie y(z)

$$(8.23) \quad y(z) = \frac{1}{\lambda Q e^{\lambda s_-} + 1} \left\{ \frac{c_2}{\lambda} + c_1 Q + e^{\lambda s_-}(cQ - c_1 Q + K + \lambda c_1 Q (\tfrac{1}{2} Q + s_-)) \right\}.$$

Om de beste (s_-,Q)-strategie te bepalen, stellen wij de partiële afgeleiden van y(z) naar s_- en Q gelijk aan nul. Na enig rekenwerk volgt dat de optimale s_- (= s_-^*) en Q (= Q^*) voldoen aan

$$(8.24) \qquad Q = A(s_-) ,$$

en

$$(8.25) \qquad s_- + (e^{\lambda s_-} - \tfrac{1}{2})\, A(s_-) - \frac{K}{\lambda c_1 A(s_-)} + \frac{c - c_2}{\lambda c_1} = 0 ,$$

waarbij

$$(8.26) \qquad A(s_-) = \frac{1 - e^{-\lambda s_-}}{\lambda} + \sqrt{\frac{2K}{\lambda c_1} + \left(\frac{1 - e^{-\lambda s_-}}{\lambda}\right)^2} .$$

Geven wij het linkerlid van vergelijking (8.25) kortweg aan met $N(s_-)$, dan vinden wij na enig rekenwerk dat $N'(s_-) > 0$ is voor alle $s_- \geq 0$. Bijgevolg is $N(s_-)$ monotoon stijgend voor $s_- \geq 0$. Voorts geldt dat $N(0) = -\frac{(c_2 - c)}{\lambda c_1} < 0$ is en $N(s_-) \to \infty$ als $s_- \to \infty$ gaat. Dus vergelijking (8.25) heeft precies één positief nulpunt; s_-^* is gelijk aan dit nulpunt.

Voor het getallenvoorbeeld, waarin

$$(8.27) \qquad \lambda = 0{,}02 \; ; \; c = c_1 = 1 \; ; \; c_2 = 15 \text{ en } K = 30 ,$$

vinden wij met behulp van de rekenautomaat

$$(8.28) \qquad s_-^* = 90{,}613 \; , \; Q^* = 110{,}758 \text{ en } y(z^*) = 207{,}675.$$

<u>Voorbeeld 8.2</u> <u>Een (S^+,S^+)-strategie voor het noodinkoopmodel met een levertijd van één periode.</u>

Beschouw het in voorbeeld 8.1 gestelde probleem. Wij nemen nu aan dat een bestelling, die geplaatst is aan het begin van een periode, wordt afgeleverd aan het begin van de volgende periode. Tevens veronderstellen wij dat de verdelingsfunktie $W(u)$ van de behoefte in een periode de kansdichtheid $w(u)$ bezit (het discrete geval kan op analoge wijze opgelost worden als het continue geval). De kostenstructuur is dezelfde als in voorbeeld 8.1. De voorraadbeheerder wenst zich te beperken tot een (S^+,S^+)-strategie; d.w.z. als aan het begin van een periode de voorraad S kleiner dan S^+ is, wordt een hoeveelheid $S^+ - S$ besteld. Dus aan het begin van elke periode is de voorraad plus de hoeveelheid in bestelling gelijk aan S^+.

Gevraagd wordt de beste (S^+,S^+)-strategie te bepalen.

Oplossing

De toestand van het systeem aan het begin van een periode wordt gedefinieerd als de grootte van de voorraad; de toestand wordt steeds gemeten vlak na de aflevering van de in de vorige periode bestelde hoeveelheid.

Als de toestand aan het begin van een periode gelijk aan S is, zijn de verwachte voorraad- en boetekosten in die periode gelijk aan

$$(8.29) \qquad h(S) = c_1 \int_0^S (S-u)\, w(u)\, du + c_2 \int_S^\infty (u-S)\, w(u)\, du \qquad \text{voor } S \geqq 0.$$

Neem aan dat strategie $z = (S^+,S^+)$ wordt toegepast. De kostenfunktie $h(S,z(S))$ wordt gegeven door

$$(8.30) \qquad h(S,z(S)) = h(S) + c(S^+-S) + K \qquad \text{voor } 0 \leqq S < S^+.$$

Stel dat de toestand nu y is. Dan wordt aan het begin van de volgende periode een bestelling ter grootte van $S^+ - y$ afgeleverd. Als de behoefte in de huidige periode $\underline{b}$ is, dan is de toestand aan het begin van de volgende periode gelijk aan $S^+ - y + \max(0,y-\underline{b})$. De gebeurtenissen $S^+ - y + \max(0,y-\underline{b}) \leqq S$ en $\max(0,y-\underline{b}) \leqq S - S^+ + y$ zijn uiteraard dezelfde. Voor twee willekeurige getallen a en b geldt $\max(a,b) \leqq c$ dan en slechts dan als $a \leqq c$ én $b \leqq c$; dus voor $S \geqq S^+ - y$ treedt de gebeurtenis $S^+ - y + \max(0,y-b) \leqq S$ dan en slechts dan op als $y - b \leqq S - S^+ + y$ is. De kans op de gebeurtenis $\underline{b} \geqq S^+ - S$ is $1 - W(S^+-S)$. Bijgevolg geldt voor elke vaste $y \in (0,S^+)$ dat

$$(8.31) \qquad K(S|y,z) = \begin{cases} 1 - W(S^+-S) & \text{voor } S^+ - y \leqq S \leqq S^+ \\ 0 & \text{voor } 0 \leqq S < S^+ - y. \end{cases}$$

Deze verdelingsfunktie is discontinu in $S = S^+ - y$ en continu in de overige S. Dit is het gevolg van het feit dat, als de toestand nu y is, er een positieve kans is dat $S^+ - y$ de toestand op het volgende beslissingstijdstip is, terwijl de overige op dat tijdstip mogelijke toestanden continu verdeeld zijn. Hoewel de verdelingsfunktie $K(S|y,z)$ geen kansdichtheid bezit, kan niettemin aangetoond worden dat de invariante kansverdeling $Q(S,z)$ een kansdichtheid $q(S,z)$ heeft, hetgeen overigens intuïtief duidelijk zal zijn. De invariante verdelingsfunktie $Q(S,z)$ voldoet aan [zie (8.6)]

$$(8.32) \qquad \begin{cases} Q(S,z) = \displaystyle\int_{S^+-S}^{S^+} \{1 - W(S^+-S)\}\, q(y,z)\, dy = \\ \qquad = \{1 - W(S^+-S)\}\, \{Q(S^+,z) - Q(S^+-S,z)\} = \\ \qquad = \{1 - W(S^+-S)\}\, \{1 - Q(S^+-S,z)\} \quad \text{voor } 0 \leqq S \leqq S^+. \end{cases}$$

Stellen wij in (8.32) de variabele $S^+ - S$ gelijk aan u, dan vinden wij

$$(8.33) \qquad Q(S^+-u,z) = \{1 - W(u)\}\, \{1 - Q(u,z)\} \qquad \text{voor } 0 \leqq u \leqq S^+.$$

Uit (8.32) en (8.33) volgt tenslotte

$$(8.34) \qquad Q(S,z) = \frac{W(S)\, \{1 - W(S^+-S)\}}{1 - \{1 - W(S)\}\, \{1 - W(S^+-S)\}} \qquad \text{voor } 0 \leqq S \leqq S^+.$$

De invariante kansdichtheid $q(S,z)$ is gelijk aan de afgeleide van $Q(S,z)$ naar S.

Als de toestand aan het begin van een periode S is, dan worden in die periode noodinkopen uitsluitend verricht als de vraag groter dan S is; de kans hierop is $1 - W(S)$. Dus de fractie van het aantal perioden waarin een noodinkoop wordt verricht, is met kans 1 gelijk aan

$$(8.35) \qquad P_{tekort} = \int_0^{S^+} \{1 - W(S)\}\ q(S,z)\ dS.$$

De kriteriumfunktie y(z) is gelijk aan

$$(8.36) \qquad y(z) = \int_0^{S^+} h(S,z(S))\ q(S,z)\ dS\ .$$

Als de vraag een exponentiële verdeling met parameter λ heeft, dus als

$$(8.37) \qquad W(u) = 1 - e^{-\lambda u}\ .$$

dan vinden wij na enig rekenwerk dat geldt

$$(8.38) \quad h(S,z(S)) = c_1 S + \frac{c_1 + c_2}{\lambda}\ e^{-\lambda S} - \frac{c_1}{\lambda} + c(S^+ - S) + K \qquad \text{voor } 0 \leqq S \leqq S^+.$$

Verder geldt dan

$$(8.39) \qquad Q(S,z) = \frac{e^{\lambda S} - 1}{e^{\lambda S^+} - 1} \qquad \text{voor } 0 \leqq S \leqq S^+$$

en

$$(8.40) \qquad q(S,z) = \frac{\lambda e^{\lambda S}}{e^{\lambda S^+} - 1} \qquad \text{voor } 0 \leqq S \leqq S^+\ .$$

De relatie (8.36) gaat over in

$$(8.41) \qquad y(z) = \frac{(2c_1 + c_2 - c)\ S^+}{e^{\lambda S^+} - 1} + c_1 S^+ - \frac{(2c_1 - c)}{\lambda} + K\ .$$

Om de beste (S^+,S^+)-strategie te bepalen, stellen wij de afgeleide van y(z) naar S^+ gelijk aan nul. Wij vinden dan dat de optimale S^+ (= S^*) voldoet aan

$$(8.42) \qquad c_1\{e^{2\lambda S^+} - 2\lambda S^+ e^{\lambda S^+} - 1\} + (c_2 - c)\ \{e^{\lambda S^+} - \lambda S^+ e^{\lambda S^+} - 1\} = 0\ .$$

Zonder op de details in te gaan, vermelden wij dat een analyse van het linkerlid van (8.42) en zijn eerste afgeleide ons leert, dat vergelijking (8.42) precies één positieve wortel heeft. S^* is gelijk aan deze wortel.

Voor het getallenvoorbeeld, waarin

$$\lambda = 0{,}02 \ ; \ c = c_1 = 1 \ ; \ c_2 = 15 \text{ en } K = 30 \ , \tag{8.43}$$

vinden wij met behulp van de rekenautomaat

$$S^* = 193{,}947 \quad \text{en} \quad y(z^*) = 239{,}452. \tag{8.44}$$

In de voorbeelden 8.3 en 8.4 behandelen wij de in voorbeeld 6.2 (zie blz. 40 e.v.) gestelde voorraadproblemen. In voorbeeld 8.3 beschouwen wij alleen het naleveringsmodel. Wij nemen in dit voorbeeld aan dat de voorraadbeheerder zich wenst te beperken tot een (s_-,nQ)-strategie. In voorbeeld 8.4 veronderstellen wij dat de voorraadbeheerder zich beperkt tot een (S^+,s_-)-strategie en wij behandelen alle modellen die in voorbeeld 6.2 aan de orde gesteld zijn.

<u>Voorbeeld 8.3</u> <u>Een (s_-,nQ)-strategie voor het naleveringsmodel.</u>

Eerst beschouwen wij het naleveringsmodel, waarin de levertijd nul is en $W(u)$ (i.e. de verdelingsfunktie van de behoefte in een periode) een kansdichtheid $w(u)$ bezit.

Een (s_-,nQ)-strategie ziet er als volgt uit: als aan het begin van een periode de voorraad groter dan of gelijk aan s_- is, wordt geen bestelling geplaatst; als de voorraad kleiner dan s_- én groter dan of gelijk aan $s_- + Q$ is, wordt een hoeveelheid Q besteld; als de voorraad kleiner dan $s_- + Q$ én groter dan of gelijk aan $s_- + 2Q$ is, wordt een bestelling van de grootte 2Q geplaatst; etc. Dus algemeen gesteld: als de voorraad kleiner dan $s_- + (n-1)Q$ én groter dan of gelijk aan $s_- + nQ$ is, wordt een hoeveelheid nQ besteld.

De voorraad vlak na eventuele aanvulling ligt dus tussen s_- en $s_- + Q$. De toestand van het systeem aan het begin van een periode wordt gedefinieerd als de grootte van de voorraad <u>na</u> een eventuele aanvulling. De toestandsruimte wordt derhalve gegeven door

(8.45) $\mathcal{S} = \{S \mid s_- \leqq S < s_- + Q\}$.

Als de toestand nu S is, zijn de verwachte inkoopkosten in de volgende periode gelijk aan

(8.46) $$\int_{S-s_-}^{\infty} \phi(Q_{S-v})\, w(v)\, dv \qquad \text{voor } s_- \leqq S < s_- + Q \ ,$$

waarin

(8.47) $$Q_{S-v} \overset{\text{def}}{=} nQ \qquad \text{als } s_- - nQ \leqq S-v < s_- - (n-1)Q.$$

De inkoopkosten in de eerste periode hebben geen invloed op de gemiddelde kosten per tijdseenheid en kunnen derhalve buiten beschouwing gelaten worden. Alle overige te maken kosten zullen wij op ondubbelzinnige wijze splitsen en wel als volgt: als aan het begin van een periode de toestand S is, kennen wij aan die periode de kosten h(S,z(S)) toe, waarbij deze kosten gedefinieerd zijn als de som van de verwachte voorraad- en boetekosten in de betreffende en de verwachte bestelkosten in de volgende periode. Bijgevolg geldt

(8.48) $$h(S,z(S)) = h(S) + \int_{S-s_-}^{\infty} \phi(Q_{S-v})\, w(v)\, dv \qquad \text{voor } s_- \leqq S < s_- + Q.$$

Stel dat de toestand nu y is. Aan het begin van de volgende periode wordt een hoeveelheid nQ besteld ($n \geqq 1$) dan en slechts dan als de gebeurtenis $\{y - s_- + (n-1)Q < \underline{b} \leqq y - s_- + nQ\}$ optreedt, waarbij $\underline{b}$ de vraag in de huidige periode is. Als een hoeveelheid nQ besteld wordt, is de voorraad vlak na aanvulling gelijk aan $y - \underline{b} + nQ$; deze voorraad is minstens S als tevens de gebeurtenis $\{\underline{b} \leqq y - S + nQ\}$ optreedt. Voor $S \in \mathcal{S}$ geldt $S \geqq s_-$, dus $y - S + nQ \leqq y - s_- + nQ$. Dus aan het begin van de volgende periode wordt een hoeveelheid nQ besteld én is de voorraad vlak na aanvulling minstens S ($\in \mathcal{S}$) dan en slechts dan als de gebeurtenis $\{y - s_- + (n-1)Q < \underline{b} \leqq$ $\leqq y - S + nQ\}$ optreedt. Door alle elkaar uitsluitende eventualiteiten te beschouwen, die tot de gebeurtenis "de toestand aan het begin van de volgende periode is minstens S" leiden, vinden wij dat voor alle y, S $\in \mathfrak{S}$ geldt

$$1 - K(S|y,z) = P\{\underline{b} \leqq y - S\} + \sum_{n=1}^{\infty} P\{y - s_- + (n-1)Q < \underline{b} \leqq y - S + nQ\} =$$

(8.49)

$$= W(y-S) + \sum_{n=1}^{\infty} \{W(y-S+nQ) - W(y-s_-+(n-1)Q)\} \quad \text{voor } S \leqq y$$

en

$$1 - K(S|y,z) = \sum_{n=1}^{\infty} P\{y - s_- + (n-1)Q < \underline{b} \leqq y - S + nQ\} =$$

(8.50)

$$= \sum_{n=1}^{\infty} \{W(y-S+nQ) - W(y-s_-+(n-1)Q)\} \quad \text{voor } S > y.$$

De verdelingsfunktie $K(S|y,z)$ heeft als kansdichtheid

$$(8.51) \qquad k(S|y,z) = \begin{cases} \sum_{n=0}^{\infty} w(y-S+nQ) & \text{voor } S \leqq y, \\ \\ \sum_{n=1}^{\infty} w(y-S+nQ) & \text{voor } S > y. \end{cases}$$

De kansdichtheid $k(S|y,z)$ heeft de volgende bijzondere eigenschap:

$$\int_{s_-}^{s_-+Q} k(S|y,z)dy = \int_{s_-}^{S} \{\sum_{n=1}^{\infty} w(y-S+nQ)\}dy + \int_{S}^{s_-+Q} \{\sum_{n=0}^{\infty} w(y-S+nQ)\}dy =$$

(8.52)

$$= \sum_{n=0}^{\infty} \left\{\int_{s_-}^{S} w(nQ+Q+y-S)dy + \int_{S}^{s_-+Q} w(nQ+y-S)dy\right\} =$$

$$= \sum_{n=0}^{\infty} \left\{\int_{s_-+Q-S}^{Q} w(nQ+v)dv + \int_{0}^{s_-+Q-S} w(nQ+v)dv\right\} =$$

$$= \sum_{n=0}^{\infty} \left\{\int_{0}^{Q} w(nQ+v)dv\right\} = \int_{0}^{\infty} w(u)du = 1 \quad \text{voor } s_- \leqq S < s_-+Q.$$

Uit bovenstaande relatie volgt dat q(S,z) = constante voldoet aan

$$(8.53)\qquad q(S,z) = \int_{s_-}^{s_-+Q} k(S|y,z)\, q(y,z)\, dy \quad \text{voor } s_- \leq S < s_-+Q.$$

De gezochte funktie q(S,z) is de kansdichtheid van een uniek bepaalde invariante kansverdeling en dus geldt

$$(8.54)\qquad q(S,z) = \begin{cases} \frac{1}{Q} & \text{voor } s_- \leq S < s_-+Q, \\ 0 & \text{anders.} \end{cases}$$

Als aan het begin van een periode de toestand (d.i. de voorraad vlak na beslissen) S is, treedt in die periode een tekort op met kans 1 - W(S), waarbij W(S) = 0 voor S < 0. De fractie van het aantal perioden waarin een tekort optreedt, is derhalve met kans 1 gelijk aan

$$(8.55)\qquad P_{tekort} = \frac{1}{Q}\int_{s_-}^{s_-+Q} \{1 - W(S)\}\, dS\ .$$

Uit (8.48) en (8.54) volgt dat de kriteriumfunktie y(z) gelijk is aan

$$(8.56)\qquad y(z) = \frac{1}{Q}\int_{s_-}^{s_-+Q} \left\{h(S) + \int_{S-s_-}^{\infty} \phi(Q_{S-v})\, w(v)\, dv\right\} dS\ .$$

Veronderstel dat

$$(8.57)\qquad \phi(u) = \begin{cases} cu + K & \text{als } u > 0 \\ 0 & \text{als } u = 0. \end{cases}$$

De bijdrage van de vaste kosten K aan y(z) wordt gegeven door

$$(8.58)\qquad \frac{1}{Q}\int_{s_-}^{s_-+Q} \left\{\int_{S-s_-}^{\infty} K\, w(v)\, dv\right\} dS = \frac{K}{Q}\int_{s_-}^{s_-+Q} \{1 - W(S-s_-)\}\, dS = $$
$$= \frac{K}{Q}\int_{0}^{Q} \{1 - W(x)\}\, dx\ .$$

De bijdrage van de inkoopkosten c aan y(z) berekenen wij met het volgende argument. In het naleveringsmodel wordt alles wat gevraagd wordt ook ingekocht. De verwachte vraag per periode is

(8.59) $$\mu \overset{\text{def}}{=} \int_0^\infty v\, w(v)\, dv.$$

De sterke wet der grote aantallen leert ons nu dat over een oneindige tijdsduur genomen de gemiddelde per periode ingekochte hoeveelheid met kans 1 gelijk is aan μ. Dus $c\mu$ is de bijdrage van c aan y(z); wij kunnen (8.56) derhalve vereenvoudigen tot

(8.60) $$y(z) = \frac{1}{Q}\int_{s_-}^{s_-+Q} h(S)\, dS + \frac{K}{Q}\int_0^Q \{1 - W(x)\}\, dx + c\mu .$$

Door de partiële afgeleiden van y(z) naar s_- en Q nul te stellen, vinden wij dat de optimale s_- (= s^*) en Q (= Q^*) gevonden kunnen worden als oplossing van de twee niet-lineaire vergelijkingen

(8.61) $$\int_0^Q h'(s_-+u)\, du = 0$$

en

(8.62) $$Q\{h(s_-+Q) + K(1-W(Q))\} - \int_0^Q \{h(s_-+u) + K(1-W(u))\}\, du = 0 .$$

Stel dat de behoefte in een periode een exponentiële kansverdeling met parameter λ heeft. Neem verder aan dat de voorraad- en boetekosten in een periode gelijk zijn aan c_1 maal de aanwezige voorraad resp. c_2 maal de achterstand aan het eind van die periode. De voorraad- en boetekostenfunktie h(S) wordt dan gegeven door de formules (6.180) en (6.181) op blz. 49. De vergelijkingen (8.61) en (8.62) kunnen (onder de aanname dat $s_- \geq 0$ is) vereenvoudigd worden tot

(8.63) $$s_- = \frac{1}{\lambda} \log \frac{(c_1+c_2)(1-e^{-\lambda Q})}{\lambda c_1 Q}$$

en

$$(8.64) \qquad (1-e^{-\lambda Q})\,\{K(\lambda Qe^{-\lambda Q}+e^{-\lambda Q}-1) - \frac{\lambda c_1 Q^2}{2} - c_1 Q\} + \lambda c_1 Q^2 = 0 \,.$$

Zonder op de details in te gaan, vermelden wij dat uit een analyse van het linkerlid van (8.64) en zijn eerste drie afgeleiden volgt, dat vergelijking (8.64) precies één positieve wortel heeft als $K > 0$ is; voor $K = 0$ heeft (8.64) geen positief nulpunt en ontaardt de optimale (s_-,nQ)-strategie in een (s_-,s_-)-strategie (zie ook in voorbeeld 8.4 formule (8.98) op blz. 157).

Voor het getallenvoorbeeld met

$$(8.65) \qquad \lambda = 0{,}02 \;;\; c = c_1 = 1 \;;\; c_2 = 15 \text{ en } K = 30$$

vinden wij met behulp van de rekenautomaat

$$(8.66) \qquad s^* = 106{,}082,\; Q^* = 74{,}079 \;\text{ en } y(z^*) = 208{,}768 \,.$$

Opmerking 8.1

Als met een bestelling een levertijd van L perioden (L vast en geheel) gepaard gaat, baseren wij de inkoopstrategie $z = (s_-,nQ)$ op de economische voorraad i.p.v. op de voorraad. De economische voorraad is gedefinieerd als de som van de voorraad en de hoeveelheden in bestelling. Omdat een tekort nageleverd wordt, is de kansverdeling van de economische voorraad in het model met een vaste levertijd dezelfde als de kansverdeling van de voorraad in het model zonder levertijd. De invariante kansverdeling van de economische voorraad S vlak na beslissen wordt derhalve gegeven door $q(S,z) = \frac{1}{Q}$ voor $s_- \leqq S < s_-+Q$.

De bestelkosten in de eerste periode en de voorraad- en boetekosten in de eerste L perioden zijn eindig; dus als wij deze kosten buiten beschouwing laten, veranderen de gemiddelde kosten per tijdseenheid genomen over een oneindige tijdsduur niet. Alle overige te maken inkoop-, voorraad- en boetekosten splitsen wij over de perioden op de volgende ondubbelzinnige wijze: als aan het begin van een periode de toestand (d.i. de economische voorraad vlak na beslissen) gelijk aan S is, kennen wij aan die periode de kosten $h(S,z(S))$ toe, waarbij deze kosten gedefinieerd zijn als de som van de verwachte bestelkosten in de volgende en de verwachte voorraad- en boetekosten in de L perioden verder gelegen periode. Als de toestand nu S is, zijn de verwachte voorraad- en boetekosten in de L perioden ($L \geq 1$) verder gelegen periode gelijk aan (vgl. blz. 54 en 55)

$$(8.67) \qquad h_L(S) = \int_0^\infty h(S-v)\, w_L(v)\, dv \ ,$$

waarbij $w_L(v)$ de kansdichtheid is van de totale behoefte in L perioden. Men kan nu gemakkelijk nagaan dat de formules (8.56) en (8.60) t/m (8.62) ook voor het naleveringsmodel met een levertijd van L perioden geldig zijn, mits wij in deze formules $h(S)$ vervangen door $h_L(S)$.

Wij merken op dat de invariante kansverdeling van de economische voorraad vlak na beslissen zowel de invariante kansverdeling van de voorraad aan het begin van een periode vlak na een eventuele aflevering als de invariante kansverdeling van de voorraad aan het eind van een periode op ondubbelzinnige wijze bepaalt. Wij zullen alleen de laatstgenoemde kansverdeling berekenen. Uit deze kansverdeling volgt hoe "vaak" een tekort ontstaat. De voorraad aan het eind van periode $n + L$ is gelijk aan de economische voorraad vlak na beslissen in periode n minus de totale behoefte in de perioden $n,n+1,\ldots,n+L$. De verdelingsfunktie van de totale behoefte in k perioden geven wij aan met $W_k(u)$, waarbij wij $W_k(u) = 0$ definiëren voor $u < 0$. Als aan het begin van een periode de economische voorraad vlak na beslissen S is, dan is de voorraad aan het eind van de L perioden verder gelegen periode $\leq v$ met kans $1 - W_{L+1}(S-v)$. De invariante kansverdeling van de voorraad aan het eind van een periode wordt dus gegeven door

$$(8.68) \qquad E(v,z) = \frac{1}{Q}\int_{S_-}^{S_-+Q} \{1 - W_{L+1}(S-v)\}\, dS \ .$$

De fractie van het aantal perioden waarin een tekort ontstaat is derhalve met kans 1 gelijk aan

$$(8.69) \qquad P_{\text{tekort}} = E(0,z).$$

Opmerking 8.2

Voor het geval de behoefte discreet verdeeld is, kunnen wij op dezelfde wijze te werk gaan als in het geval met een continu verdeelde behoefte. Ook dan vinden wij voor de invariante kansverdeling van de voorraad resp. economische voorraad vlak na bestellen een homogene verdeling; dus elk der Q toestanden $s_-, s_-+1, \ldots, s_-+Q-1$ heeft al invariante kans $\frac{1}{Q}$.

Voorbeeld 8 4 De (S^+, s_-)-strategie voor diverse voorraadmodellen.

a) Het nalevenringsmodel zonder l vertijd

a) Continu verdeelde vraag

Wij nemen aan dat de verdelingsfunktie W(u) van de behoefte in een periode een kansdichtheid w(u) bezit. Laat μ de verwachting van de behoefte in een periode zijn. De verdelingsfunktie en de kansdichtheid van de totale behoefte in j perioden geven wij aan met $W_j(u)$ resp. $w_j(u)$ (dus $W_1(u) = W(u)$, $w_1(u) = w(u)$). Wij definiëren

$$(8.70) \qquad W(u;1) = \sum_{j=1}^{\infty} W_j(u) \quad \text{en} \quad w(u;1) = \sum_{j=1}^{\infty} w_j(u) \qquad \text{voor } u \geq 0.$$

Voor de funkties W(u;1) en w(u,1) geldt

$$(8.71) \qquad W(u;1) = \int_0^u w(v;1)\, dv \qquad \text{voor } u \geq 0.$$

Met behulp van relatie (6.135) op blz 38 kunnen wij aantonen dat w(u;1) voldoet aan de integraalvergelijking

$$(8.72) \qquad w(u;1) = w(u) + \int_0^u w(u-v;1)\, w(v)\, dv \qquad \text{voor } u \geq 0.$$

Een beroemde stelling uit de waarschijnlijkheidsrekening leert ons dat $W(x;1)/x$ naar $1/\mu$ convergeert voor $x \to \infty$; m.a.w. voor grote x wordt W(x;1) benaderd door x/μ. Voorts vermelden wij zonder bewijs dat 1 + W(x;1) geïnterpreteerd kan worden als het verwachte aantal perioden nodig voor een totale vraag groter dan x. Dus als een (S^+,s_-)-strategie wordt toegepast, dan is $1 + W(S^+-s_-;1)$ het verwachte aantal perioden tussen twee opeenvolgende bestellingen.

Een (S^+,s_-)-strategie heeft de volgende eenvoudige vorm: als aan het begin van een periode de voorraad kleiner dan s_- is, wordt de voorraad tot S^+ aangevuld; is de voorraad s_- of meer, dan wordt niet aangevuld.

De toestand van het systeem aan het begin van een periode definiëren wij als de omvang van de voorraad vlak voor beslissen. Stel dat strategie $z = (S^+,s_-)$ wordt toegepast. De toestand neemt dan alleen waarden $\leq S^+$ aan. Als aan het begin van een periode de toestand S is, zijn de verwachte kosten in die periode

$$(8.73) \qquad h(S,z(S)) = \begin{cases} h(S) & \text{voor } s_- \leq S \leq S^+, \\ \phi(S^+-S) + h(S^+) & \text{voor } \quad S < s_-. \end{cases}$$

Voor de verdelingsfunktie $K(S|y,z)$ geldt

$$(8.74) \qquad K(S|y,z) = \begin{cases} 1 - W(S^+-S) & \text{voor } \quad y < s_- \text{ en } \quad S \leq S^+, \\ 1 - W(y-S) & \text{voor } s_- \leq y \leq S^+ \text{ en } \quad S \leq y\ , \\ 1 & \text{voor } s_- \leq y \leq S^+ \text{ en } y < S \leq S^+. \end{cases}$$

Voor de kansdichtheid van de verdelingsfunktie $K(S|y,z)$ geldt dus

$$(8.75) \qquad k(S|y,z) = \begin{cases} w(S^+-S) & \text{voor } \quad y < s_- \text{ en } \quad S \leq S^+, \\ w(y-S) & \text{voor } s_- \leq y \leq S^+ \text{ en } \quad S \leq y\ , \\ 0 & \text{voor } s_- \leq y \leq S^+ \text{ en } y < S \leq S^+. \end{cases}$$

De invariante kansdichtheid $q(S,z)$ voldoet derhalve aan

$$(8.76) \quad q(S,z) = \int_{-\infty}^{s_-} w(S^+-S)q(y,z)dy + \int_{S}^{S^+} w(y-S)q(y,z)dy \quad \text{voor } s_- \leq S \leq S^+,$$

resp.

$$(8.77)\quad q(S,z) = \int_{-\infty}^{s_-} w(S^+-S)q(y,z)dy + \int_{s_-}^{S^+} w(y-S)q(y,z)dy \qquad \text{voor } S < s_- .$$

Voeren wij de variabelen $u = S^+ - S$ en $v = y - S^+ + u$ in, dan kunnen wij (8.76) herleiden tot

$$(8.78)\quad q(S^+-u,z) = w(u)\int_{-\infty}^{s_-} q(y,z)dy + \int_0^u q(S^+-u+v,z)w(v)dv \quad \text{voor } 0 \leqq u \leqq S^+-s_-$$

Indien we definiëren

$$(8.79)\qquad q^*(x,z) = \frac{q(S^+-x,z)}{\int_{-\infty}^{s_-} q(y,z)dy} \qquad \text{voor } 0 \leqq x \leqq S^+-s_- ,$$

dan kunnen wij (8.78) schrijven als

$$(8.80)\qquad q^*(u,z) = w(u) + \int_0^u q^*(u-v,z)w(v)dv \quad \text{voor } 0 \leqq u \leqq S^+-s_- .$$

Uit (8.72) volgt dat

$$(8.81)\qquad q^*(u,z) = w(u;1) \qquad \text{voor } 0 \leqq u \leqq S^+-s_-$$

voldoet aan (8.80). Het bewijs dat deze oplossing uniek is, laten wij achterwege. Uit (8.79), (8.81) en (8.71) volgt

$$(8.82)\qquad \int_0^{S^+-s_-} q(S^+-u,z)\, du = W(S^+-s_-;1) \int_{-\infty}^{s_-} q(y,z)\, dy .$$

Door gebruik te maken van het feit dat $q(S,z)$ een kansdichtheid is, kunnen wij (8.82) herleiden tot

$$(8.83)\qquad 1 - \int_{-\infty}^{s_-} q(y,z)\, dy = W(S^+-s_-;1) \int_{-\infty}^{s_-} q(y,z)\, dy .$$

Dus

$$(8.84) \qquad \int_{-\infty}^{s_-} q(y,z)\, dy = \frac{1}{1 + W(S^+ - s_-;1)} .$$

Uit (8.77), (8.79), (8.81) en (8.84) volgt tenslotte

$$(8.85)\ q(S,z) = \begin{cases} \dfrac{w(S^+-S;1)}{1 + W(S^+-s_-;1)} & \text{voor } s_- \leq S \leq S^+, \\[2ex] \dfrac{w(S^+-S) + \int_0^{S^+-s_-} w(S^+-S-u)w(u;1)du}{1 + W(S^+-s_-;1)} & \text{voor } S < s_- . \end{cases}$$

De fractie van het aantal perioden waarin een tekort optreedt, is met kans 1 gelijk aan

$$(8.86) \qquad P_{\text{tekort}} = \int_{-\infty}^{0} q(S,z)\, dS.$$

De kriteriumfunktie $y(z)$ wordt gegeven door

$$(8.87)\ y(z) = \int_{-\infty}^{s_-} \{\phi(S^+-S) + h(S^+)\}\, q(S,z)dS + \int_{s_-}^{S^+} h(S)\, q(S,z)\, dS .$$

Met behulp van (8.84) kunnen wij (8.87) herschrijven tot

$$(8.88)\ y(z) = \frac{h(S^+) + \int_0^{S^+-s_-} h(S^+-u)w(u;1)du}{1 + W(S^+-s_-;1)} + \int_{-\infty}^{s_-} \phi(S^+-S)\, q(S,z)\, dS .$$

Veronderstel dat

$$(8.89) \qquad \phi(x) = \begin{cases} cx + K & \text{voor } x > 0, \\ 0 & \text{voor } x = 0. \end{cases}$$

Dan geldt (vgl. (8.84))

$$(8.90)\quad \int_{-\infty}^{s_-} \phi(S^+-S)\; q(S,z)\; dS = \frac{K}{1 + W(S^+-s_-;1)} + c \int_{-\infty}^{s_-} (S^+-S)\; q(S,z)\; dS \; .$$

De bijdrage van de inkoopkosten c aan y(z) is gelijk aan $c\mu$ (zie blz.149). Hieruit kan men afleiden dat de tweede term uit het rechterlid van (8.90) gelijk aan $c\mu$ is. Onder de veronderstelling (8.89) kan de kriteriumfunktie y(z) derhalve vereenvoudigd worden tot

$$(8.91)\qquad y(z) = \frac{h(S^+) + \int_0^{S^+-s_-} h(S^+-u)w(u;1)du + K}{1 + W(S^+-s_-;1)} + c\mu \; .$$

Kiezen wij S^+-s_- en S^+ als variabelen i.p.v. s_- en S^+ en stellen wij de partiële afgeleiden van y(z) naar S^+-s_- en S^+ gelijk aan nul, dan vinden wij dat de optimale s_- (= s^*) en S^+ (= S^*) voldoen aan

$$(8.92)\qquad h(s_-)\; \{1 + W(S^+-s_-;1)\} = h(S^+) + \int_0^{S^+-s_-} h(S^+-u)w(u;1)du + K$$

en

$$(8.93)\qquad h'(S^+) + \int_0^{S^+-s_-} h'(S^+-u)\; w(u;1)\; du = 0 \; .$$

Deze twee niet-lineaire vergelijkingen kunnen i.h.a. alleen numeriek opgelost worden. Wij merken op dat de vergelijkingen (8.92) en (8.93) een sterke overeenkomst vertonen met de vergelijkingen (6.174) en (6.175) op blz. 47 resp. 48, die behoren bij het voorraadmodel waarin de kosten verdisconteerd worden. Stellen wij de verdisconteringsfactor $\alpha = 1$ in (6.174) en (6.175), dan gaan deze vergelijkingen over in (8.92) en (8.93).

Als K = 0 is en wij ons beperken tot (S^+,s_-)-strategieën met $S^+ = s_-$, dan voldoet de optimale s_- (= S^+) aan

(8.94) $$h'(s_-) = 0 \ .$$

Ter illustratie veronderstellen wij dat de behoefte in een periode exponentieel verdeeld is met als kansdichtheid

(8.95) $$w(u) = \lambda e^{-\lambda u} \qquad \text{voor } u \geqq 0 \ .$$

Dan geldt

(8.96) $$W(u;1) = \lambda u \ , \ w(u;1) = \lambda \qquad \text{voor } u \geqq 0 \ .$$

Nemen wij tevens aan dat de voorraad- en boetekosten in een periode evenredig zijn met de aanwezige voorraad resp. achterstand aan het eind van die periode (evenredigheidsconstanten c_1 resp. c_2), dan wordt de funktie h(S) gegeven door (6.180) en (6.181) op blz. 49 en kunnen de vergelijkingen (8.92) en (8.93) analytisch opgelost worden. Onder de veronderstelling dat $s^* \geqq 0$ is, vinden wij na enig rekenwerk voor S^*-s^* de befaamde wortelformule

(8.97) $$S^* - s^* = \sqrt{\frac{2K}{\lambda c_1}}$$

en blijkt s^* gelijk te zijn aan

(8.98) $$s^* = \frac{1}{\lambda} \log \frac{c_1 + c_2}{c_1 + \sqrt{2\lambda c_1 K}} \ .$$

Deze formules zijn alleen geldig als $\sqrt{2\lambda c_1 K} \leqq c_2$ is. Voor het getallenvoorbeeld met

(8.99) $$\lambda = 0{,}02 \ ; \ c = c_1 = 1 \ ; \ c_2 = 15 \text{ en } K = 30$$

vinden wij

(8.100) $$s^* = 101{,}641 \ , \ S^* = 156{,}413 \text{ en } y(z^*) = 206{,}413 \ .$$

a2) Discreet verdeelde vraag

Laat p(i) de kans zijn dat de behoefte in een periode gelijk aan i is. De verwachting van de behoefte in een periode geven wij wederom aan met μ. De kans dat de totale behoefte in j perioden gelijk aan i is, geven wij aan met $p_j(i)$ (dus $p_1(i) = p(i)$). Wij definiëren

(8.101) $$p(i;1) = \sum_{j=1}^{\infty} p_j(i) \qquad \text{voor } i = 0,1,\ldots .$$

Met behulp van de relatie (6.144) op blz. 40 vinden wij dat $p(i;1)$ voldoet aan de recursieve betrekking

(8.102) $$p(i;1) = p(i) + \sum_{r=0}^{i} p(i-r;1)\, p(r) \qquad \text{voor } i = 0,1,\ldots .$$

Wij vermelden zonder bewijs dat $\sum_{i=0}^{k} p(i;1)$ voor grote k benaderd wordt door k/μ en dat $1 + \sum_{i=0}^{k} p(i;1)$ geïnterpreteerd kan worden als het verwachte aantal perioden nodig voor een totale vraag groter dan k. Dus als een (S^+,s_-)-strategie wordt toegepast, dan is $1 + \sum_{i=0}^{S^+-s_-} p(i;1)$ het verwachte aantal perioden tussen twee opeenvolgende bestellingen.

Op analoge wijze als (8.85) afgeleid is, kan men aantonen dat de invariante kans op een voorraad S vlak voor beslissen gelijk is aan

(8.103) $$q_S[z] = \begin{cases} \dfrac{p(S^+-S;1)}{1 + \sum_{k=0}^{S^+-s_-} p(k;1)} & \text{voor } S = s_-,\ldots,S^+, \\[2ex] \dfrac{p(S^+-S) + \sum_{k=0}^{S^+-s_-} p(S^+-S-k)p(k;1)}{1 + \sum_{k=0}^{S^+-s_-} p(k;1)} & \text{voor } S < s_-. \end{cases}$$

De fractie van het aantal perioden waarin een tekort optreedt, is met kans 1 gelijk aan

(8.104) $$P_{tekort} = \sum_{S=-\infty}^{-1} q_S[z] .$$

De kriteriumfunktie y(z) is gelijk aan

$$(8.105)\quad y(z) = \frac{h(S^+) + \sum_{k=0}^{S^+-s_-} h(S^+-k)p(k;1)}{1 + \sum_{k=0}^{S^+-s_-} p(k;1)} + \sum_{S=-\infty}^{s_--1} \phi(S^+-S)\, q_S[z] \; .$$

Als

$$(8.106)\qquad \phi(i) = \begin{cases} ci + K & \text{voor } i = 1,2,\ldots \\ 0 & \text{voor } i = 0 \; , \end{cases}$$

dan kan y(z) vereenvoudigd worden tot

$$(8.107)\qquad y(z) = \frac{h(S^+) + \sum_{k=0}^{S^+-s_-} h(S^+-k)\; p(k;1) + K}{1 + \sum_{k=0}^{S^+-s_-} p(k;1)} + c\mu \; .$$

Als K = 0 is en wij ons beperken tot (S^+,s_-)-strategieën met $S^+ = s_-$, dan volgt uit (8.107) dat de optimale s_- (= S^+) gelijk is aan het gehele getal S_0 waarvoor h(S) minimaal is.

b) <u>Het naleveringsmodel met vaste levertijd.</u>

Als met een bestelling een vaste levertijd van L (L $\geq$ 1 en geheel) perioden gepaard gaat, wordt de (S^+,s_-)-strategie gebaseerd op de economische voorraad i.p.v. op de voorraad. De invariante kansverdeling van de economische voorraad vlak <u>voor</u> beslissen wordt gegeven door (8.85) als de vraag een kansdichtheid bezit en door (8.103) als de vraag discreet verdeeld is. Uit de invariante kansverdeling van de economische voorraad vlak voor beslissen kunnen wij zowel de invariante kansverdeling van de voorraad aan het begin van een periode als de invariante kansverdeling van de voorraad aan het eind van een periode berekenen. Men gaat gemakkelijk na dat de invariante kansverdeling van de voorraad aan het eind van een periode gegeven wordt door

$$(8.108)\qquad E(v,z) = \int_{-\infty}^{S^+} \{1 - W_L(S-v)\}\, q(S,z)\, dS$$

als de vraag een kansdichtheid bezit en door

$$(8.109)\qquad e_v[z] = \sum_{S=-\infty}^{S^+} p_L(S-v)\, q_S[z]$$

als de vraag discreet verdeeld is. De kans P_{tekort} is gelijk aan $E(0,z)$ resp. $\sum_{v=-\infty}^{-1} e_v[z]$.

Voor het naleveringsmodel met een levertijd van L perioden kennen wij als kosten aan een periode toe de som van de bestelkosten in die periode en de voorraad- en boetekosten in de L perioden verder gelegen periode. Men gaat gemakkelijk na dat de formules behorende bij het model met een levertijd van L ($L \geqq 1$) perioden verkregen worden door in de formules behorende bij het model zonder levertijd de funktie h(S) te vervangen door

$$(8.110)\qquad h_L(S) = \int_0^{\infty} h(S-v)\, w_L(v)\, dv$$

in het geval van een continu verdeelde vraag, en door

$$(8.111)\qquad h_L(S) = \sum_{k=0}^{\infty} h(S-k)\, p_L(k)$$

in het geval van een discreet verdeelde vraag.

c. Schattingen en benaderingsformules voor de optimale s_- en S^+

Schattingen en benaderingsformules voor de optimale s_- ($= s^*$) en S^+ ($= S^*$) zullen worden gegeven onder de veronderstelling dat

$$(8.112)\qquad \phi(x) = \begin{cases} cx + K & \text{voor } x > 0, \\ 0 & \text{voor } x = 0. \end{cases}$$

De te geven schattingen en benaderingsformules gelden voor elke niet-negatieve, gehele waarde van de levertijd L.

De schattingen worden berekend uit de funktie $h_L(S)$, die voor elke $L \geqq 0$ gedefinieerd is als de verwachte voorraad- en boetekosten in periode n+L als aan het begin van periode n de economische voorraad vlak <u>na</u> beslissen gelijk aan S is (vgl. blz. 55). Om de schattingen te kunnen maken moeten wij verder aannemen dat er een getal S_0 bestaat zodanig dat de funktie $h_L(S)$ voor $S \leqq S_0$ monotoon dalend en voor $S \geqq S_0$ monotoon stijgend is. Wij merken hierbij op dat onder bovenstaande veronderstellingen aangetoond kan worden dat een optimale voorraadstrategie bestaat, die tot de klasse van de (S^+, s_-)-strategieën behoort. Voor bewijzen wordt de geïnteresseerde lezer verwezen naar de twee in de voetnoot op blz. 59 genoemde artikelen. Verder merken wij op dat op eenvoudige wijze kan worden aangetoond dat er een getal S_0 met bovengenoemde eigenschap bestaat als de voorraad- en boetekosten in een periode gelijk zijn aan c_1 resp. c_2 vermenigvuldigd met de aanwezige voorraad resp. achterstand aan het eind van die periode; in het geval van een continu resp. discreet verdeelde vraag voldoet S_0 aan (vgl. (6.207) en (6.214))

(8.113a) $$W_{L+1}(S_0) = \frac{c_2}{c_1 + c_2}$$

resp.

(8.113b) $$\sum_{k=0}^{S_0-1} p_{L+1}(k) < \frac{c_2}{c_1 + c_2} \leqq \sum_{k=0}^{S_0} p_{L+1}(k).$$

Wij zullen nu de in het artikel van VEINOTT en WAGNER gevonden schattingen voor de optimale s^* en S^* geven. Daartoe definiëren wij voor het geval de vraag continu verdeeld is de getallen $s_1 \leqq S_0$ en $S_1 \geqq S_0$ door

(8.114) $$h_L(s_1) = h_L(S_0) + K \quad \text{en} \quad h_L(S_1) = h_L(S_0) + K.$$

Als de vraag echter discreet verdeeld is, wordt s_1 gedefinieerd als het kleinste gehele getal $\leqq S_0$ dat voldoet aan

(8.115) $$h_L(s_1) \leqq h_L(S_0) + K,$$

en wordt S_1 gedefinieerd als het kleinste gehele getal $\geqq S_0$ dat voldoet aan

$$(8.116) \qquad h_L(S_1+1) \geqq h_L(S_0) + K.$$

Zowel voor het continue als het discrete geval geldt

$$(8.117) \qquad s_1 \leqq s^* \leqq S_0 \leqq S^* \leqq S_1.$$

Als de inkoopkosten gegeven worden door (8.112) en de voorraad- en boetekosten evenredig zijn met de aanwezige voorraad resp. achterstand aan het eind van een periode, kunnen tevens benaderingsformules voor s^* en S^* gegeven worden. Daarbij veronderstellen wij dat $s^* \geqq 0$ is. Onder de veronderstelling dat de vaste bestelkosten K en de boetekosten c_2 beide voldoende groot zijn, zijn benaderingsformules gevonden door D.M. ROBERTS.*) Voor het continue geval geldt bij benadering

$$(8.118) \qquad S^* - s^* \approx \sqrt{\frac{2K\mu}{c_1}} \quad \text{en} \quad c_2 \int_{s^*}^{\infty} (v-s^*)\, w_{L+1}(v)\, dv \approx \sqrt{2K\mu c_1}$$

en voor het discrete geval

$$(8.119) \qquad S^* - s^* \approx \sqrt{\frac{2K\mu}{c_1}}$$

en

$$(8.120) \qquad c_2 \sum_{k=s^*+1}^{\infty} (k-s^*)\, p_{L+1}(k) < \sqrt{2K\mu c_1} \leqq c_2 \sum_{k=s^*}^{\infty} (k-s^*+1)\, p_{L+1}(k),$$

waarin μ de verwachte vraag per periode is en c_1 de voorraadkosten per eenheid voorstelt. Het artikel van ROBERTS bevat tevens meer verfijnde benaderingsformules, die zich echter minder gemakkelijk tot snelle berekeningen lenen. De door (8.119) en (8.120) bepaalde benaderingen van $S^* - s^*$ en s^* geven wij aan met D_R resp. s_R. De benaderingen D_R en s_R zijn zowel voor

---

*) D.M. ROBERTS, Approximations to optimal policies in a dynamic inventory model; hoofdstuk 13 uit het boek van K.J. ARROW, S. KARLIN and H, SCARF (ed.), Studies in applied probability and management science, Stanford University Press, Stanford, (1962), p. 207-230.

een Poisson-, een negatief binomiale als een geometrische vraagverdeling $\{p(k),\ k \geq 0\}$ uitvoerig getest door H.M. WAGNER e.a. *) In dit artikel hebben de auteurs daarnaast nog onderzocht hoe gevoelig de optimale (S^+,s_-)-strategie is voor de verschillende parameters en bovendien zijn andere numerieke methoden getest, waarmee de optimale (S^+,s_-)-strategie benaderd kan worden. Het testkriterium is de verhouding tussen de gemiddelde kosten per tijdseenheid van de door de benaderingsmethode gevonden (S^+,s_-)-strategie en de gemiddelde kosten per tijdseenheid van een optimale (S^+,s_-)-strategie. Wij zullen enkele algemene conclusies uit het bovengenoemde artikel geven zonder in details te treden. De merites van de diverse benaderingsmethoden blijken significant afhankelijk van de beschouwde vraagverdeling. De benadering van ROBERTS is voor een groot aantal representatieve voorbeelden getest en wel voor 240 voorbeelden met een Poissonverdeling, 240 voorbeelden met een negatief binomiale verdeling en voor 216 voorbeelden met een geometrische vraagverdeling. In tabel 8.1 is voor elk van de vraagverdelingen aangegeven voor welk percentage van de beschouwde voorbeelden de gemiddelde kosten per tijdseenheid behorende bij de benadering van ROBERTS de minimale gemiddelde kosten per tijdseenheid met minder dan 1, 5, 10 resp. 25% overschrijden. De percentages tussen haakjes hebben betrekking op de empirische benadering, die op blz. 164 geformuleerd zal worden.

verdeling / afwijking	Poisson	Negatief binomiaal	Geometrisch
< 1%	46% (75%)	33% (64%)	17%
< 5%	68% (95%)	69% (95%)	58%
< 10%	77% (98,3%)	84% (99,6%)	75%
< 25%	88% (100%)	95% (100%)	94%

Tabel 8.1

*) H.M. WAGNER, M. O'HAGAN and B. LUNDH, An empirical study of exactly and approximately optimal inventory policies, Management Science, 11, (1965), p. 690-723.

Uit numerieke berekeningen blijkt verder dat de benadering van ROBERTS beter wordt naarmate K toeneemt. Bovendien kan geconcludeerd worden dat de benadering slechter wordt als μ toeneemt. Opmerkelijk is dat in bijna alle voorbeelden met een Poisson- of een negatief binomiale verdeling s_R weinig verschilt van s^*, zelfs als D_R significant groter dan S^*-s^* is. Voor de geometrische verdeling daarentegen blijkt s_R vaak aanzienlijk groter te zijn dan s^* en is D_R een uitstekende benadering van S^*-s^*.

In tegenstelling tot de geometrische verdeling is voor de Poisson- en de negatief binomiale verdeling de afwijking van D_R met S^*-s^* de voornaamste reden voor het slechter worden van de benadering van ROBERTS als μ toeneemt. Zoals uit relatie (8.119) volgt, is de grootheid D_R een monotoon stijgende funktie van μ. Uit numerieke berekeningen blijkt echter dat S^*-s^*, gezien als funktie van μ, niet behoeft toe te nemen als μ toeneemt; S^*-s^* kan bij toenemende μ plotseling een sprong naar beneden maken. Naar aanleiding van dit empirische gedrag van S^*-s^* als funktie van μ heben WAGNER c.s. een empirische benaderingsformule afgeleid. Als de grootheid D_R/μ, die een maat is voor de frekwentie waarmee besteld wordt, kleiner dan of gelijk aan 1,5 is of, equivalent hiermee, als

$$(8.126) \qquad \mu \geqq 0,8888 \frac{K}{c_1}$$

is, dan stellen zij de empirische benaderingsformule

$$(8.122) \qquad s^* \approx \min(s_R, S_0) \ , \quad S^* \approx \min(s_R+D_R, S_0)$$

voor, waarin S_0 bepaald wordt door de tweede relatie uit (8.113). Zowel voor de Poisson- als de negatief binomiale verdeling leidt deze empirische benaderingsformule tot aanzienlijk betere resultaten (zie de percentages tussen haakjes in tabel 8.1).

Tenslotte merken wij nog op dat uit de in het artikel van WAGNER e.a. gedane berekeningen blijkt dat de beste (s_-, nQ)-strategie (zie voorbeeld 8.3) een bijzonder goede approximatie is van de optimale (S^+, s_-)-strategie; d.w.z. de waarden voor de gemiddelde kosten per tijdseenheid zijn vrijwel dezelfde. Voor details verwijzen wij naar het genoemde artikel.

d) Het noodinkoopmodel zonder levertijd

Het noodinkoopmodel zonder levertijd kan op analoge wijze aangepakt worden als het naleveringsmodel. Wij definiëren nu de toestand van het systeem aan het begin van een periode als de omvang van de vlak voor beslissen aanwezige voorraad minus de grootte van de in de vorige periode verrichte hoeveelheid noodinkopen. De invariante kansverdeling van de toestand wordt voor een continu verdeelde vraag gegeven door (8.85) en voor een discreet verdeelde vraag door (8.103). De fractie van het aantal perioden waarin een noodinkoop wordt verricht, is derhalve met kans 1 gelijk aan (8.86) resp. (8.104).

Men gaat gemakkelijk na dat de formules (8.88) resp. (8.105) ook voor het noodinkoopmodel zonder levertijd gelden, mits wij in deze formules

$$(8.123) \qquad \int_{-\infty}^{S_-} \phi(S^+-S)\, q(S,z)\, dS \quad \text{resp.} \quad \sum_{S=-\infty}^{S_--1} \phi(S^+-S)\, q_S[z]$$

vervangen door

$$(8.124) \qquad \int_{-\infty}^{S_-} \phi^*(S^+-S)\, q(S,z)\, dS \quad \text{resp.} \quad \sum_{S=-\infty}^{S_--1} \phi^*(S^+-S)\, q_S[z],$$

waarbij

$$(8.125) \qquad \phi^*(S^+-S) = \begin{cases} \phi(S^+-S) & \text{voor } S > 0, \\ \phi(S^+) & \text{voor } S \leq 0. \end{cases}$$

Door de formules uit (8.124) te herschrijven (vgl. (6.201) op blz. 53) kan men op eenvoudige wijze aantonen, dat als

$$(8.126) \qquad \phi(x) = \begin{cases} cx + K & \text{voor } x > 0, \\ 0 & \text{voor } x = 0, \end{cases}$$

het noodinkoopmodel zonder levertijd kan worden beschouwd als een naleveringsmodel zonder levertijd, indien h(S) vervangen wordt door

$$(8.127) \qquad h^*(S) = h(S) - c \int_S^\infty (v-S)\, w(v)\, dv$$

resp.

$$(8.128) \qquad h^*(S) = h(S) - c \sum_{k=S+1}^{\infty} (k-S)\, p(k).$$

Dus als de vraag exponentieel verdeeld is met parameter λ, de voorraadkosten in een periode c_1 maal de aan het eind van die periode aanwezige voorraad zijn en de noodinkoopkosten c_2 per eenheid bedragen, dan volgt uit (6.180), (8.127), (8.97) en (8.98) dat de optimale s^* en S^* gegeven worden door

$$(8.129) \qquad S^* - s^* = \sqrt{\frac{2K}{\lambda c_1}}\,, \quad s^* = \frac{1}{\lambda} \log \frac{c_1 + c_2 - c}{c_1 + \sqrt{2\lambda c_1 K}}\,.$$

Deze formules zijn alleen geldig als $\sqrt{2\lambda c_1 K} \leqq c_2 - c$ is. Voor het getallenvoorbeeld met

$$(8.130) \qquad \lambda = 0{,}02\,, \quad c = c_1 = 1\,, \quad c_2 = 15 \quad \text{en} \quad K = 30$$

vinden wij

$$(8.131) \qquad s^* = 98{,}414\,, \quad S^* = 153{,}186 \quad \text{en} \quad y(z^*) = 203{,}186.$$

<u>Slotopmerking</u> <u>De totale verwachte verdisconteerde opbrengst en de verwachte gemiddelde opbrengst per tijdseenheid.</u>

In dit deeltje hebben wij voor de beschouwde ∞-stapsbeslissingsproblemen zowel de totale verwachte verdisconteerde opbrengst als de verwachte gemiddelde opbrengst per tijdseenheid als mogelijk optimaliteitskriterium genomen. Bij het eerstgenoemde kriterium worden de te verkrijgen opbrengsten verdisconteerd met een vaste factor α ($0 \leqq \alpha < 1$). Afgezien van een mogelijk economische noodzaak, kan het te prefereren zijn het verdisconteringsmodel te beschouwen. In dit model doet de structuur van het beslissingsproces (= Markov-proces) er niet toe; d.w.z. het is dan voor de bepaling van een optimale strategie niet van belang te weten of het proces disjuncte fuiken *) bevat of niet. Als de verwachte gemiddelde opbrengst per tijdseenheid het kriterium is, dan is het veelal wel nodig te weten of er disjuncte fuiken zijn of niet.

Als de periode tussen twee opeenvolgende beslissingstijdstippen klein is, zal de verdisconteringsfactor α veelal dicht bij 1 liggen. Als α heel dicht bij 1 ligt, kunnen in het verdisconteringsmodel moeilijkheden optreden bij de numerieke berekeningen. Men kan zich dan afvragen of het niet beter is de opbrengsten niet te verdisconteren en de verwachte gemiddelde opbrengst per tijdseenheid als kriterium te nemen. Dit werpt de vraag op of een strategie, die optimaal is met betrekking tot de verwachte gemiddelde opbrengst per tijdseenheid, ook optimaal of "bijna" optimaal is in het verdisconteringsmodel met α voldoende dicht bij 1. In vele beslissingsproblemen blijkt dit inderdaad het geval te zijn. Dit verschijnsel kan nader plausibel worden gemaakt als men weet dat mede op grond van een klassieke stelling uit de analyse **) voor vele beslissingsproblemen kan worden aangetoond dat

$$\lim_{\alpha \uparrow 1} (1-\alpha)\, f(S;z) = y_S(z) \qquad \text{voor alle } S \in \mathcal{S} \text{ en alle } z \in \mathcal{Z}, \tag{8.132}$$

waarin $f(S;z)$ resp. $y_S(z)$ de totale verwachte verdisconteerde opbrengst resp. de verwachte gemiddelde opbrengst per tijdseenheid is als de begintoestand S is en strategie z wordt toegepast. Zoals wij reeds geconstateerd hebben, zullen in vele praktische problemen de beslissingsprocessen geen disjuncte fuiken bevatten en de grootheden $y_S(z)$ niet afhangen van de begintoestand S.

Wij merken op dat voor de voorraadstrategieën $z = (S^+, s_-)$, behandeld in §6 (blz. 40-59) en §8 (blz. 152-166), de relatie (8.132) direct geveri-

*) Een fuik is een (niet-lege) verzameling van toestanden met de eigenschap dat het systeem daar voor altijd in blijft als het er eenmaal in is.

**) Als $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} a_k = a$ en $\sum_{n=0}^{\infty} a_n x^n$ absoluut convergeert voor $0 < x < 1$, dan is $\lim_{x \uparrow 1} (1-x) \sum_{n=0}^{\infty} a_n x^n = a$.

fieerd kan worden met behulp van de gevonden expliciete uitdrukkingen voor $f(S;z)$ en $y(z)$.

Een andere vraag die men zich kan stellen, is in welke mate in het verdisconteringsmodel de optimale strategie afhangt van de verdisconteringsfactor α. Hierover kunnen wij ruwweg zeggen dat in vele gevallen een strategie die optimaal is voor een bepaalde waarde van α, ook optimaal blijft als α iets verandert. In dit verband verwijzen wij ook naar de artikelen genoemd in de voetnoten op blz. 104 en 105, waarin nader op deze kwestie wordt ingegaan.

Het verdisconteringsmodel kan als volgt geïnterpreteerd worden als een model waarin de opbrengsten niet verdisconteerd worden. In plaats van het model met een verdisconteringsfactor $\alpha < 1$ kunnen wij het equivalente model beschouwen waarin de opbrengsten weliswaar niet verdisconteerd worden, maar op elk beslissingstijdstip het beslissingsproces eindigt met kans $1-\alpha$, gegeven dat het proces op het vorige tijdstip nog niet beëindigd was. Als geen opbrengsten meer worden verkregen vanaf het moment dat het beslissingsproces beëindigd is en voor het overige de overgangskansen (onder de voorwaarde dat het proces nog gaande is) en opbrengsten hetzelfde zijn als in het verdisconteringsmodel, dan is in dit absorptiemodel de totale verwachte (niet-verdisconteerde) opbrengst gelijk aan de totale verwachte verdisconteerde opbrengst in het verdisconteringsmodel. Dus het verdisconteringsmodel kan in principe vertaald worden in een model waarin niet verdisconteerd wordt, door een absorptietoestand met opbrengst 0 in te voeren. Voor het absorptiemodel geldt dat, ongeacht de toegepaste strategie en ongeacht de begintoestand, het systeem met kans 1 na een eindige tijd in de absorptietoestand komt om daarin dan voor altijd te blijven; de verwachte gemiddelde opbrengst per tijdseenheid is in het absorptiemodel dus altijd nul. Tenslotte merken wij nog op dat op eenvoudige wijze nagegaan kan worden dat een toepassing van HOWARD's iteratiemethode uit §7 op het eindige absorptiemodel, waarbij in elke iteratiestap de v-waarde van de absorptietoestand nul wordt gesteld, resulteert in een speciaal geval van de in §6 (blz. 18 e.v.) beschreven iteratiemethode voor strategie-verbetering in het verdisconteringsmodel.

Literatuur

G.L. NEMHAUSER, Introduction to dynamic programming, John Wiley & Sons, New York, (1966).

A. KAUFMANN and R. CRUON, Dynamic programming: sequential scientific management, Academic Press, New York, (1967).

R. BELLMAN, Dynamic programming, Princeton University Press, Princeton, (1957).

R. BELLMAN, Adaptive control processes: a guided tour, Princeton University Press, Princeton, (1961).

R. BELLMAN and S.E. DREYFUS, Applied dynamic programming, Princeton University Press, Princeton, (1962).

G. HADLEY, Non-linear and dynamic programming, Addison-Wesley, Reading (Mass.), (1964).

R. HOWARD, Dynamic programming and Markov processes, Technology Press, Cambridge, (1960).

F.S. HILLIER and G.J. LIEBERMAN, Introduction to operations research, Holden-Day, San Francisco, (1967).

O.L.R. JACOBS, An introduction to dynamic programming: the theory of multistage decision processes, Chapman & Hall Ltd., London, (1967).

D.J. WHITE, Dynamic programming, Oliver & Boyd, Edinburgh, (1969).

UITGAVEN IN DE SERIE MC SYLLABUS

Onderstaande uitgaven zijn verkrijgbaar bij het Mathematisch Centrum, 2e Boerhaavestraat 49 te Amsterdam-1005, tel. 020-947272.

MCS 1.1	F. GOBEL & J. VAN DE LUNE, *Leergang Besliskunde, deel 1: Wiskundige basiskennis*, 1965.
MCS 1.2	J. HEMELRIJK & J. KRIENS, *Leergang Besliskunde, deel 2: Kansberekening*, 1965.
MCS 1.3	J. HEMELRIJK & J. KRIENS, *Leergang Besliskunde, deel 3: Statistiek*, 1966.
MCS 1.4	G. DE LEVE & W. MOLENAAR, *Leergang Besliskunde, deel 4: Markovketen, en wachttijden*, 1966.
MCS 1.5	G. DE LEVE & J. KRIENS, *Leergang Besliskunde, deel 5: Inleiding tot de mathematische besliskunde*, 1966.
MCS 1.6a	B. DORHOUT & J. KRIENS, *Leergang Besliskunde, deel 6a: Wiskundige programmering 1*, 1968.
MCS 1.7a	G. DE LEVE, *Leergang Besliskunde, deel 7a: Dynamische programmering 1*, 1968.
MCS 1.7b	G. DE LEVE & H.C. TIJMS, *Leergang Besliskunde, deel 7b: Dynamische programmering 2*, 1970.
MCS 1.7c	G. DE LEVE & H.C. TIJMS, *Leergang Besliskunde, deel 7c: Dynamische programmering 3*, 1971.
MCS 1.8	J. KRIENS, F. GOBEL & W. MOLENAAR, *Leergang Besliskunde, deel 8: Minimaxmethode, netwerkplanning, simulatie*, 1968.
MCS 2.1	G.J.R. FORCH, P.J. VAN DER HOUWEN & R.P. VAN DE RIET, *Colloquium stabiliteit van differentieschema's, deel 1*, 1967.
MCS 2.2	L. DEKKER, T.J. DEKKER, P.J. VAN DER HOUWEN & M.N. SPIJKER, *Colloquium stabiliteit van differentieschema's, deel 2*, 1968.
MCS 3.1	H.A. LAUWERIER, *Randwaardeproblemen, deel 1*, 1967.
MCS 3.2	H.A. LAUWERIER, *Randwaardeproblemen, deel 2*, 1968.
MCS 3.3	H.A. LAUWERIER, *Randwaardeproblemen, deel 3*, 1968.
MCS 4	H.A. LAUWERIER, *Representaties van groepen*, 1968.
MCS 5	J.H. VAN LINT, J.J. SEIDEL, P.C. BAAYEN, *Colloquium discrete wiskunde*, 1968.
MCS 6	K.K. KOKSMA, *Cursus ALGOL 60*, 1969.
MCS 7.1	*Colloquium moderne rekenmachines, deel 1*, 1969.
MCS 7.2	*Colloquium moderne rekenmachines, deel 2*, 1969.
MCS 8	H. BAVINCK & J. GRASMAN, *Relaxatietrillingen*, 1969.
MCS 9.1	T.M.T. COOLEN, G.J.R. FORCH, E.M. DE JAGER & H.G.J. PIJLS, *Colloquium elliptische differentiaalvergelijkingen, deel 1*, 1969.
MCS 9.2	W.P. VAN DE BRINK, T.M.T. COOLEN, B. DIJKHUIS, P.P.N. DE GROEN, P.J. VAN DER HOUWEN, E.M. DE JAGER, N.M. TEMME & R.J. DE VOGELAERE, *Colloquium elliptische differentiaalvergelijkingen, deel 2*, 1970.
MCS 10	J. FABIUS & W.R. VAN ZWET, *Grondbegrippen van de waarschijnlijkheidsrekening*, 1970.

MCS 11 H. BART, M.A. KAASHOEK, H.G.J. PIJLS, W.J. DE SCHIPPER & J. DE VRIES, *Colloquium halfalgebra's en positieve operatoren*, 1971.

MCS 12 T.J. DEKKER, *Numerieke algebra*, 1971.

MCS 13 F.E.J. KRUSEMAN ARETZ, *Programmeren voor rekenautomaten; De MC ALGOL 60 vertaler voor de EL X8*, 1971.

MCS 14 H. BAVINCK, W. GAUTSCHI & G.M. WILLEMS, *Colloquium approximatie-theorie*, 1971.

MCS 15.1 T.J. DEKKER, P.W. HEMKER & P.J. VAN DER HOUWEN, *Colloquium stijve differentiaalvergelijkingen, deel 1*, 1972.

MCS 15.2 P.A. BEENTJES, K. DEKKER, H.C. HEMKER, S.P.N. VAN KAMPEN & G.M. WILLEMS, *Colloquium stijve differentiaalvergelijkingen, deel 2*, 1973.

* MCS 15.3 P.A. BEENTJES e.a., *Colloquium stijve differentiaalvergelijkingen, deel 3.*

MCS 16.1 L. GEURTS, *Cursus programmeren, deel 1: De elementen van het programmeren*, 1973.

MCS 16.2 L. GEURTS, *Cursus programmeren, deel 2: De programmeertaal ALGOL 60*, 1973.

MCS 17.1 P.S. STOBBE, *Lineaire algebra, deel 1*, 1974.

MCS 17.2 P.S. STOBBE, *Lineaire algebra, deel 2*, 1974.

MCS 18 F. VAN DER BLIJ, H. FREUDENTHAL, J.J. DE IONGH, J.J. SEIDEL & A. VAN WIJNGAARDEN, *Een kwart eeuw wiskunde 1946-1971, Syllabus van de Vakantiecursus 1971*, 1974.

MCS 19 A. HORDIJK, R. POTHARST & J. TH. RUNNENBURG, *Optimaal stoppen van Markovketens*, 1974.

* MCS 20 T.M.T. COOLEN, P.W. HEMKER, P.J. VAN DER HOUWEN & E. SLAGT, *ALGOL 60 procedures voor begin- en randwaardeproblemen.*

* MCS 21 L.J.M. GEURTS, D. GRUNE, Z. MANNA, L.G.L.T. MEERTENS & W.P. DE ROEVER, *Colloquium Programmacorrectheid.*

* MCS 22 R. HELMERS, J. OOSTERHOFF, F.H. RUYMGAART & M.C.A. VAN ZUYLEN, *Werkweek Statistiek* 1973.

* MCS 23.1 J. GRASMAN, J.H.B. KEMPERMAN, J.W. DE ROEVER & G.M. WILLEMS, *Colloquium Onderwerpen uit de Biomathematica, deel 1.*

* MCS 23.2 J. GRASMAN, J.H.B. KEMPERMAN, J.W. DE ROEVER & G.M. WILLEMS, *Colloquium Onderwerpen uit de Biomathematica, deel 2.*

MCS 24.1 P.J. VAN DER HOUWEN, *Numerieke integratie van differentiaalvergelijkingen, deel 1: Eenstapsmethoden*, 1974.

* MCS 25 *Colloquium Structuur Programmeertalen.*

De met een * gemerkte uitgaven moeten nog verschijnen.