

Stochastisch epidemisch modelleren met  
continue-tijd-Markovketens  
Bachelorscriptie Wiskunde & Toepassingen  
7.5 ECTS



**Universiteit Utrecht**

Fleur Slegers (5605407)  
Onder begeleiding van dhr. S. Dirksen

7 november 2020

# Inhoudsopgave

<b>1</b>	<b>Introductie</b>	<b>2</b>
<b>2</b>	<b>Continue-tijd-Markovketens</b>	<b>6</b>
2.1	Q-matrix . . . . .	9
2.2	Sprongmatrix . . . . .	10
2.3	Stochastische matrix $P(t)$ . . . . .	11
2.4	Definities van een CTMK . . . . .	17
<b>3</b>	<b>Het SIR-CTMK-model</b>	<b>22</b>
3.1	Toestandsruimte, Q-matrix en sprongmatrix . . . . .	22
3.2	Definities van een SIR-CTMK . . . . .	26
<b>4</b>	<b>Simulaties</b>	<b>29</b>
4.1	Gillespie-algoritme . . . . .	29
4.2	Resultaten . . . . .	30
<b>5</b>	<b>Deterministisch SIR-model</b>	<b>33</b>
<b>6</b>	<b>Het SIS-CTMK-model</b>	<b>35</b>
6.1	Simulaties . . . . .	38
<b>7</b>	<b>Conclusie</b>	<b>40</b>

# Hoofdstuk 1

## Introductie

Wanneer een besmettelijke ziekte uitbreekt, is er informatie nodig over hoe de ziekte-uitbraak zich zal ontwikkelen. Zo is het bijvoorbeeld belangrijk om een goede schatting te maken van het aantal besmette personen op een bepaald tijdstip, een voorspelling te kunnen doen over de effectiviteit van eventuele maatregelen, en de kans op een grote uitbraak van de ziekte. Er is sprake van een grote uitbraak wanneer het aantal besmettelijke individuen substantieel toeneemt tot een groot aantal [1].

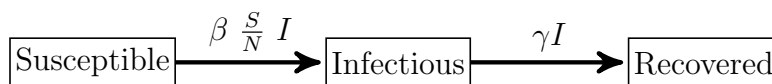
Een wiskundig epidemiologisch model dat vaak wordt toegepast om epidemieën van een ziekte die van mens op mens wordt overgedragen te modelleren, is het SIR (*S*uceptible, *I*nfectious, *R*ecovered)-model. Het SIR-model modelleert een epidemie van een infectieziekte in een grote populatie van constante grootte. De populatie wordt in 3 onderling disjuncte groepen verdeeld van vatbare (*susceptible*), besmettelijke (*infectious*) en genezen (*recovered*) individuen. Er ontstaan nieuwe besmettingen door interactie tussen vatbare en besmettelijke personen, en besmettelijke individuen kunnen van de ziekte genezen. Het SIR-model beschrijft de stroming van individuen tussen de 3 groepen [4].

De snelheid waarmee nieuwe besmettingen optreden wordt de infectiekracht genoemd. De infectiekracht is gelijk aan  $\beta \left( \frac{S(t)}{N} \right) I(t)$ , voor een positieve constante  $\beta$ , waarbij  $S(t)$  het aantal vatbare en  $I(t)$  het aantal besmettelijke individuen is op tijdstip  $t$ , en  $N$  de populatiegrootte.

De genezingsnelheid is de snelheid waarmee individuen uit de *infectious*-groep overgaan naar de *recovered*-groep. De genezingsnelheid is  $\gamma I(t)$ , voor een posi-

tieve constante  $\gamma$  [2].

In figuur 1.1 is een stroomdiagram weergegeven voor het SIR-model.



Figuur 1.1: Stroomdiagram tussen de 3 groepen van het SIR-model.

De infectiekracht en genezingsnelheid komen samen in het basaal reproductiegetal  $\mathbb{R}_0$ . Het basaal reproductiegetal is het verwachte aantal secundaire besmettingen dat wordt veroorzaakt doordat één besmet individu in een verder volledig vatbare populatie wordt geïntroduceerd [2].

**Definitie 1.0.1** (Basaal reproductiegetal).  $\mathbb{R}_0 = \beta \cdot \frac{1}{\gamma}$ .

Het basaal reproductiegetal is een belangrijke maat voor het analyseren van een epidemie. Wanneer  $\mathbb{R}_0$  groter is dan 1, is de infectiekracht groter dan de genezingsnelheid, en groeit de infectie binnen de populatie. Wanneer  $\mathbb{R}_0$  kleiner is dan 1, genezen individuën sneller dan zij anderen besmetten, en sterft de infectie uit [1].

Het epidemische SIR-model wordt vaak als eerste benadering toegepast tijdens een ziekte-uitbraak om schattingen te verkrijgen over de potentiële verspreiding van de ziekte. Zo is via het SIR-model nuttige informatie verkregen over de verspreiding van de Ebola-uitbraak in Afrika van 2014 [2]. Ook voor de pandemie waarin wij ons nu bevinden, blijkt het SIR-model nuttig te zijn. Zo construeerden [5] een epidemiologisch model voor de COVID-19-epidemie in Italië, met als basis het SIR-model.

[7] formuleerden het SIR-model oorspronkelijk als deterministisch model; een stel van gewone differentiaalvergelijkingen dat de veranderingen van het aantal individuën per groep beschrijft [6]. Voor een populatie van constante grootte

$N = S(t) + I(t) + R(t)$  geldt

$$\begin{aligned}\frac{dS}{dt} &= -\beta I \frac{S}{N}, \\ \frac{dI}{dt} &= \beta I \frac{S}{N} - \gamma I, \\ \frac{dR}{dt} &= \gamma I.\end{aligned}\tag{1.1}$$

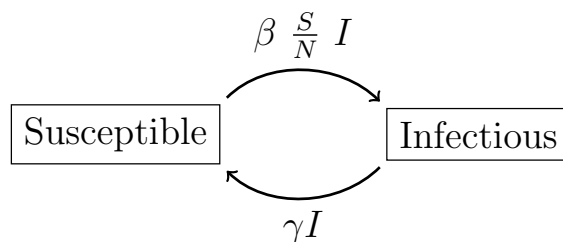
De stochastische variant van het SIR-model is een continue-tijd-Markovketen waarbij de transitiesnelheden van het proces gebaseerd zijn op 1.1. Voor het opstellen van het stochastische SIR-model worden twee kansvariabelen  $S$  en  $I$  gedefinieerd. Hoe de waarden van  $S$  en  $I$  veranderen, hangt af van de infectiekracht en genezingsnelheid van het proces. Wanneer het aantal besmettelijke individuen binnen een populatie klein is, of wanneer variabiliteit in infectiekracht, genezingsnelheid, geboortes en sterftes invloed heeft op het verloop van een epidemie, is een stochastische aanpak gewenst. Daarnaast is het voor multivariate processen moeilijk om een analytische oplossing te vinden voor stelsels van differentiaalvergelijkingen. Het is dan vaak eenvoudiger om de oplossing van het stelsel numeriek te benaderen [1].

Het SIR-model is slechts één voorbeeld van een wiskundig epidemiologisch model. Een model dat veel lijkt op het SIR-model, is het SIS (*Susceptible, Infectious, Susceptible*)-model [8]. In dit model bouwen individuen geen immuniteit op voor de infectie nadat zij besmet zijn geraakt. Dit houdt in dat genezen individuen terugkeren in de groep van vatbare individuen. De populatie wordt dus onderverdeeld in slechts 2 groepen; vatbaren en besmettelijken. Het stelsel van gewone differentiaalvergelijkingen dat het SIS-model kenmerkt, is

$$\begin{aligned}\frac{dS}{dt} &= -\beta I \frac{S}{N} + \gamma I, \\ \frac{dI}{dt} &= \beta I \frac{S}{N} - \gamma I.\end{aligned}\tag{1.2}$$

Het stroomdiagram dat bij 1.2 hoort, is weergegeven in figuur 1.2.

Wanneer individuen tijdelijk immuun zijn voor een ziekte nadat ze zijn genezen, kan het SIRS (*Susceptible, Infectious, Recovered, Susceptible*)-model worden toegepast [2]. Zo is er nog een aantal vergelijkbare epidemiologische modellen die berusten op een stelsel van gewone differentiaalvergelijkingen.



Figuur 1.2: Stroomdiagram tussen de 2 groepen van het SIS-model.

In deze scriptie laten we zien hoe stochastische processen voor het SIR- en SIS-model opgesteld kunnen worden vanuit het deterministische raamwerk van de modellen. De stochastische modellen kunnen worden gebruikt om voorspellingen te doen over het aantal vatbare en besmettelijke individuën op een bepaald tijdstip en over de kans op een grote uitbraak.

De structuur van deze scriptie is als volgt: in Hoofdstuk 2 wordt de theorie achter continue-tijd-Markovketens besproken. Er worden 3 manieren om een continue-tijd-Markovproces te definiëren gegeven. In Hoofdstuk 3 laten we zien hoe het SIR-model als toepassing van een continue-tijd-Markovketen geïnterpreteerd kan worden. We geven de 3 definities van een SIR-continue-tijd-Markovproces. In Hoofdstuk 4 laten we zien hoe het stochastische SIR-model gesimuleerd kan worden, en bespreken we de invloed van de parameters van het SIR-model op de simulaties. We laten in Hoofdstuk 5 zien hoe de stochastische interpretatie van het SIR-model zich verhoudt tot de deterministische variant van het model. In Hoofdstuk 6 behandelen we het SIS-CTMK-model.

# Hoofdstuk 2

## Continue-tijd-Markovketens

Een continue-tijd-Markovketen (CTMK) is een rechts-continu stochastisch proces  $(X_t)_{t \geq 0}$  met een aftelbare toestandsruimte en continue tijdsparameter  $t$  dat voldoet aan de Markov-eigenschap [9]. Dit houdt in dat het gedrag van het proces in de toekomst alleen afhangt van de huidige toestand van het proces, en niet van het verleden. We geven nu een formele definitie van een CTMK. Definities uit deze sectie zijn verkregen uit sectie 2.2 van [9].

**Definitie 2.0.1** (Continue-tijd-Markovketen). *Een proces  $(X_t)_{t \geq 0}$  is een continue-tijd-Markovketen als voor alle  $s, t \geq 0$  en alle elementen  $i, j, x(u) \in I$  waarvoor  $0 \leq u < s$  geldt dat*

$$\mathbb{P}(X_{t+s} = j | X_s = i, X_u = x_u, 0 \leq u < s) = \mathbb{P}(X_{t+s} = j | X_s = i).$$

Als bovendien geldt dat

$$\mathbb{P}(X_{t+s} = j | X_s = i)$$

onafhankelijk is van  $s$ , dan heeft het proces *stationaire transitiekansen*.

Elk pad  $t \mapsto X_t(\omega)$  van een CTMK maakt een aftelbaar aantal sprongen tussen toestanden. De tijden  $J_0, J_1, J_2, \dots$  waarop zo'n sprong wordt gemaakt, zijn de sprongtijden van  $(X_t)_{t \geq 0}$ .

**Definitie 2.0.2** (Sprongtijden).

$$J_0 = 0, J_{n+1} = \inf\{t \geq J_n | X_t \neq X_{J_n}\}, n \geq 0.$$

Tussen twee sprongen in bevindt het proces zich voor enige tijd in dezelfde toestand. De tijdsduren dat het proces zich in dezelfde toestand bevindt, oftewel de tijden tot de volgende sprong wordt gemaakt, zijn de wachttijden  $S_1, S_2, S_3, \dots$  van  $(X_t)_{t \geq 0}$ .

**Definitie 2.0.3** (Wachttijden).

$$S_n = \begin{cases} J_n - J_{n-1} & \text{als } J_{n-1} < \infty \\ \infty & \text{anders} \end{cases} \quad (2.1)$$

voor  $n \geq 1$ .

**Definitie 2.0.4** (Sprongketen). *De sprongketen van het proces  $(X_t)_{t \geq 0}$  is de keten  $(Y_n)_{n \geq 0}$  gedefinieerd door*

$$Y_n = X_{J_n}, \quad n \geq 0.$$

De waarden van de sprongketen veranderen niet continu in de tijd, maar in stapjes. Daarom is de sprongketen van een CTMK een *discrete-tijd-Markovketen*. In sectie 4.1 uit [10] wordt de volgende definitie van een discrete-tijd-Markovketen gegeven.

**Definitie 2.0.5** (Discrete-tijd-Markovketen). *Een discrete-tijd-Markovketen is een stochastisch proces  $\{X_n, n = 0, 1, 2, \dots\}$  waarvoor voor alle  $i, j \in I$  en alle  $n \geq 0$  geldt dat*

$$\mathbb{P}(X_{n+1} = j \mid X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i_n) = \mathbb{P}(X_{n+1} = j \mid X_n = i).$$

De volgende stelling over de wachttijden van een CTMK komt uit sectie 6.2 van [10].

**Stelling 2.0.1.** *De wachttijden  $(S_n)_{n=1,2,\dots}$  van een continue-tijd-Markovketen zijn exponentieel verdeeld.*

*Bewijs.* Stel dat een continue-tijd-Markovketen  $(X_t)_{t \geq 0}$  op tijdstip  $\tau$  naar toestand  $i$  transitioneert. Zij  $S$  de tijd voor de eerstvolgende sprong van het proces. Stel dat het proces toestand  $i$  niet verlaat tussen  $\tau$  en  $\tau + s$  voor een  $s \geq 0$ , oftewel  $S > s$ . De kans dat het proces de toestand niet verlaat tussen  $\tau + s$  en  $\tau + s + t$  voor een  $t \geq 0$  is, vanwege de Markov-eigenschap van het proces, onafhankelijk van het verleden, oftewel

$$\mathbb{P}(S > s + t \mid S > s) = \mathbb{P}(S > t).$$



Stel  $g(t) = \mathbb{P}(S > t)$  zodat

$$g(s+t) = g(s)g(t), \text{ voor alle } s, t \geq 0. \quad (2.2)$$

Omdat het proces  $(X_t)_{t \geq 0}$  rechts-continu is, kunnen we aannemen dat  $S > 0$ . Dus er is een  $n > 0$  zodat  $g\left(\frac{1}{n}\right) = \mathbb{P}\left(S > \frac{1}{n}\right) > 0$ . Herhaaldelijk toepassen van 2.2 geeft

$$g(1) = g\left(\frac{1}{n} + \dots + \frac{1}{n}\right) = g^n\left(\frac{1}{n}\right) > 0,$$

dus  $g(1) = e^{-\lambda}$  voor een  $0 \leq \lambda < \infty$ . Zij  $p, q \in \mathbb{N}_0$ . We hebben dan op dezelfde manier

$$g(1) = g^q\left(\frac{1}{q}\right),$$

waaruit volgt dat

$$g\left(\frac{p}{q}\right) = g^{p/q}(1).$$

Dus voor alle rationale getallen  $r > 0$  geldt  $g(r) = g(1)^r > 0$ , dus  $g(r) = e^{-\lambda r}$ . Merk op dat  $g$  een dalende functie is. Voor  $t \geq 0$ , kies  $r, s > 0$  zodat  $r \leq t \leq s$ . We hebben

$$e^{-r} = g(r) \geq g(t) \geq g(s) = e^{-\lambda s}.$$

Omdat we  $s$  en  $r$  willekeurig dichtbij  $t$  kunnen kiezen, kunnen we concluderen dat  $g(t) = e^{-\lambda t}$ , dus  $S$  is exponentieel verdeeld met parameter  $\lambda$ , oftewel  $S \sim E(\lambda)$ .  $\square$

Als een proces  $(X_t)_{t \geq 0}$  een eindige toestandsruimte heeft, wordt de kans op een gebeurtenis volledig bepaald door de gezamenlijke distributie van de wachttijden en sprongketen van het proces [9]. De kans  $\mathbb{P}(X_t = i, i \in I)$  wordt in dit geval gegeven door

$$\mathbb{P}(X_t = i) = \sum_{n=0}^{\infty} \mathbb{P}(Y_n = i, J_n \leq t < J_{n+1})$$

en

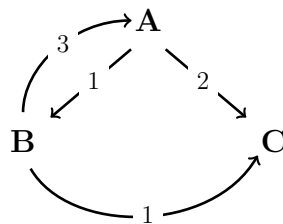
$$\mathbb{P}(X_t = i \text{ voor een } t \in [0, \infty)) = \mathbb{P}(Y_n = i \text{ voor een } n \geq 0) [9].$$

Een proces met een eindige toestandsruimte kan dus volledig geconstrueerd worden aan de hand van zijn wachttijden en sprongketen. Continue-tijd-Markovketens kunnen echter op verschillende manieren worden gedefinieerd, ook aan de hand

van de de transitiesnelheden of transitiekansen van het proces. In sectie 2.4 geven we 3 formele definities voor een CTMK, en laten we zien dat deze equivalent zijn. In de volgende secties bespreken we de theorie die nodig is om tot deze definities te komen.

## 2.1 Q-matrix

De snelheid waarmee een CTMK tussen toestanden overgaat, kan worden beschreven aan de hand van een transitiediagram. In een transitiediagram zijn de toestanden uit de toestandruimte van het proces verbonden op basis van transitiesnelheden. Een pijl van toestand  $i$  naar  $j$  heeft als waarde de snelheid waarmee het proces van toestand  $i$  naar toestand  $j$  overgaat [9]. In figuur 2.1 is een simpel voorbeeld van een transitiediagram voor een continue-tijd-Markovketen met toestandruimte  $I = \{A, B, C\}$  weergegeven.



Figuur 2.1: Voorbeeld van een transitiediagram.

De informatie uit het transitiediagram kan ook worden genoteerd in een zogeheten *Q-matrix*. De volgende definitie komt uit sectie 2.1 van [9].

**Definitie 2.1.1** (Q-matrix). *Zij  $I$  een aftelbare verzameling. Een Q-matrix op  $I$  is een matrix  $Q = (q_{ij} \mid i, j \in I)$  die voldoet aan:*

- (i)  $0 \leq -q_{ii} < \infty$  voor alle  $i$ ;
- (ii)  $q_{ij} \geq 0$  voor alle  $i \neq j$ ;
- (iii)  $\sum_{j \in I} q_{ij} = 0$  voor alle  $i$ .

Een element  $q_{ij}, i \neq j$  van een Q-matrix representeert de snelheid waarmee het proces van toestand  $i$  naar  $j$  gaat. Voorwaarde (ii) houdt dus in dat alle transitie-snelheden niet-negatief zijn.

Merk op dat een transitiediagram geen waarde toekent aan de transitie van een toestand naar zichzelf. Echter, voorwaarde (iii) houdt in dat in een Q-matrix de diagonaalelementen  $q_{ii}$  zo worden gedefinieerd dat

$$q_{ii} = - \sum_{j \neq i} q_{ij}.$$

Een transitiediagram correspondeert met een unieke Q-matrix. Het transitiediagram uit 2.1 correspondeert met de volgende Q-matrix:

$$Q = \begin{pmatrix} -3 & 1 & 2 \\ 3 & -4 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.3)$$

We weten nu iets over de snelheden waarmee een CTMK tussen toestanden transitioneert, maar we zijn op zoek naar een stochastische matrix die informatie bevat over de transitiekansen van het proces. De volgende definitie van een stochastische matrix is verkregen uit sectie 2.1 van [9].

**Definitie 2.1.2** (Stochastische matrix). *Een matrix  $P = (p_{ij} | i, j \in I)$  is stochastisch wanneer deze voldoet aan*

- (i)  $0 \leq p_{ij} < \infty$  voor alle  $i, j$
- (ii)  $\sum_{j \in I} p_{ij} = 1$  voor alle  $i$ .

Een stochastische matrix geeft voor alle  $i, j \in I$  de kansen dat een proces vanuit toestand  $i$  naar toestand  $j$  transitioneert. Uit een Q-matrix kan een zogeheten *sprongmatrix* geconstrueerd worden. In de volgende sectie zullen we zien dat een sprongmatrix een stochastische matrix is.

## 2.2 Sprongmatrix

Sectie 2.6 van [9] bevat de volgende definitie van een sprongmatrix.

**Definitie 2.2.1** (Sprongmatrix). De sprongmatrix  $\Pi = (\pi_{ij} | i, j \in I)$  behorend tot de  $Q$ -matrix  $Q$  wordt gedefinieerd door

$$\pi_{ij} = \begin{cases} q_{ij}/-q_{ii} & \text{als } j \neq i \text{ en } -q_{ii} \neq 0; \\ 0 & \text{als } j \neq i \text{ en } -q_{ii} = 0. \end{cases}$$

$$\pi_{ii} = \begin{cases} 0 & \text{als } -q_{ii} \neq 0; \\ 1 & \text{als } -q_{ii} = 0. \end{cases}$$

Voor elke rij in  $Q$  bekijken we de niet-diagonale elementen en schalen deze zodat ze optellen tot 1. Wanneer alle niet-diagonale elementen 0 zijn, is dit niet mogelijk. In dit geval krijgt het diagonaalelement de waarde 1. Anders krijgt het diagonaalelement de waarde 0.

$\Pi$  is een stochastische matrix. Immers, voor alle  $i \in I$  geldt dat

$$\sum_{j \in I} \pi_{ij} = 1,$$

en voor alle  $i, j \in I$  geldt dat  $0 \leq \pi_{ij} < \infty$ .

De matrix  $\Pi$  die bij transitiediagram 2.1 en  $Q$ -matrix 2.3 hoort, is

$$\Pi = \begin{pmatrix} 0 & 1/3 & 2/3 \\ 3/4 & 0 & 1/4 \\ 0 & 0 & 1 \end{pmatrix}. \quad (2.4)$$

## 2.3 Stochastische matrix $P(t)$

In deze sectie bekijken we de matrix  $P(t) = e^{tQ}$ . We zullen laten zien dat  $P(t)$  een stochastische matrix is voor alle  $t$ , en dat  $P(t)$  de unieke oplossing is van de voorwaartse vergelijking

$$\frac{d}{dt}P(t) = P(t)Q, \quad P(0) = I,$$

en van de achterwaartse vergelijking

$$\frac{d}{dt}P(t) = QP(t), \quad P(0) = I.$$

De volgende stelling komt overeen met stelling 2.1.1 van [9]

**Stelling 2.3.1.** *Zij  $I$  een eindige verzameling met  $|I|$  elementen. Zij  $Q \in \mathbb{R}^{|I| \times |I|}$ . Stel  $P(t) = e^{tQ}$ . Dan voldoet  $(P(t) \mid t \geq 0)$  aan de volgende eigenschappen:*

(i)  $P(s+t) = P(s)P(t)$  voor alle  $s, t$ ;

(ii)  $(P(t) \mid t \geq 0)$  is de unieke oplossing van de voorwaartse vergelijking

$$\frac{d}{dt}P(t) = P(t)Q, \quad P(0) = I; \quad (2.5)$$

(iii)  $(P(t) \mid t \geq 0)$  is de unieke oplossing van de achterwaartse vergelijking

$$\frac{d}{dt}P(t) = QP(t), \quad P(0) = I;$$

(iv) Voor  $k = 0, 1, 2, \dots$ , geldt

$$\left( \frac{d}{dt} \right)^k \Big|_{t=0} P(t) = Q^k.$$

Het bewijs van stelling 2.3.1 volgt dat van [9].

*Bewijs.*

(i) Voor alle  $s, t \in \mathbb{R}$  geldt dat  $sQ$  en  $tQ$  commuteren, dus

$$e^{sQ+tQ} = e^{sQ}e^{tQ},$$

dus

$$P(s+t) = e^{(s+t)Q} = e^{sQ+tQ} = e^{sQ}e^{tQ} = P(s)P(t).$$

(ii) Er geldt

$$P(t) = e^{tQ} = \sum_{k=0}^{\infty} \frac{(tQ)^k}{k!}. \quad (2.6)$$

We laten nu zien dat 2.6 een oneindige convergentieradius heeft, naar voorbeeld van Appendix 2.10 van [9]. Hiervoor gebruiken we het gegeven dat  $I$  eindig is. Zij  $N = |I|$ . Dan heeft  $Q$  dimensie  $N \times N$ . We passen de volgende twee matrixnormen toe op  $Q$ :

$$|Q| = \sup_{v \neq 0} \left\{ \frac{|Qv|}{|v|} \right\}, \quad (2.7)$$

$$\|Q\|_\infty = \sup_{i,j} |q_{ij}|, \quad (2.8)$$

waarbij  $v$  een vector in  $\mathbb{R}^N$  is. Merk op dat  $q_{ij} < \infty$  voor alle  $i, j \in I$ , dus  $\|Q\|_\infty < \infty$ .

Voor alle vectoren  $w$  geldt

$$|Q||w| = \sup_{v \neq 0} \left\{ \frac{|Qv|}{|v|} \right\} |w| \geq \frac{|Qw|}{|w|} |w| = |Qw|.$$

In het bijzonder geldt voor de eenheidsvectoren  $e_i, 1 \leq i \leq N$  met norm 1 dat

$$|Qe_j| \leq |Q|.$$

Er is een  $j \in I$  zodat  $|q_{jj}| = \sup_{i,j \in I} |q_{ij}|$  voor een  $j \in I$ . Hieruit volgt dat

$$\|Q\|_\infty^2 = |q_{jj}|^2 \leq \sum_{i \in I} q_{ij}^2 = |Qe_j|^2 \leq |Q|^2.$$

Aan de andere kant hebben we ook

$$\begin{aligned} |Qv|^2 &= \left( \sum_{i \in I} \sum_{j \in I} q_{ij} v_j \right)^2 \\ &\leq \sum_{i \in I} \left( \sum_{j \in I} \sup_{i,j \in I} |q_{ij}| |v_j| \right)^2 \\ &= N \|Q\|_\infty^2 \left( \sum_{j \in I} |v_j| \right)^2. \end{aligned}$$

Toepassen van de ongelijkheid van Cauchy-Schwarz geeft ons nu

$$|Qv|^2 \leq N^2 \|Q\|_\infty^2 |v|^2,$$

waaruit volgt dat  $|Q| \leq N \|Q\|_\infty$ . Dus

$$\|Q\|_\infty \leq |Q| \leq N \|Q\|_\infty^2 \leq \infty.$$

Voor alle  $N \times N$  matrices  $M_1, M_2$  en alle vectoren  $v \in \mathbb{R}^N$  geldt

$$|(M_1 + M_2)v| \leq |M_1v| + |M_2v| \leq (|M_1| + |M_2|)|v|$$

en

$$|M_1M_2v| \leq |M_1||M_2v| \leq |M_1||M_2||v|.$$

We bekijken nu voor  $n = 0, 1, 2, \dots$  de eindige som

$$E(n) = \sum_{k=0}^n \frac{Q^k}{k!}.$$

Voor alle  $m, n$  met  $m \leq n$  en alle  $i, j \in I$  geldt

$$\begin{aligned} |(E(n) - E(m))_{ij}| &\leq \|E(n) - E(m)\|_\infty \\ &\leq |E(n) - E(m)| \\ &= \left| \sum_{k=m+1}^n \frac{Q^k}{k!} \right| \\ &\leq \sum_{k=m+1}^n \frac{|Q|^k}{k!}. \end{aligned}$$

Omdat  $|Q| \leq \infty$ , convergeert  $\sum_{k=0}^{\infty} \frac{|Q|^k}{k!}$  volgens de ratiotest. Dus als  $m, n \rightarrow \infty$ , dan

$$\sum_{k=m+1}^n \frac{|Q|^k}{k!} \rightarrow 0.$$

Hieruit kunnen we concluderen dat elk component van  $E(n)$  een cauchy-reeks vormt. Dus elk component van  $E(n)$  convergeert. Hieruit kunnen we concluderen dat

$$e^Q = \sum_{k=0}^{\infty} \frac{Q^k}{k!}$$

goed gedefinieerd is en de machtreeks

$$(e^{tQ})_{ij} = \frac{(tQ_{ij})^k}{k!}$$

een oneindige convergentieradius heeft voor alle  $i, j \in I$ . Hieruit volgt dat elk component van  $e^{tQ}$  differentieerbaar is. De afgeleide van de machtreeks is de reeks van afgeleiden van de termen:

$$P'(t) = \sum_{k=1}^{\infty} \frac{t^{k-1}Q^k}{(k-1)!} = P(t)Q = QP(t).$$

Dus  $P(t)$  voldoet aan de voor- en achterwaartse vergelijkingen.

We laten nu zien dat  $P(t)$  de unieke oplossing van de voor- en achterwaartse vergelijkingen is. Stel dat de matrix  $M(t)$  ook voldoet aan de voorwaartse vergelijking, dus

$$\frac{d}{dt}M(t) = M(t)Q.$$

Er geldt dan, volgens de kettingregel voor differentiëren, dat

$$\begin{aligned} \frac{d}{dt}(M(t)e^{-tQ}) &= \left(\frac{d}{dt}M(t)\right)e^{-tQ} + M(t)\left(\frac{d}{dt}e^{-tQ}\right) \\ &= M(t)Qe^{-tQ} - M(t)Qe^{-tQ} \\ &= 0. \end{aligned}$$

Dus  $M(t)e^{-tQ} = \frac{M(t)}{P(t)}$  is constant, dus  $M(t) = P(t)$ . Hieruit concluderen we dat  $P(t)$  de unieke oplossing is van de voorwaartse vergelijking. Eenzelfde bewijs kan gegeven worden voor de uitspraak dat  $P(t)$  de unieke oplossing is van de achterwaartse vergelijking. Hiermee zijn (ii) en (iii) aangetoond.

□

Aan de hand van stelling 2.1.2 uit [9] laten we zien dat de matrix  $P(t) = e^{tQ}$ , met  $Q$  een Q-matrix, een stochastische matrix is. Het bewijs voor deze stelling volgt dat van [9].

**Stelling 2.3.2.** *Als  $I$  een eindige verzameling is, en  $Q \in \mathbb{R}^{|I| \times |I|}$  een Q-matrix is, dan is  $P(t) = e^{tQ}$  een stochastische matrix voor alle  $t \geq 0$ .*



*Bewijs.* Stel dat een matrix  $Q$  op een eindige verzameling  $I$  een Q-matrix is. Zij  $P(t) = e^{tQ}$ . Per definitie is

$$e^{tQ} = \sum_{k=0}^{\infty} \frac{(tQ)^k}{k!},$$

dus we kunnen  $P(t)$  schrijven als

$$P(t) = \sum_{k=0}^{\infty} \frac{(tQ)^k}{k!} = I + tQ + \frac{(tQ)^2}{2} + \dots \quad (2.9)$$

Voor  $t \downarrow 0$  geldt

$$P(t) = I + tQ + \mathcal{O}(t^2)$$

Uit het gegeven dat  $Q$  een Q-matrix is, volgt dat voor alle elementen  $(i, j)$  van  $Q$  geldt dat  $q_{ij} \geq 0$ . Voor voldoende kleine  $t \geq 0$  geldt ook  $1 + tq_{ii} \geq 0$ . Dus, voor  $t \geq 0$  voldoende klein, geldt

$$0 \leq p_{ij}(t) < \infty \text{ voor alle } i, j \in I. \quad (2.10)$$

Uit 2.9 volgt dat

$$p_{ij}(t) = \begin{cases} \sum_{n=1}^{\infty} \frac{t^n (q_{ij})^n}{n!} & \text{als } i \neq j; \\ 1 + \sum_{n=1}^{\infty} \frac{t^n (q_{ii})^n}{n!} & \text{als } i = j. \end{cases} \quad (2.11)$$

We bekijken nu de rijssommen van de matrix  $P$ . Uit 2.11 volgt dat voor alle  $i \in I$  de bijbehorende rijssom als volgt kan worden geschreven:

$$\begin{aligned} \sum_{j \in I} p_{ij}(t) &= 1 + \sum_{n=1}^{\infty} \frac{(t q_{ii})^n}{n!} + \sum_{j \in I, j \neq i} \sum_{n=1}^{\infty} \frac{(t q_{ij})^{(n)}}{n!} \\ &= 1 + \sum_{j \in I} \sum_{n=1}^{\infty} \frac{(t q_{ij})^{(n)}}{n!} \\ &= 1 + \sum_{n=1}^{\infty} \frac{t^n}{n!} \sum_{j \in I} (q_{ij})^{(n)} \end{aligned}$$

Merk op dat

$$\begin{aligned}
 \sum_{j \in I} (q_{ij})^{(n)} &= \sum_{j \in I} \sum_{k \in I} (q_{ik})^{n-1} q_{kj} \\
 &= \sum_{k \in I} (q_{ik})^{n-1} \sum_{j \in I} q_{kj} \\
 &= 0 \quad (\text{def 2.1.1}(iii)),
 \end{aligned}$$

dus voor alle  $i \in I$  geldt

$$\sum_{j \in I} p_{ij}(t) = 1. \tag{2.12}$$

Uit 2.10 en 2.12 volgt stelling 2.3.2. □

## 2.4 Definities van een CTMK

In deze sectie geven we 3 verschillende definities van een CTMK-model, aan de hand van de Q-matrix, sprongmatrix en stochastische matrix  $P(t) = e^{tQ}$ .

Zij  $(Y_n)_{n \geq 0}$  een discrete-tijd-Markovketen met initiële distributie  $\delta_{(s_0, i_0)}$  en transitie matrix  $\Pi$ . Zij  $T_1, T_2, \dots$  onafhankelijke exponentiële kansvariabelen met parameter 1, onafhankelijk van  $(Y_n)_{n \geq 0}$ . Definieer

$$\begin{aligned}
 S_n &= \frac{T_n}{q(Y_{n-1})}, \\
 J_n &= S_1 + \dots + S_n
 \end{aligned}$$

en

$$X_t = \begin{cases} Y_n & \text{als } J_n \leq t < J_{n+1} \text{ voor een } n \\ \infty & \text{anders.} \end{cases}$$

Dan geldt voor alle  $n \geq 1$ , conditioneel op  $Y_0, \dots, Y_{n-1}$ , dat de wachttijden  $S_1, \dots, S_n$  onafhankelijke exponentiële kansvariabelen zijn met parameters  $q(Y_0), \dots, q(Y_{n-1})$  respectievelijk [9]. We zeggen dan dat  $(X_t)_{t \geq 0}$  *Markov*( $\delta_{(s_0, i_0)}, Q$ ) is. Dit noemen we de sprongketen-en-wachttijddefinitie van  $(X_t)_{t \geq 0}$ .

De volgende stelling komt overeen met stelling 2.8.2 van [9].

**Stelling 2.4.1** (Hoofdstelling). *Zij  $(X_t)_{t \geq 0}$  een rechts-continu proces met een eindige toestandsruimte  $I$  en initiële toestand  $X_0 = (s_0, i_0)$ . Zij  $Q$  een  $Q$ -matrix op  $I$  met bijbehorende sprongmatrix  $\Pi$ . Dan zijn de volgende 3 voorwaarden equivalent:*

(i) (Sprongketen-en-wachttijddefinitie)

$(X_t)_{t \geq 0}$  is Markov( $\delta_{(s_0, i_0)}, Q$ ).

(ii) (Infinitesimale definitie)

Voor alle  $t, h \geq 0$  geldt dat  $X_{t+h}$ , conditioneel op  $X_t = i$ , onafhankelijk is van  $(X_s)_{s \leq t}$ . Bovendien geldt voor alle  $j \in I$  dat als  $h \downarrow 0$  uniform in  $t$ , dan

$$\mathbb{P}(X_{t+h} = j \mid X_t = i) = \delta_{ij} + q_{ij}h + o(h);$$

(iii) (Transitiekansendefinitie)

Voor alle  $n = 0, 1, 2, \dots$ , alle tijden  $0 \leq t_1 \leq t_2 \leq \dots \leq t_{n+1}$  en alle toestanden  $i_0, \dots, i_{n+1}$ , geldt

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} \mid X_{t_0} = i_0, \dots, X_{t_n} = i_n) = p_{i_n i_{n+1}}(t_{n+1} - t_n)$$

waarbij  $(p_{ij} \mid i, j \in I, t \geq 0)$  de unieke oplossing is van de voorwaartse vergelijking

$$P'(t) = P(t)Q, \quad P(0) = I. \tag{2.13}$$

Als  $(X_t)_{t \geq 0}$  aan één van bovenstaande voorwaarden voldoet, wordt  $(X_t)_{t \geq 0}$  een Markovketen met generatormatrix  $Q$  en initiële toestand  $(s_0, i_0)$  genoemd.

We zullen nu stelling 2.4.1 bewijzen. Hiervoor gebruiken we de notatie  $q_i = -q_{ii}$ . Het bewijs volgt het bewijs van [9].

*Bewijs.*

(i)  $\Rightarrow$  (ii) Stel dat (i) waar is. Dan  $X_0 = i$ . Als  $h \downarrow 0$ , dan geldt

$$\begin{aligned} \mathbb{P}_i(X_h = i) &= \mathbb{P}_i(J_1 > h) + \mathbb{P}_i(J_1 < h, X_h = i) \\ &\geq \mathbb{P}_i(J_1 > h) \\ &= 1 - \mathbb{P}_i(S_1 \leq h) \\ &= e^{q_i h} \\ &= 1 - q_i h + o(h). \end{aligned}$$

Voor  $j \neq i$  geldt

$$\begin{aligned}\mathbb{P}_i(X_h = j) &\geq \mathbb{P}_i(J_1 \leq h, Y_1 = j, S_2 > h) \\ &= (1 - e^{-q_i h}) \cdot \pi_{ij} \cdot e^{-q_j h} \\ &= (q_i h) (1 - q_j h) \pi_{ij} + o(h)\end{aligned}$$

Stel  $q_i \neq 0$ . Dan geldt dat  $\pi_{ij} = q_{ij}/q_i$ . We krijgen dan

$$\begin{aligned}\mathbb{P}_i(X_h = j) &\geq (q_i h - q_i q_j h^2) \frac{q_{ij}}{q_i} + o(h) \\ &= q_{ij} h + o(h)\end{aligned}$$

Stel dat  $q_i = 0$ . Dan geldt dat  $\pi_{ij} = 0$ . Hieruit volgt dat

$$\mathbb{P}_i(X_h = j) = 0.$$

Voor alle toestanden  $j \in I$  geldt dus

$$\mathbb{P}_i(X_h = j) \geq \delta_{ij} + q_{ij} h + o(h),$$

waarbij

$$\delta_{ij} = \begin{cases} 1 & \text{als } i = j \\ 0 & \text{als } i \neq j. \end{cases}$$

Er moet gelden dat

$$\sum_{j \in I} \mathbb{P}_i(X_h = j) = 1,$$

waaruit volgt dat

$$\mathbb{P}_i(X_h = j) = \delta_{ij} + q_{ij} h + o(h)$$

voor alle  $j \in I$ .

Volgens de Markov-eigenschap geldt voor alle  $t, h \geq 0$ , conditioneel op  $X_t = i$ , dat  $X_{t+h}$  onafhankelijk is van  $(X_s | s \leq t)$ , en als  $h \downarrow 0$  uniform in  $t$ , dan

$$\mathbb{P}(X_{t+h} = j | X_t = i) = \mathbb{P}_i(X_h = j) = \delta_{ij} + q_{ij} h + o(h).$$

(ii)  $\Rightarrow$  (iii) Zij  $P(t)$  de matrix zodat

$$p_{ij}(t) = \mathbb{P}_i(X_t = j) = \mathbb{P}(X_t = j | X_0 = i). \quad (2.14)$$

Als (ii) waar is, dan geldt

$$\begin{aligned} p_{ij}(t+h) &= \mathbb{P}_i(X_{t+h} = j) \\ &= \sum_{k \in I} \mathbb{P}(X_{t+h} = j | X_t = k) \mathbb{P}(X_t = k | X_0 = i) \\ &= \sum_{k \in I} \mathbb{P}(X_{t+h} = j | X_t = k) p_{ik}(t) \\ &= \sum_{k \in I} (\delta_{kj} + q_{kj}h + o(h)) p_{ik}(t) \\ &= p_{ij}(t) + \sum_{k \in I} (q_{kj} p_{ik}(t) h + o(h)). \end{aligned}$$

Hieruit volgt dat

$$\frac{p_{ij}(t+h) - p_{ij}(t)}{h} = \sum_{k \in I} (q_{kj} p_{ik}(t)) + \mathcal{O}(h),$$

dus  $p_{ij}(t)$  is rechts-differentieerbaar.

Op dezelfde manier zijn we in dat

$$\begin{aligned} p_{ij}(t) &= \sum_{k \in I} \mathbb{P}(X_t = j | X_t = k) \mathbb{P}(X_{t-h} = k | X_0 = i) \\ &= p_{ij}(t-h) + \sum_{k \in I} q_{kj} p_{ik}(t) h + o(h), \end{aligned}$$

waaruit volgt dat

$$\frac{p_{ij}(t-h) - p_{ij}(t)}{h} = \sum_{k \in I} q_{kj} p_{ik}(t) + \mathcal{O}(h).$$

Dus  $p_{ij}(t)$  is links-differentieerbaar.

We kunnen nu concluderen dat  $p_{ij}(t)$  differentieerbaar is en voldoet aan de voorwaartse vergelijking

$$p_{ij}'(t) = \sum_{k \in I} p_{ij}(t) q_{kj}, \quad p_{ij}(0) = \delta_{ij}. \quad (2.15)$$

We weten dat  $I$  een eindig aantal elementen bevat. Uit stelling 2.3.1 volgt dat  $p_{ij}(t)$  de unieke oplossing is die voldoet aan 2.15.

Uit (ii) volgt dat

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \dots, X_{t_n} = i_n) = \mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n).$$

Bovenstaande vergelijking geldt voor alle  $(X_{t_{n+1}})_{t \geq 0}$ .

Uit 2.14 volgt dat voor alle  $n \geq 0, 0 \leq t_1 \leq \dots \leq t_{n+1}$  en toestanden  $i_0, \dots, i_{n+1}$

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_t = i_n) = p_{i_n i_{n+1}}(t_{n+1} - t_n),$$

waarmee (ii)  $\Rightarrow$  (iii) is aangetoond.

(iii)  $\Rightarrow$  (i) Er is een proces  $(X_t)_{t \geq 0}$  dat voldoet aan de omschrijving (i). We hebben net laten zien dat dit proces dan ook aan (iii) voldoet. Dus er is een proces  $(X_t)_{t \geq 0}$  dat zowel aan (i) als aan (iii) voldoet.

Stel nu dat (iii)  $\not\Rightarrow$  (i). Dan bestaat er een proces  $(X_t)_{t \geq 0}$  dat wel aan (iii) voldoet, maar niet aan (i). Echter, voorwaarde (iii) bepaalt volledig de eindig-dimensionale distributies van het proces en daarmee ook zijn wachttijden en sprongketen. Dus er bestaat geen proces dat aan (iii) maar niet aan (i) voldoet.

We kunnen nu concluderen dat (iii)  $\Rightarrow$  (i).

□

# Hoofdstuk 3

## Het SIR-CTMK-model

### 3.1 Toestandsruimte, Q-matrix en sprongmatrix

Een CTMK zou een geschikte methode kunnen zijn om het deterministische SIR-model te representeren. In dit hoofdstuk stellen we een stochastisch proces samen dat berust op 1.1. We zullen zien dat de toestandsruimte van het proces aftelbaar is. De Markov-eigenschap van het proces is een redelijke aanname. De snelheid waarmee nieuwe besmettingen plaatsvinden is alleen afhankelijk van het aantal vatbaren en besmettelijken in de populatie op dat moment, dus geheugen. Of de ziekteduur per persoon exponentieel verdeeld is, zou onderzocht kunnen worden. We zullen 3 definities geven voor het SIR-CTMK-model zoals in stelling 2.4.1.

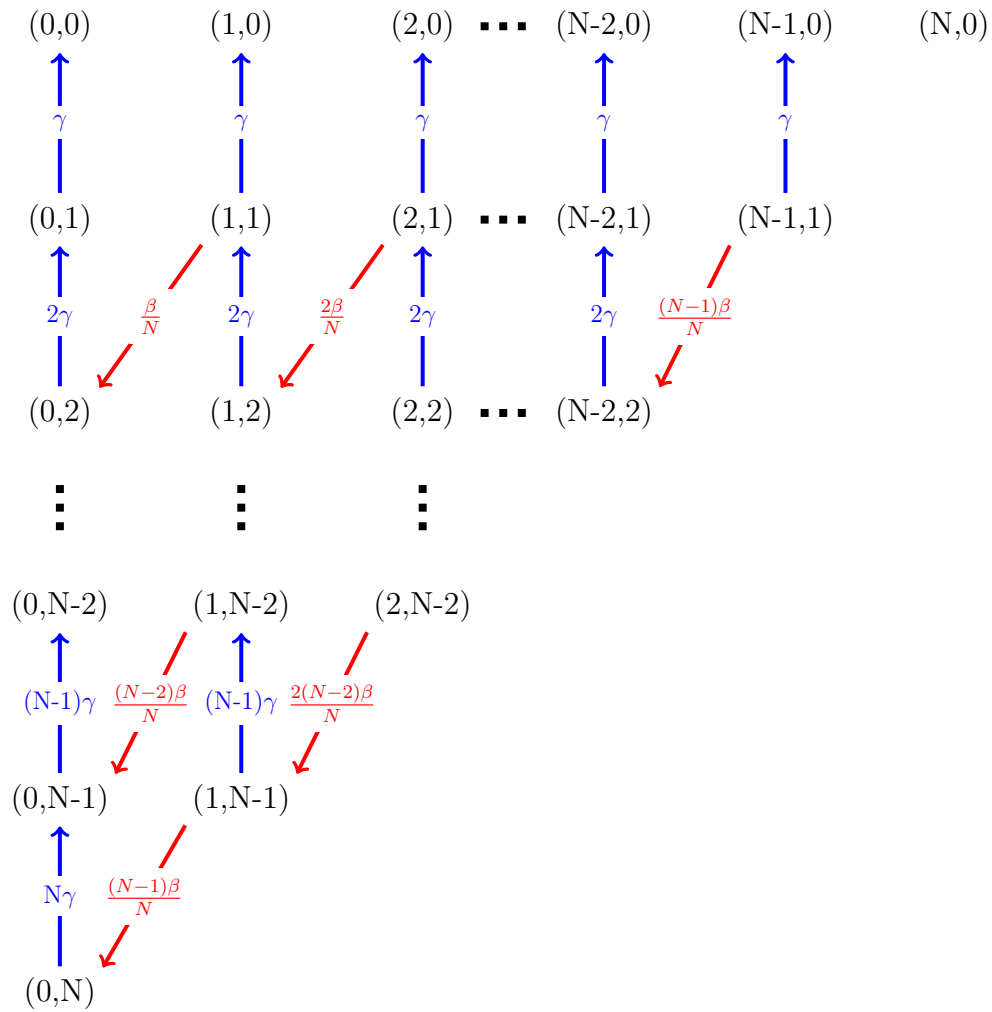
De toestandsruimte  $I$  van het SIR-CTMK-model is

$$I = \{(s, i) \mid 0 \leq s + i \leq N, s, i \in \mathbb{N}_0\}.$$

De kansvariabelen van het proces zijn  $S(t)$  en  $I(t)$ .

De transitiekansen van het proces zijn gebaseerd op 1.1. Het proces is stationair. Door aan elke toestand uit de toestandsruimte de bijbehorende infectiekracht en genezingsnelheid toe te kennen, komt het transitiediagram uit figuur 3.1 tot stand.

De toestanden  $(s, i)$  waarvoor  $i = 0$  zijn de absorberende toestanden van het SIR-CTMK-model. Wanneer het proces in zo'n toestand terechtkomt, blijft het in die toestand. Een epidemie eindigt wanneer het proces zo'n absorberende toestand bereikt [1].



Figuur 3.1: Transitiediagram voor het SIR-model met populatiegrootte  $N$ . Rode pijlen corresponderen met de infectiekracht  $\beta(s/N)i$ , blauwe pijlen met de genezingsnelheid  $\gamma i$ .



We stellen nu de  $Q$ -matrix  $Q$  op voor 3.1 op. Voor de elementen van  $Q$  geldt:

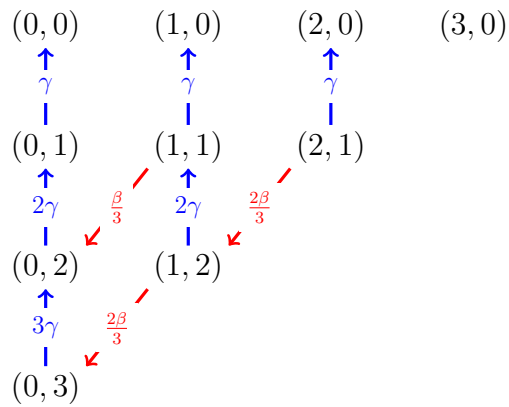
$$q_{(s,i)(r,j)} = \begin{cases} \beta \frac{s}{N} i & \text{als } (r, j) = (s - 1, i + 1), \\ \gamma i & \text{als } (r, j) = (s, i - 1), \\ -\beta \frac{s}{N} i - \gamma i & \text{als } (r, j) = (s, i), \\ 0 & \text{anders.} \end{cases} \quad (3.1)$$

Voor de elementen van de bijbehorende sprongmatrix  $\Pi$  geldt

$$\pi_{(s,i)(r,j)} = \begin{cases} 1 & \text{als } (r, j) = (s, i) \text{ en } i = 0, \\ \frac{\beta s}{\beta s + N\gamma} & \text{als } (r, j) = (s - 1, i + 1) \text{ en } i \neq 0, \\ \frac{N\gamma}{\beta s + N\gamma} & \text{als } (r, j) = (s, i - 1) \text{ en } i \neq 0, \\ 0 & \text{anders .} \end{cases} \quad (3.2)$$

$Q$  en  $\Pi$  zijn beide  $\left(\frac{(N+1)(N+2)}{2} \times \frac{(N+1)((N+2))}{2}\right)$ -matrices. Voor grote waarden van  $N$  is het dus moeilijk deze matrices volledig uit te schrijven. Als voorbeeld bekijken we een populatie van grootte  $N = 3$ .

**Voorbeeld 1.** *Stel dat  $N = 3$ . De toestandsruimte met voor elke toestand de bijbehorende infectiekracht  $\beta \frac{s}{3} i$  en genezingsnelheid  $\gamma i$ , is afgebeeld in 3.2. De  $Q$ -matrix voor dit transitiediagram is gegeven in vergelijking 3.3. De bijbehorende sprongmatrix, onder de aanname dat  $\gamma > 0$  of  $\beta > 0$ , wordt gegeven in vergelijking 3.4.*



Figuur 3.2: Transitiediagram van het SIR-model voor  $N=3$ .

$$Q = \begin{bmatrix}
 & (0,0) & (1,0) & (2,0) & (3,0) & (0,1) & (1,1) & (2,1) & (0,2) & (1,2) & (0,3) \\
 (0,0) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 (1,0) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 (2,0) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 (3,0) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 (0,1) & \gamma & 0 & 0 & 0 & -\gamma & 0 & 0 & 0 & 0 & 0 \\
 (1,1) & 0 & \gamma & 0 & 0 & 0 & -\gamma - \frac{\beta}{3} & 0 & \frac{\beta}{3} & 0 & 0 \\
 (2,1) & 0 & 0 & \gamma & 0 & 0 & 0 & -\gamma - \frac{2\beta}{3} & 0 & \frac{2\beta}{3} & 0 \\
 (0,2) & 0 & 0 & 0 & 0 & 2\gamma & 0 & 0 & -2\gamma & 0 & 0 \\
 (1,2) & 0 & 0 & 0 & 0 & 0 & 2\gamma & 0 & 0 & -2\gamma - \frac{2\beta}{3} & \frac{2\beta}{3} \\
 (3,0) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3\gamma & 0 & -3\gamma
 \end{bmatrix} \tag{3.3}$$

$$\Pi = \begin{bmatrix} & (0,0) & (1,0) & (2,0) & (3,0) & (0,1) & (1,1) & (2,1) & (0,2) & (1,2) & (0,3) \\ (0,0) & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ (1,0) & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ (2,0) & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ (3,0) & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ (0,1) & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ (1,1) & 0 & \frac{3\gamma}{3\gamma+\beta} & 0 & 0 & 0 & 0 & 0 & \frac{\beta}{3\gamma+\beta} & 0 & 0 \\ (2,1) & 0 & 0 & \frac{3\gamma}{3\gamma+2\beta} & 0 & 0 & 0 & 0 & 0 & \frac{2\beta}{3\gamma+2\beta} & 0 \\ (0,2) & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ (1,2) & 0 & 0 & 0 & 0 & 0 & \frac{3\gamma}{3\gamma+\beta} & 0 & 0 & 0 & \frac{\beta}{3\gamma+\beta} \\ (3,0) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad (3.4)$$

## 3.2 Definities van een SIR-CTMK

We gaan nu aan de hand van 3.1 en 3.2 een SIR-CTMK-proces construeren. We hebben de volgende 3 definities van een SIR-CTMK-model.

**Stelling 3.2.1** (Definities SIR-CTMK-model). *Zij  $(X_t)_{t \geq 0}$  een rechts-continu proces met toestandsruimte*

$$I = \{(s, i) \mid 0 \leq s + i \leq N, s, i \in \mathbb{N}_0\},$$

*en initiële distributie  $\delta_{(s_0, i_0)}$ . Zij  $Q$  een  $Q$ -matrix op  $I$  met bijbehorende sprongmatrix  $\Pi$ , die voldoen aan 3.1 en 3.2, respectievelijk. Dan zijn de volgende 3 uitspraken equivalent:*

(i) (Sprongketen-en-wachttijddefinitie)

$(X_t)_{t \geq 0}$  is Markov( $\delta_{(s_0, i_0)}, Q$ ).

(ii) (Infinitesimale definitie)

Voor alle  $t, h \geq 0$  geldt dat  $X_{t+h}$ , conditioneel op  $X_t = (s, i)$ , onafhankelijk is van  $(X_u | u \leq t)$ . Bovendien geldt voor alle  $j \in I$  dat als  $h \downarrow 0$  uniform in  $t$ , dan

$$\mathbb{P}(X_{t+h} = (r, j) | X_t = (s, i)) = \delta_{(s,i)(r,j)} + q_{(s,i)(r,j)}h + o(h);$$

oftewel

$$\mathbb{P}(X_{t+h} = (s-1, i+1) | X_t = (s, i)) = h \left( \beta \frac{s}{N} i \right) + o(h)$$

$$\mathbb{P}(X_{t+h} = (s, i-1) | X_t = (s, i)) = h(\gamma i) + o(h)$$

$$\mathbb{P}(X_{t+h} = (s, i) | X_t = (s, i)) = 1 - h \left( \beta \frac{s}{N} i + \gamma i \right) + o(h).$$

(iii) (Transitiekansendefinitie)

Voor alle  $n = 0, 1, 2, \dots$ , alle tijden  $0 \leq t_1 \leq t_2 \leq \dots \leq t_{n+1}$  en alle toestanden  $(s, i)_0, \dots, (s, i)_{n+1}$ , geldt

$$\begin{aligned} \mathbb{P}(X_{t_{n+1}} = (s, i)_{n+1} | X_{t_0} = (s_0, i_0), \dots, X_{t_n} = (s_n, i_n)) \\ = p_{(s_n, i_n)(s_{n+1}, i_{n+1})}(t_{n+1} - t_n), \end{aligned}$$

waarbij  $(p_{(s,i)(r,j)} | (s, i), (r, j) \in I, t \geq 0)$  de unieke oplossing is van de voorwaartse vergelijking

$$P'(t) = P(t)Q, \quad P(0) = I. \quad (3.5)$$

3.5 uitgeschreven in coëfficiënten is

$$P'_{(s,t)(r,j)}(t) = \begin{cases} p_{(s,t)(s-1,i+1)}(t) \left( \beta \frac{s}{N} i \right) + p_{(s,t)(s,i-1)}(t) \gamma i - p_{(s,t)(s,t)}(t) \left( -\beta \frac{s}{N} i - \gamma i \right) \\ \text{als } 1 \leq i \leq N-1, s \geq 1, \\ \\ p_{(0,i)(0,i-1)}(t) \gamma i - p_{(0,i)(0,i)}(t) \gamma i \\ \text{als } 1 \leq i \leq N, s = 0, \\ \\ 0 \text{ anders.} \end{cases}$$

Voor de beginwaarden van 3.5 geldt

$$p_{(s,i)(r,j)}(0) = \delta_{(s,i)(r,j)}.$$

Als  $(X_t)_{t \geq 0}$  aan één van bovenstaande voorwaarden voldoet, is  $(X_t)_{t \geq 0}$  een SIR-CTMK-proces. De parameters die het proces vastleggen zijn  $N$ ,  $\beta$ ,  $\gamma$  en de initiële toestand  $(s_0, i_0)$ .

Er is geen simpele manier om de matrix  $P(t)$  in een expliciete gesloten vorm te vinden. We kunnen echter  $P(t)$  wel benaderen voor een eindige toestandruimte door de oneindige som af te kappen tot een eindige som. Voor een benadering van de kans op een grote ziekte-uitbraak aan de hand van het SIR-CTMK-model verwijzen we naar [1].

# Hoofdstuk 4

## Simulaties

### 4.1 Gillespie-algoritme

Voor processen met twee variabelen is het vaak lastig een analytische oplossing te vinden voor de transitiekansen aan de hand van de voorwaartse vergelijking 2.5. Het is dan vaak eenvoudiger om numeriek voorbeeldpaden van het proces te simuleren [1]. Een methode die hiervoor gebruikt kan worden is het Gillespie-algoritme. De manier waarop we het algoritme hebben geïmplementeerd, komt overeen met [1]. Voor elke transitie van het proces worden twee uniforme random getallen  $u_1, u_2 \in U[0, 1]$  gegenereerd; één voor de wachttijd tussen twee transities, en één om te bepalen welke gebeurtenis heeft plaatsgevonden.

Voor het SIR-CTMK-model geldt dat een wachttijd  $T$  exponentieel verdeeld is met parameter  $\lambda$ ,  $T \sim \lambda e^{-\lambda t}$ , waarbij

$$\lambda = \beta \frac{s}{N} i + \gamma i.$$

De waarde van de wachttijd  $\tau$  wordt als volgt berekend:

$$\tau = -\frac{\ln u_1}{\lambda}.$$

Het tweede uniforme random getal  $u_2$  bepaalt welke gebeurtenis heeft plaatsgevonden; een besmetting of een genezing. Het interval  $[0, 1]$  wordt verdeeld in twee stukken:  $[0, p_1]$  en  $[p_1, p_1 + p_2]$ . De grootte van de stukken wordt bepaald aan de hand van de kans op een gebeurtenis. De kans op een bepaalde gebeurtenis wordt bepaald door de relatieve grootte van de bijbehorende transitiesnelheid. Voor het

SIR-CMTK-model geldt

$$p_1 = \mathbb{P}(\text{besmetting}) = \frac{\beta(s/N)i}{\beta(s/N)i + \gamma i} = \frac{\beta s}{\beta s + \gamma N}$$

en

$$p_2 = \mathbb{P}(\text{genezing}) = \frac{\gamma i}{\beta(s/N)i + \gamma i} = \frac{\gamma N}{\beta s + \gamma N}.$$

Merk op dat  $p_1 + p_2 = 1$ . In het geval dat  $u_2 \in [0, p_1]$ , vindt er een besmetting plaats. Als  $u_2 \in (p_1, 1]$  geneest een besmettelijk persoon. Het proces stopt wanneer  $i = 0$ . In dit geval is een absorberende toestand bereikt, en zal het proces niet meer transitioneren. [1]

In algoritme 2 is pseudocode gegeven die gebruikt is om voorbeeldpaden van het SIR-CTMK-model te simuleren.

Stel dat de gemiddelde besmettingsduur voor een bepaalde ziekte 10 dagen is, en we dagen als tijdseenheid gebruiken. Dan kiezen we  $\gamma = \frac{1}{10}$ . We zijn geïnteresseerd in het verloop van de epidemie voor verschillende basale reproductiegetallen, vooral in het verschil tussen  $R_0 < 1$  en  $R_0 > 1$ . We testen  $\beta = \{0.05, 0.1, 0.2, 1\}$ . Voor onze simulaties hebben we gekozen voor  $N = 10.000$  en  $I(0) = 100$ . In dit geval is het initiële aantal vatbaren relatief groot. Voor elke waarde  $\beta$  wordt 5 keer een voorbeeldpad gegenereerd. Voor elk pad worden de waarden  $S(t)$  en  $I(t)$  op de sprongtijden van het proces afgebeeld. Elke simulatie stopt op tijdstip  $t$  waarvoor  $I(t) = 0$ . Op dit moment is de epidemie uitgestorven.

## 4.2 Resultaten

In de resultaten is terug te zien dat voor  $R_0 > 1$  (dus voor  $\beta = 0.2$  en  $\beta = 1$ ) het aantal besmettelijken  $I(t)$  eerst stijgt. Hoewel er verschil te zien is tussen de voorbeeldpaden voor een bepaalde parametersetting, hebben de paden met dezelfde waarden  $\beta$  en  $\gamma$  ongeveer dezelfde vorm en richting. Voorbeeldpaden liggen dicht bij elkaar voor  $R_0$ -waarden verder verwijderd van 1. Dit komt doordat de kans op een besmetting en genezing voor deze waarden niet gelijk is verdeeld, maar één van de twee vaker wordt gekozen door het algoritme, in alle voorbeeldpaden. Voor alle waarden geldt dat de epidemie uiteindelijk uitsterft. Dit komt doordat de populatie constant is en immuniteit voor de ziekte opbouwt na een besmetting.

---

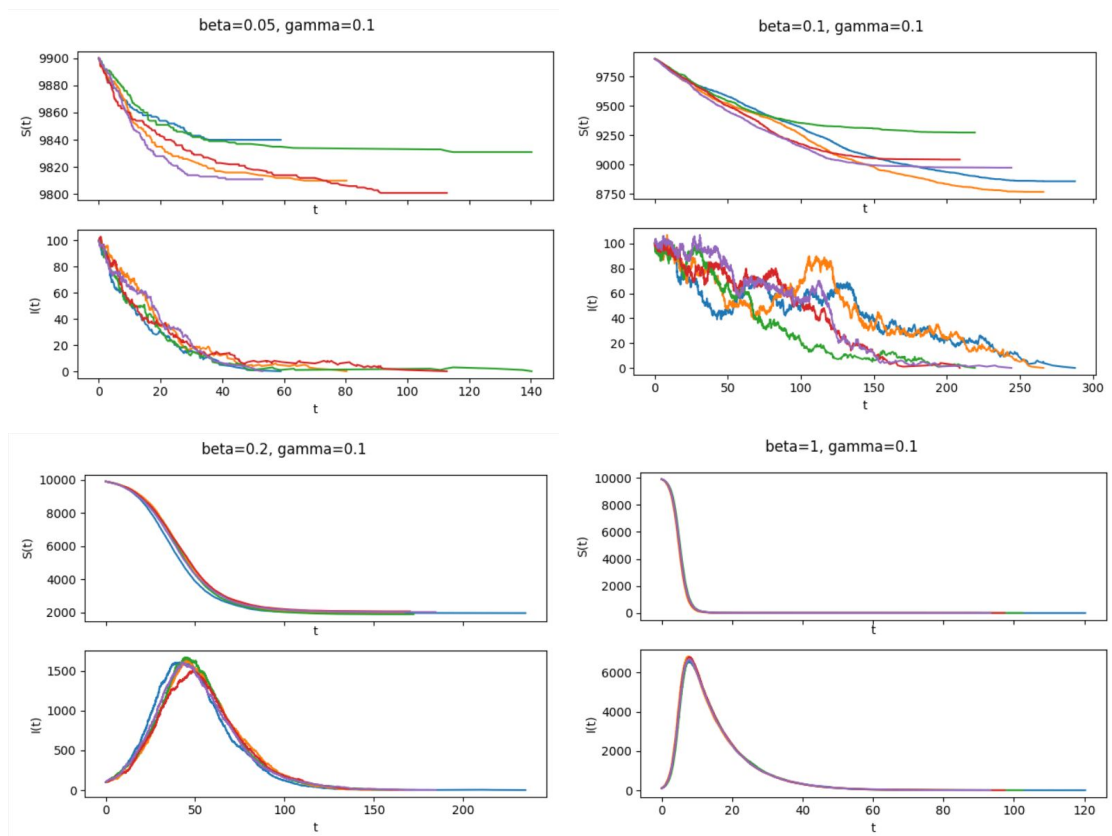
**Algorithm 2** Pseudocode van het Gillespie-algoritme voor het SIR-CTMK-model.

---

```
1: Stel parameters  $\beta, \gamma, I(0)$  en  $N$  in.
2:
3: for aantal simulaties do
4:    $i = I(0)$ ;
5:    $s = N - I(0)$ ;
6:
7:   while  $0 < i < N - 1$  do
8:     infectiekracht =  $\beta \frac{s}{N} i$ 
9:     genezingsnelheid =  $\gamma i$ 
10:    totaal = infectiekracht + genezingsnelheid
11:
12:    Genereer een wachttijd:
13:     $u_1, u_2 = U[0, 1]$ 
14:
15:    wachttijd =  $\frac{-\ln u_1}{\text{totaal}}$ 
16:
17:    Bepaal of een besmetting of genezing heeft plaatsgevonden:
18:     $\mathbb{P}(\text{besmetting}) = \text{infectiekracht}/\text{totaal}$ 
19:     $\mathbb{P}(\text{genezing}) = \text{genezingsnelheid}/\text{totaal}$ 
20:
21:    if  $u_2 \leq \mathbb{P}(\text{besmetting})$  then
22:       $i = i + 1$ 
23:       $s = s - 1$ 
24:    else
25:       $i = i - 1$ 
```

---





Figuur 4.1: Voorbeeldpaden van  $S(t)$  en  $I(t)$  voor het SIR-CTMK-model voor verschillende waarden van  $\beta$  en  $\gamma$ .

# Hoofdstuk 5

## Deterministisch SIR-model

Het deterministische SIR-model, dat berust op het volgende stelsel van differentiaalvergelijkingen:

$$\frac{dS}{dt} = -\beta I \frac{S}{N}, \quad (5.1)$$

$$\frac{dI}{dt} = \beta I \frac{S}{N} - \gamma I, \quad (5.2)$$

$$\frac{dR}{dt} = \gamma I, \quad (5.3)$$

is de *mean-field* benadering van het stochastische SIR-CTMK-model. In 1970 is voor het eerst aangetoond dat voor het SIR-CTMK-model op een compleet netwerk (waarin alle individuen onderling verbonden zijn) de fracties van het aantal vatbare, besmettelijke en genezen individuen convergeert naar de oplossing van het deterministische model als de populatiegrootte  $N$  groter wordt. In 2015 gaf [3] een eenvoudiger bewijs voor deze eigenschap van het SIR-CTMK-model. In 5.0.1 geven we de hoofdstelling van [3].

Zij  $X_n(t)$  een SIR-CTMK-model voor een populatie van grootte  $n$ .  $S_n(t)$  is het aantal vatbaren op tijdstip  $t$ ,  $I_n(t)$  het aantal besmettelijken. We definiëren

$$s_n(t) = \frac{S_n(t)}{n}, \quad i_n(t) = \frac{I_n(t)}{n}$$

als de fracties van vatbaren en besmettelijken op tijdstip  $t$ , respectievelijk. De exacte waarden van de verwachte waarden van  $S_n(t)$  en  $I_n(t)$  kunnen worden berekend met definitie (iii) van 3.2.1.

**Stelling 5.0.1.** *Als  $s_n(0) \rightarrow s_0$  en  $i_n(0) \rightarrow i_0$  als  $n \rightarrow \infty$ , en  $s_0, i_0 \in [0, 1]$ , dan convergeert  $(s_n(t), i_n(t))$  uniform in gemiddelde kwadratische fout, oftewel*

$$E [|| (s_n(t), i_n(t)) - y(t) ||^2] \rightarrow 0,$$

*op elk eindige tijdsinterval  $[0, T]$  naar de oplossing van de mean-field vergelijkingen*

$$y_1' = -\beta y_1 y_2, \quad y_1(0) = s_0, \tag{5.4}$$

$$y_2' = \beta y_1 y_2 - \gamma y_2, \quad y_2(0) = i_0. \tag{5.5}$$

$$\tag{5.6}$$

Door vergelijkingen 5.5 en 5.6 door de populatiegrootte te delen, krijgen we precies vergelijkingen 5.2 en 5.3 van het deterministische model. Voor het bewijs van stelling 5.0.1 verwijzen we naar [3].

# Hoofdstuk 6

## Het SIS-CTMK-model

Het voornaamste verschil tussen het SIR- en SIS-model is dat in het SIS-model besmettelijke individuen niet genezen, maar opnieuw vatbaar worden voor de ziekte. In dit hoofdstuk stellen we een SIS-CTMK-model op. We gebruiken hiervoor dezelfde notaties voor de kansvariabelen als voor het SIR-model.

De toestandruimte  $I$  voor het SIS-CTMK-model is gelijk aan die van het SIR-model, oftewel

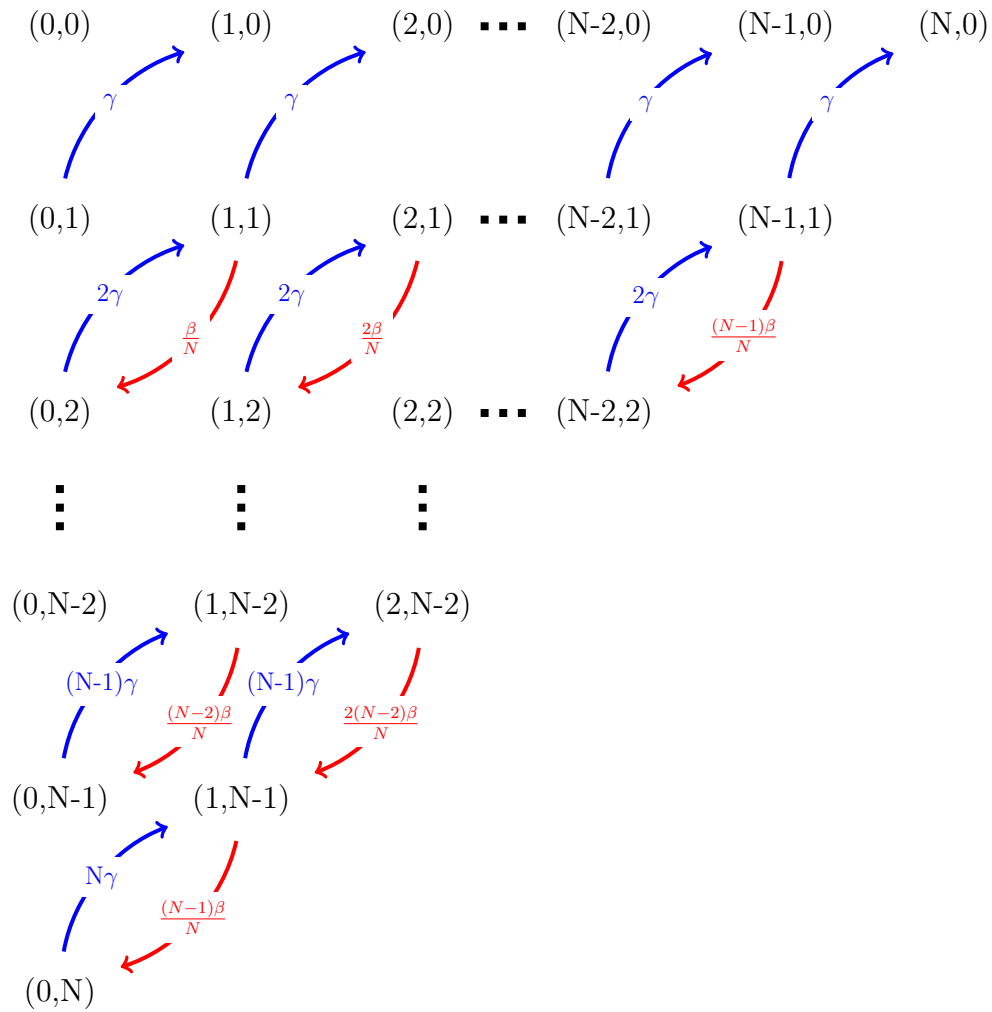
$$I = \{(s, i) \mid 0 \leq s + i \leq N, s, i \in \mathbb{N}_0\}.$$

Het transitiediagram voor het SIS-CTMK-model voor een populatie van grootte  $N$  is gegeven in figuur 6.1.

De elementen  $q_{(s,i)(t,j)}$  van de Q-matrix  $Q$  van het SIS-CTMK-model voldoen aan

$$q_{(s,i)(t,j)} = \begin{cases} \beta \frac{s}{N} i & \text{als } (t, j) = (s - 1, i + 1), \\ \gamma i & \text{als } (t, j) = (s + 1, i - 1), \\ -\beta \frac{s}{N} i - \gamma i & \text{als } (t, j) = (s, i), \\ 0 & \text{anders .} \end{cases} \quad (6.1)$$

Voor de elementen  $\pi_{(s,i)(t,j)}$  van de bijbehorende sprongmatrix  $\Pi$  geldt



Figuur 6.1: Transitiediagram voor het SIS-model met populatiegrootte  $N$ . Rode pijlen corresponderen met de infectiekracht, blauwe pijlen met de genezingsnelheid.

$$\pi_{(s,i)(t,j)} = \begin{cases} \frac{\beta s}{\beta s + N\gamma} & \text{als } (t, j) = (s - 1, i + 1) \text{ en } i \neq 0, \\ \frac{N\gamma}{\beta s + N\gamma} & \text{als } (t, j) = (s + 1, i - 1) \text{ en } i \neq 0, \\ 1 & \text{als } (t, j) = (s, j) \text{ en } i = 0, \\ 0 & \text{anders .} \end{cases} \quad (6.2)$$

We kunnen nu stelling 2.4.1 toepassen. In stelling 6.0.1 geven we de 3 definities van het SIS-CTMK-model.

**Stelling 6.0.1.** *Zij  $(X_t)_{t \geq 0}$  een rechts-continu proces met toestandsruimte*

$$I = \{(s, i) \mid 0 \leq s + i \leq N, s, i \in \mathbb{N}_0\},$$

*en initiële distributie  $\delta_{(s,i)}$ . Zij  $Q$  een  $Q$ -matrix op  $I$  met bijbehorende sprongmatrix  $\Pi$ , die voldoen aan 6.1 en 6.2, respectievelijk. Dan zijn de volgende 3 uitspraken equivalent:*

(i) (Sprongketen-en-wachttijddefinitie)

$(X_t)_{t \geq 0}$  is Markov( $\delta_{(s,i)}, Q$ ).

(ii) (Infinitesimale definitie)

*Voor alle  $t, h \geq 0$  geldt dat  $X_{t+h}$ , conditioneel op  $X_t = (s, i)$ , onafhankelijk is van  $(X_u \mid u \leq t)$ . Bovendien geldt voor alle  $j \in I$  dat als  $h \downarrow 0$  uniform in  $t$ , dan*

$$\mathbb{P}(X_{t+h} = (r, j) \mid X_t = (s, i)) = \delta_{(s,i)(r,j)} + q_{(s,i)(r,j)}h + o(h);$$

*oftewel*

$$\mathbb{P}(X_{t+h} = (s - 1, i + 1) \mid X_t = (s, i)) = h \left( \beta \frac{s}{N} i \right) + o(h)$$

$$\mathbb{P}(X_{t+h} = (s + 1, i - 1) \mid X_t = (s, i)) = h (\gamma i) + o(h)$$

$$\mathbb{P}(X_{t+h} = (s, i) \mid X_t = (s, i)) = 1 - h \left( \beta \frac{s}{N} i + \gamma i \right) + o(h).$$

(iii) (Transitiekansdefinitie)

Voor alle  $n = 0, 1, 2, \dots$ , alle tijden  $0 \leq t_1 \leq t_2 \leq \dots \leq t_{n+1}$  en alle toestanden  $(s, i)_0, \dots, (s, i)_{n+1}$ , geldt

$$\mathbb{P}(X_{t_{n+1}} = (s, i)_{n+1} \mid X_{t_0} = (s, i)_0, \dots, X_{t_n} = (s, i)_n) = p_{(s,i)_n(s,i)_{n+1}}(t_{n+1} - t_n),$$

waarbij  $(p_{(s,i)(r,j)} \mid (s, i), (r, j) \in I, t \geq 0)$  de unieke oplossing is van de voorwaartse vergelijking

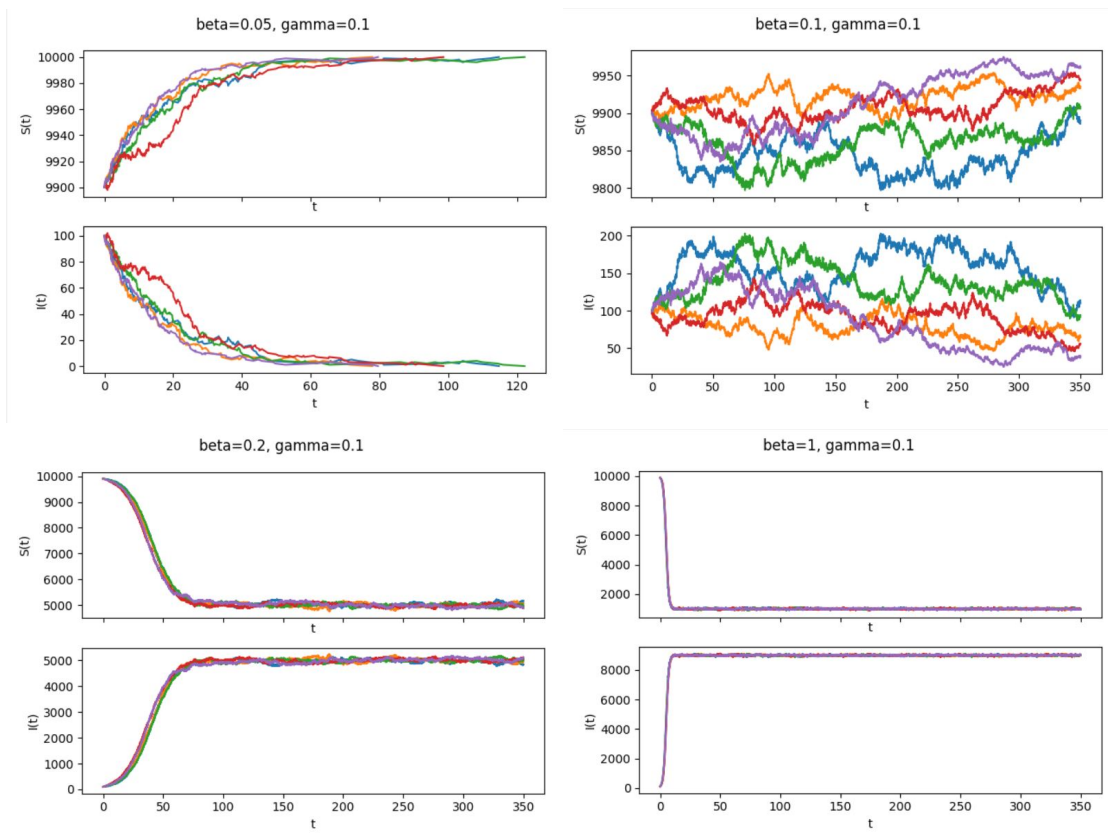
$$P'(t) = P(t)Q, \quad P(0) = I. \quad (6.3)$$

Als  $(X_t)_{t \geq 0}$  aan één van bovenstaande voorwaarden voldoet, is  $(X_t)_{t \geq 0}$  een SIS-CTMK-proces voor een populatie van constante grootte  $N$ , infectiekracht  $\beta(s/N)I$ , genezingskracht  $\gamma i$  en initiële distributie  $\delta_{(s,i)}$ .

## 6.1 Simulaties

We hebben een kleine aanpassing gemaakt aan algoritme 2 zodat we voorbeeldpaden van het SIS-CTMK-model kunnen simuleren.

In de resultaten zien we dat alleen in het geval dat  $\beta$  kleiner is dan  $\gamma$  de ziekte uit de populatie verdwijnt. De voorbeeldpaden voor  $\beta = \gamma = 0.1$  verschillen veel van elkaar. Voor  $R_0 > 1$  zien we dat het aantal besmettelijken eerst sterk toeneemt. Daarna vinden er nog wel kleine veranderingen in  $I(t)$  en  $S(t)$  plaats, maar deze blijven ongeveer dezelfde waarde.



Figuur 6.2: Voorbeeldpaden van  $S(t)$  en  $I(t)$  voor het SIS-CTMK-model voor verschillende waarden van  $\beta$  en  $\gamma$ .



# Hoofdstuk 7

## Conclusie

In deze scriptie hebben we 3 definities van het SIR-CTMK-model voor een populatie van constante grootte  $N$  met infectiekracht  $\beta(S(t)/N)I(t)$  en genezingsnelheid  $\gamma I(t)$  gegeven. We hebben laten zien hoe de kans op een grote uitbraak benaderd kan worden aan het begin van een ziekte-uitbraak. Vervolgens hebben we voorbeeldpaden van het SIR-CTMK-model gesimuleerd aan de hand van het Gillespie-algoritme. Uit de resultaten van de simulaties van het model kan geconcludeerd worden dat de keuze van  $\beta$  en  $\gamma$  van grote invloed zijn op het verloop van de epidemie. Om goede voorspellingen te kunnen doen over het verloop van een ziekte-uitbraak, is het dus van belang goede schattingen te maken van  $\beta$  en  $\gamma$ .

In deze scriptie is een stochastische versie van het SIR-model gepresenteerd. Er zijn echter nog meer wiskundige epidemische modellen die als continue-tijd-Markovketen opgesteld kunnen worden door de toestandruimte, Q-matrix en sprongmatrix te identificeren. Door de algemene opzet van deze scriptie is het mogelijk om andere epidemische modellen met eindige toestandruimte als continue-tijd-Markovketen op te zetten. In Bijlage A laten we zien hoe het SIS (*Susceptible, Infectious, Susceptible*)-CTMK-model gedefinieerd wordt.

# Bibliografie

- [1] Allen, L. J. S. (2017). A Primer on Stochastic Epidemic Models: Formulation, Numerical Simulation, and Analysis. *Infectious Disease Modelling* 2, p 128-142.
- [2] Allen, L. J. (2015). Stochastic Population and Epidemic Models. *Mathematical Biosciences Lecture Series, Stochastics in Biological Systems*.
- [3] Armbruster, B., Beck, E. (2017). Elementary Proof of Convergence to the Mean-field Model for the SIR Process. *Journal of Mathematical Biology* 75.2, p 327-339.
- [4] Donsimoni, J., Glawion, R., Plachter, B., Wälde, K. (2020). Projecting the Spread of COVID-19 for Germany. *German Economic Review*, 21(2), 181-216.
- [5] Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., Colaneri, M. (2020). Modelling the COVID-19 Epidemic and Implementation of Population-wide Interventions in Italy. *Nature Medicine*, 1-6.
- [6] Harko, T., Lobo, F. S. N., Mak, M. K. (2014). Exact Analytical Solutions of the Susceptible-Infected-Recovered (SIR) Epidemic Model and of the SIR Model with Equal Death and Birth Rates. *Applied Mathematics and Computation* 236, p 184-194.
- [7] Kermack, W. O., McKendrick, A. G. (1927). A Contribution to the Mathematical Theory of Epidemics. *Proceeding of the Royal Society of London, Serie A Containing Papers of Mathematics and Physical Character*, 115, 700-721.
- [8] Jacquez, J. A., Simon, C. P. (1993). The Stochastic SI Model with Recruitment and Deaths I. Comparison with the Closed SIS Model. *Mathematical Biosciences* 117(1-2), p 77-125
- [9] Norris, J. R. (1998). *Markov Chains (No.2)*. Cambridge university press.

[10] Ross, S. M. (2014). Introduction to Probability Models. Academic press.