Printed at the Mathematical Centre, 413 Kruislaan, 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.).

MC SYLLABUS 50

INLEIDING SYSTEEMTHEORIE

H. NIJMEIJER

MATHEMATISCH CENTRUM AMSTERDAM 1982

1980 Mathematics subject classification: 93XX01

ISBN 90 6196 240 0

INHOUD

VOORWOORD

Deze syllabus is een herziene en uitgebreide versie van een kollegediktaat voor een kollege Inleiding Systeemtheorie, gegeven aan de Universiteit van Amsterdam (voorjaar 1981). De aanzet tot deze syllabus is gegeven door Prof.dr. G. de Leve; bij de voorbereiding hiervan heb ik veel medewerking gehad van Prof.dr.ir. J.C. Willems en dr.ir. J.H. van Schuppen, die evenals drs. W. Ravenek, ook kritisch kommentaar op het oorspronkelijk manuskript geleverd hebben.

INLEIDING

Sinds het begin van de zestiger jaren mag de systeemtheorie zich verheugen in een sterk toegenomen belangstelling. Deze interesse is voor een belangrijk deel gekoppeld aan de grote technische vooruitgang in onze maatschappij. We hoeven hierbij slechts te denken aan de stormachtige ontwikkeling op het gebied van de ruimtevaart.

Systeemtheorie, of liever regeltheorie behoorde oorspronkelijk tot het vakgebied der ingenieurs. Tegenwoordig wordt de systeemtheorie, mede door zijn moderne wiskunde-onderbouw, toegepast in vele verschillende vakgebieden (zoals de natuurkunde, scheikunde, biologie en de ekonomie).

Bij het wiskundig modelleren van een ekonomisch (ekonometrisch) model maakt men onderscheid tussen verschillende grootheden, namelijk de *besturingsgrootheden* u (instrumentele variabelen, inputs, controls), *uitgangsgrootheden* y ("targets" en endogene variabelen, outputs) en *storingen* q (exogene variabelen). Het verband tussen deze grootheden wordt door de gebruikte ekonomische wetten gegeven. De systeemtheorie houdt zich nu bezig met de wiskundige achtergrond van deze dynamische verschijnselen. Karakteristieke (ekonomische) problemen kunnen op deze manier bestudeerd worden, bijvoorbeeld:

- *stabiliteit*: Als $u \equiv 0$ en $q \equiv 0$ konvergeert de uitgangsgrootheid y dan ook naar nul?
- *storingsonderdrukking*: Kunnen we $u(t)$ zo bepalen dat $q(t)$ (bijna) geen invloed heeft op $y(t)$?
- *regelbaarheid*: Kunnen we met behulp van een geschikte besturing iedere gewenste uitgangswaarde bereiken?

Deze syllabus geeft een inleiding tot de lineaire systeemtheorie. De belangrijkste begrippen uit de systeemtheorie komen erin aan de orde. Door middel van vele voorbeelden en opgaven aan het einde van elk hoofdstuk wordt de theorie geïllustreerd. De meeste van de wiskundige bewijzen van stellingen zijn hier achterwege gebleven. Voor de geïnteresseerde lezer is daarom een beknopte literatuurlijst toegevoegd. De noodzakelijke voorkennis (nivo: kandidaats wiskunde) op het gebied van de wiskunde is in een drietal appendices samengevat.

HOOFDSTUK 1

LINEAIRE DYNAMISCHE SYSTEMEN

Alvorens ik een definitie zal geven van wat we hier onder een dynamisch systeem zullen verstaan, wil ik eerst een aantal voorbeelden (modellen) behandelen, welke als motivatie voor de opbouw van het vak systeemtheorie kunnen dienen. (Met opzet zijn deze voorbeelden gekozen uit verschillende takken van de wetenschap).

<u>VOORBEELD 1.1.</u>

Een auto rijdt wrijvingsloos op een rechte weg. We willen de plaats op een willekeurig tijdstip t bepalen aan de hand van de aandrijvingskracht van de auto.

Stellen we de plaats op tijdstip t voor door $x(t)$ met beginpositie $x(0) = 0$ en beginsnelheid $\dot{x}(0) = 0$, dan krijgen we de volgende grootheden

$\dot{x}(t)$: snelheid ten tijde t

$\ddot{x}(t)$: versnelling op tijdstip t.

Volgens de tweede wet van Newton voldoet de aandrijvingskracht van de auto op tijdstip t, $u(t)$, aan:

$$u(t) = m \cdot \ddot{x}(t), \qquad m = \text{massa van de auto.}$$

We kunnen dit als systeem schrijven door te definiëren:

$$x_1(t) := x(t), \qquad x_2(t) := \dot{x}(t)$$

$$\Rightarrow \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{1}{m} \begin{pmatrix} 0 \\ u \end{pmatrix}, \quad \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

waarneming (d.w.z. plaats van auto op tijdstip t)

$$y = (1 \ 0) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \qquad \square$$

<u>VOORBEELD 1.2.</u> 's Lands ekonomie.
We definiëren de volgende grootheden:

$Y(k)$: Nationale produkt op tijdstip k,
$C(k)$: Konsumptie op tijdstip k,
$I(k)$: Investeringen op tijdstip k,
$G(k)$: Staatsuitgaven op tijdstip k.

We denken hierbij aan maandelijkse perioden, dus $Y(k)$ stelt het nationale produkt over de k^e maand voor en evenzo voor $C(k)$, $I(k)$ en $G(k)$.
Als basisvergelijking nemen we:

$$Y(k) = C(k) + I(k) + G(k).$$

Verder nemen we aan dat

$$C(k) = m\,Y(k), \qquad 0 < m < 1$$

en

$$Y(k+1) - Y(k) = r\,I(k), \qquad r > 0.$$

Kombinatie van deze vergelijkingen geeft

$$Y(k+1) = [1 + r(1-m)]Y(k) - r\,G(k).$$

Opnieuw kunnen we deze vergelijking interpreteren als een relatie tussen besturingsgrootheden $G(k)$ en uitgangsgrootheden $Y(k)$. □

<u>VOORBEELD 1.3.</u> Populatiedynamika.
We willen de totale populatie N uitdrukken als funktie van het aantal geboorten per tijdseenheid B. Demografen leveren ons een funktie P waarbij $P(x,t)$ de waarschijnlijkheid is dat iemand die geboren is op het tijdstip $t-x$ op het tijdstip t nog in leven is. Dan is

$$N(t) = \int_{-\infty}^{t} P(t-\tau,\tau)B(\tau)d\tau$$

Opnieuw zien we weer een verband tussen besturingsgrootheid $B(t)$ en uitgangsgrootheid $N(t)$. □

We zien een belangrijk verschil tussen de voorbeelden 1.1 en 1.2 enerzijds en voorbeeld 1.3 anderzijds. Het laatste voorbeeld geeft een direkt verband tussen de besturingsfunktie en de waarnemingsfunktie, terwijl in de andere voorbeelden dit verband via een *geheugenfunktie* gegeven wordt. Dit verschil zullen we formaliseren middels een tweetal definities van een systeem.

DEFINITIE 1.4.

Een *lineair dynamisch systeem in Ingangs-Uitgangsvorm (Input-Output)* is gedefiniëerd door $\sum_{I/O} := \{T,U,\mathcal{U},Y,\mathcal{Y},F\}$

waarbij:

(i) $T \subseteq \mathbb{R}$, de tijdsas meestal $T = \mathbb{Z}$ of $T = \mathbb{R}$,

(ii) U, Y vektorruimten over $\mathbb{R}$,

U, verzameling van ingangswaarden, Y die van uitgangswaarden,

$\mathcal{U}$ (resp. $\mathcal{Y}$) is de verzameling van funkties

$u : T \to U$ ($y : T \to Y$) en wordt de ingangsverzameling (uitgangsverzameling) genoemd. We nemen aan dat $\mathcal{U}(\mathcal{Y})$ gesloten is onder *concatenatie* d.w.z. als $u_1,u_2 \in \mathcal{U}$, $t' \in T$ dan is

dan is $\bar{u} \in \mathcal{U}$ waarbij

$$\bar{u}(t) = \begin{cases} u_1(t) & t < t' \\ u_2(t) & t \geq t' \end{cases}$$

(idem voor $\mathcal{Y}$).

(iii) $F : \mathcal{U} \to \mathcal{Y}$ de *systeemfunktie*.

F voldoet aan de volgende eigenschappen:

niet-anticiperend d.w.z. als $u_1,u_2 \in \mathcal{U}$ met $u_1(t) = u_2(t)$ voor $t \leq t'$ ($t' \in T$) dan geldt $(Fu_1)(t') = (Fu_2)(t')$.

$F : \mathcal{U} \to \mathcal{Y}$ is een *lineaire afbeelding*.

OPMERKINGEN.

(i) Uit het feit dat F niet-anticiperend is volgt dat het verleden van de uitgang niet afhangt van de toekomst en het heden van de ingang. (causaliteitsprincipe)

(ii) Al naar gelang we werken met $T = \mathbb{R}$ of $T = \mathbb{Z}$ spreken we van (tijds)-kontinue of diskrete systemen.

(iii) Een I/O systeem heet *tijdsinvariant* (ook wel stationáir genoemd) als geldt $u \in \mathcal{U} \Rightarrow S_t u \in \mathcal{U}$ $(t \in T)$ waarbij $(S_t u)(t') = u(t+t')$ en $y \in \mathcal{Y} \Rightarrow S_t y \in \mathcal{Y}$ en er geldt dat

$$S_t F = F S_t.$$

DEFINITIE 1.5.

Een *lineair dynamisch systeem met toestandsruimte* is gedefinieerd door $\sum_M := \{T,U,\mathcal{U},Y,\mathcal{Y},\mathcal{X},\phi,r\}$, waarbij

(i) $T,U,\mathcal{U},Y,\mathcal{Y}$ als in definitie 1.4.

(ii) $\mathcal{X}$ vektorruimte over $\mathbb{R}$: de toestandsruimte.

(iii) ϕ is de *toestandsevolutiefunktie*,

$\phi : T_+^2 \times X \times U = X \qquad T_+^2 = \{(t_1,t_2) \in T^2 \mid t_2 \geq t_1\}$

voldoet aan de volgende eisen:

1. (konsistentie) $\phi(t,t,x,u) = x$,
2. (semi-groepseigenschap) $\phi(t_2,t_1,\phi(t_1,t_0,x_0,u),u) = \phi(t_2,t_0,x_0,u)$.
3. (determinisme) als $u_1,u_2 \in U$ met $u_1(t) = u_2(t)$ voor $t_0 \leq t < t_1$ dan is $\phi(t_1,t_0,x_0,u_1) = \phi(t_1,t_0,x_0,u_2)$.
4. ϕ is lineair.

(iv) $r : X \times U \times T \to Y$ is de *uitleesfunktie*,
r is lineair.

- $\phi(t_1,t_0,x_0,u)$ is de toestand die bereikt wordt op tijdstip t_1 door in toetoestand x_0 te starten op tijdstip t_0 en de ingangsfunktie u toe te passen.

OPMERKINGEN.

(i) $\sum_M$ wordt *tijdsinvariant* (*stationair*) genoemd als geldt

$\forall u \in U \quad S_t u \in U \quad en \quad \forall y \in Y \quad S_t u \in Y$

en bovendien

$\phi(t_1+t,t_0+t,x_0,u) = \phi(t_1,t_0,x_0,S_t u)$

terwijl: $r(x,u,t)$ niet expliciet afhangt van t.

(ii) Als $U = \{0\}$, dus geen mogelijkheid tot besturing, dan spreken we van een *autonoom* systeem; de eigenschappen waaraan ϕ dan moet voldoen komen dan precies overeen met de eigenschappen van een *flow* van een differentiaal- of differentievergelijking.

(iii) De systemen die beschreven worden door definitie 1.4 respektievelijk definitie 1.5 zijn in zekere zin de eenvoudigst denkbare. Toch bestaan er simpele voorbeelden, die niet binnen dit kader vallen (zie opgave 2); we zullen hier verder geen aandacht besteden aan deze *niet-lineaire* systemen.

Voorbeeld 1.3 is een illustratie van definitie 1.4, terwijl de voorbeelden 1.1 en 1.2 vallen onder het hoofdje lineair dynamisch systeem met toestandsruimte. (Nagaan.)

Zoals we uit de voorbeelden tot nu toe hebben gezien wordt een systeem vaak beschreven door een differentie- of differentiaalvergelijking. We zullen dat iets nader uitwerken.

(Diskrete tijd)

DEFINITIE 1.6.

Een systeem Σ_M in *rekursieve vorm* wordt beschreven door $\Sigma_M = \{T,U,\mathcal{U},Y,\mathcal{Y},X,f,r\}$ waarbij $T = \mathbb{Z}$, $U,\mathcal{U},Y,\mathcal{Y},X$ en r als in definitie 1.5 en $f : X \times U \times T \to X$ de *toestandsovergangsfunktie*, die voldoet aan de eigenschap dat voor alle $t \in T$ de funktie $f(\cdot,\cdot,t) : X \times U \to X$ lineair is.

OPMERKING.

Uit de lineariteit van f volgt dat

(*) $f(x,u,k) = A_k x + B_k u$, waarbij A_k, B_k matrices van geschikte dimensie zijn.

Voor gegeven $k_0, k_1 \in \mathbb{Z}$ met $k_1 > k_0$, $x_0 \in X$ en $u \in \mathcal{U}$ kunnen we dan $x_1 \in X$ rekursief definiëren:

$$x(k+1) = A_k\, x(k) + B_k\, u(k) \qquad x(k_0) = x_0$$
$$k = k_0,\dots\dots,k_1-1 \qquad x(k_1) := x_1.$$

Uiteraard induceert de afbeelding f een toestandsevolutiefunktie ϕ als in definitie 1.5, terwijl ook omgekeerd een toestandsevolutiefunktie ϕ aanleiding geeft tot een toestandsovergangsfunktie

$$f(x,u,k) = \phi(k+1,k,x,\bar{u}) \text{ voor een } \bar{u} \in \mathcal{U} \text{ met } \bar{u}(k) = u.$$

Het zij tenslotte nog opgemerkt dat indien we ook nog aannemen dat we met een een stationair systeem aan het werk zijn, dat dan (*) de vorm krijgt

$$f(x,u,k) = Ax + Bu.$$

(Kontinue tijd)

DEFINITIE 1.7.

Een *differentiaalsysteem* Σ_M wordt beschreven door $\Sigma_M = \{T,U,\mathcal{U},Y,\mathcal{Y},X,f,r\}$ waarbij $T = \mathbb{R}$, $U,\mathcal{U},Y,\mathcal{Y},X$ en r als in definitie 1.5 en de *snelheidsfunktie* $f : X \times U \times T \to X$ die voldoet aan de volgende eigenschappen:

(i) Voor alle $t \in T$ is de funktie $f(\cdot,\cdot,t) : X \times U \to X$ lineair.

(ii) Voor elke $x_0 \in X$, $u \in \mathcal{U}$ en $t_0 \in T$ bestaat er precies één funktie $x(\cdot) : T \to X$ met $x(t_0) = x_0$ en die voldoet aan de differentiaalvergelijking $\dot{x}(t) = f(x(t),u(t),t)$.

OPMERKING.

Uit de lineariteit (i) van f volgt dat $f(x,u,t) = A(t)x + B(t)u$, waarbij $A(t)$ en $B(t)$ matrices van geschikte dimensies zijn. Eis (ii) impliceert dat er een unieke oplossing bestaat van de differentiaalvergelijking $\dot{x}(t) = A(t)x(t) + B(t)u(t)$, $x(t_0) = x_0$. Analoog aan wat we voor diskrete systemen gezien hebben, volgt ook hier weer dat een snelheidsfunktie f een toestandsevolutiefunktie ϕ induceert (en omgekeerd).

VOORBEELD 1.8. (kapitaalgroei)

Solow's wet voor de groei van het geïnvesteerde kapitaal per arbeider luidt

$$\dot{k} = v\, f(k) - \lambda k$$
$$y = k$$

waarbij $k = K/L$, K : totale investeringen, L : aantal arbeiders,

f : produktiviteitsfunktie; f(k) geeft de produktie per arbeider bij geïnvesteerd kapitaal k per arbeider.

λ : groeifactor van L ($\dot{L} = \lambda L$), λ = constant,

v : neiging tot sparen; de fraktie van de produktie die wordt geïnvesteerd. We kunnen v als input interpreteren. (N.B. $v \in [0,1]$!) □

OPMERKING.

We zullen ons in de volgende hoofdstukken vrijwel alleen nog maar bezighouden met tijdsinvariante systemen. De theorie voor deze systemen is eenvoudiger te begrijpen, terwijl toch de meeste ingrediënten voor niet-stationaire systemen aanwezig zijn. In hoofdstuk 6 zullen we ook niet-stationaire systemen onder de loep nemen.

OPGAVEN.

1. Bewijs dat de flow van een gewone lineaire differentiaalvergelijking een lineair dynamisch systeem met toestandsruimte is.
2. Een investeringsmaatschappij belegt geld in een N-tal fondsen en ontvangt daarover per jaar per kapitaalseenheid rente. Probeer dit als een systeem te formaliseren.
3. Bewijs dat uit stationariteit van het autonome systeem $x_{k+1} = A_k x_k$ volgt dat de matrices A_k niet van k afhangen.

HOOFDSTUK 2

LINEAIRE SYSTEMEN MET TOESTANDSRUIMTE

We zullen in dit hoofdstuk onderscheid maken tussen diskrete-tijd en kontinue-tijd systemen. Enige kennis van differentie- en differentiaalvergelijkingen vereenvoudigt de theorie aanzienlijk.

2.1. Diskrete systemen met toestandsruimte

We gaan definitie 1.5 nader specificeren:

$T = \mathbb{Z}$,
$U = \mathbb{R}^m$, $Y = \mathbb{R}^p$,
$\mathcal{U}$ = alle rijtjes $\{u_k\}_{-\infty}^{\infty}$ met $u_k \in \mathbb{R}^m$ en $u_k = 0$ voor $k \leq k_1$,
$\mathcal{Y}$ = " " $\{y_k\}$ " $y_k \in \mathbb{R}^p$ en $y_k = 0$ voor $k \leq k_1'$,
$\mathcal{X} = \mathbb{R}^n$.

Analoog als in definitie 1.6 krijgen we in *rekursieve* vorm (N.B. stationair)

$$\begin{cases} x_{k+1} = Ax_k + Bu_k & (\rightarrow \text{definieert } f \text{, en dus } \phi) \\ y_k = Cx_k + Du_k & (\rightarrow \text{definieert } r) \end{cases}$$
$\forall k \in \mathbb{Z}$.

A : (n,n)-matrix, B : (n,m)-matrix, C : (p,n)-matrix,
D : (p,m)-matrix.

We kunnen eenvoudig de toestandsevolutiefunktie berekenen:
Voor $k_1 > k_0$ krijgen we

$$x_{k_1} = A^{k_1-k_0} x_{k_0} + \sum_{\ell=k_0}^{\ell=k_1-1} A^{k_1-\ell-1} B u_\ell$$

en dus

$$y_{k_1} = C A^{k_1-k_0} x_{k_0} + C \sum_{k_0}^{k_1-1} A^{k_1-\ell-1} B u_\ell + D u_k .$$

OPMERKING.

Bovenstaande evolutiefunktie laat zich gemakkelijk uitbreiden voor niet-stationaire systemen. Omdat we in hoofdstuk 6 hiermee te maken krijgen zullen we de oplossing hier geven:

Laat

$$x_{k+1} = A_k x_k + B_k u_k, \qquad k \in \mathbb{Z},$$

waarbij

$$A_k : (n,n)\text{-matrix}, \qquad k \in \mathbb{Z}$$
$$B_k : (n,m)\text{-matrix}, \qquad k \in \mathbb{Z}.$$

Voor $k > \ell$ is dan

$$x_k = A_{k-1} \cdot A_{k-2} \cdot \ldots A_\ell x_\ell + A_{k-1} \ldots A_{\ell-1} B_\ell u_\ell + \ldots$$
$$\ldots + A_{k-1} B_{k-2} u_{k-2} + B_{k-1} u_{k-1}.$$

VOORBEELD 2.1.

Vraag en aanbod voor een bepaald goed geven aanleiding tot een differentievergelijking.

Stel de prijs/eenheid = p

vraag := d(p)

aanbod := s(p)

We veronderstellen

$$d(p) = p_0 - ap \quad \text{en} \quad s(p) = s_0 + bp \qquad (a, d_0, b > 0).$$

en als dynamische vergelijkingen:

$$\begin{cases} s(k+1) = s_0 + b\, p(k) \\ d(k+1) = d_0 - a\, p(k+1) \end{cases}$$

evenwicht als

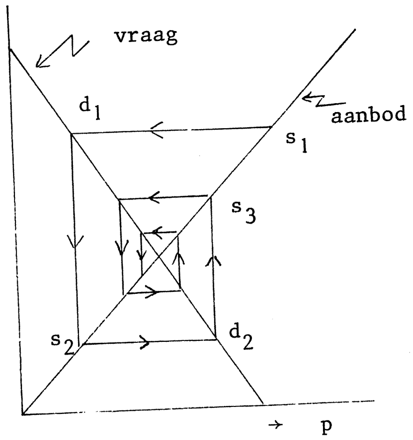
$$s_0 + b(k) = d_0 - a\, p(k+1) \qquad (*)$$

$$\Rightarrow \quad p(k+1) = -(\tfrac{b}{a})\, p(k) + \frac{d_0 - s_0}{a}$$

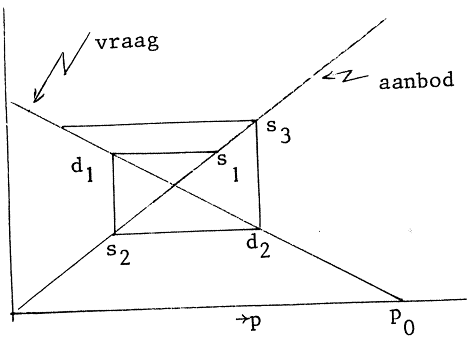
$$\Rightarrow \quad p(k) = \left(-\frac{b}{a}\right)^k \cdot p(0) + \frac{1 - \left(-\frac{b}{a}\right)^k}{a + b} (d_0 - s_0) \qquad (**)$$

(*) evenwichtsprijs $p(k+1) = p(k) \Rightarrow p = \dfrac{d_0 - s_0}{a + b}$.

Uit (**) volgt dat voor $b < a$ de prijs $p(k)$ voor $k \to \infty$ naar de evenwichtsprijs zal convergeren; voor $b > a$ treedt *geen* convergentie op.



$b < a$: de producent is minder gevoelig voor prijsschommelingen dan de konsument.



$b > a$: de producent is gevoeliger voor prijsschommelingen dan de konsument. □

2.2. Kontinue systemen met toestandsruimte

We nemen in definitie 1.5 (en definitie 1.7)

$T = \mathbb{R}$,
$U = \mathbb{R}^m$,
$Y = \mathbb{R}^p$,
$\mathcal{U}$: alle stuksgewijs kontinue funkties $u : \mathbb{R} \to \mathbb{R}^m$,
$\mathcal{Y}$ analoog,
$X = \mathbb{R}^n$.

Het stationaire geval reduceert dan tot

$$(*) \quad \begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases}$$

A : (n,n)-matrix   C : (p,n)-matrix
B : (n,m)-matrix   D : (p,m)-matrix.

De toestandsevolutiefunktie laat zich, met behulp van de transitiematrix, direkt uitrekenen (zie Appendix A).

STELLING 2.2. *De toestandsevolutiefunktie behorende bij (*) is éénduidig gedefiniëerd door de differentiaalvergelijking* $\dot{x} = Ax + Bu$ *en wordt gegeven door*

$$(**)\qquad x(t_1) = \phi(t_1,t_0,x_0,u) = e^{A(t_1-t_0)}x_0 + \int_{t_0}^{t_1} e^{A(t_1-\sigma)} B\,u(\sigma)d\sigma.$$

"BEWIJS"

1) Dat deze oplossing aan de beginwaarde voldoet is triviaal (invullen $t_1 = t_0$).
2) Dat de oplossing aan $\dot{x}(t) = A\,x(t) + B\,u(t)$ voldoet volgt rechtstreeks door $\dot{x}(t)$ uit te rekenen.
3) Uniciteit volgt uit de theorie van gewone differentiaalvergelijkingen. □

OPMERKING.

De formule (**) wordt wel de variatie van konstanten formule genoemd (zie gewone differentiaalvergelijkingen).

VOORBEELD 2.3. (= voorbeeld 1.1)

Werken we voorbeeld 1.1 op bovenstaande manier uit dan hebben we in dat geval

$$U = \mathbb{R},$$
$$Y = \mathbb{R},$$
$$X = \mathbb{R}^2,$$
$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad C = (1\ 0).$$

Bij een gegeven beginwaarde $\begin{pmatrix} x_1(t_0) \\ x_2(t_0) \end{pmatrix} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix}$ krijgen we:

$$\begin{pmatrix} x_1\ (t_1) \\ x_2\ (t_1) \end{pmatrix} = e^{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}(t_1-t_0)} \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} + \int_{t_0}^{t_1} e^{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}(t_1-\sigma)} \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(\sigma)d\sigma$$

$$= \begin{pmatrix} 1 & t_1-t_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} + \int_{t_0}^{t_1} \begin{pmatrix} 1 & t_1-\sigma \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(\sigma)d\sigma$$

$$= \begin{pmatrix} \bar{x}_1 + (t_1-t_0)\bar{x}_2 \\ \bar{x}_2 \end{pmatrix} \begin{pmatrix} \int_{t_0}^{t_1} (t_1-\sigma)u(\sigma)d\sigma \\ \int_{t_0}^{t_1} u(\sigma)d\sigma \end{pmatrix}. \qquad \square$$

Het is direkt duidelijk dat als we eenmaal de variatie van konstanten formule hebben, dat we dan ook onmiddellijk de uitleesfunktie kunnen geven:

$$y(t_1) = r(x(t_1),u(t_1),t_1) = r(x(t_1),u(t_1)) =$$
$$= C\, e^{A(t_1-t_0)} x_0 + C \int_{t_0}^{t_1} e^{A(t_1-\sigma)} B\, u(\sigma)d\sigma + D\, u(t_1).$$

<u>OPGAVEN</u>.

1. Beschouw het diskrete systeem $x_{k+1} = Ax_k + Bu_k$ met

$$A = \begin{pmatrix} 0 & 1 & 0 & & & O \\ 0 & 0 & 1 & 0 & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & 0 \\ & & O & & & 1 \\ & & & & & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

   Bepaal, voor willekeurige x_0, x_k $(k \geq 0)$.

2. Laat

$$A = \begin{pmatrix} 1 & 6 \\ 2 & 5 \end{pmatrix}.$$

   Bepaal e^{At}.

3. Stel

$$A = \begin{pmatrix} \alpha_1 & & O \\ & \ddots & \\ O & & \alpha_n \end{pmatrix}, \quad \alpha_1,\ldots,\alpha_n \in \mathbb{R}.$$

   Bepaal e^{At}.

   Wanneer is $\underline{x} = \underline{0}$ een stabiele oplossing voor $\dot{x} = Ax$?

HOOFDSTUK 3

REGELBAARHEID EN WAARNEEMBAARHEID

In dit hoofdstuk zullen we twee begrippen invoeren welke van fundamenteel belang zijn voor de lineaire systeemtheorie. Ze spelen een essentiële rol in het construeren van *regelwetten*. (Een onderwerp dat in hoofdstuk 8 aan de orde zal komen). We zullen de theorie behandelen voor tijdskontinue systemen; voor tijdsdiskrete systemen treden er geen wezenlijke veranderingen op.

We beschouwen het volgende lineaire systeem

$$(1) \qquad \begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases}$$

$x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$,
A : (n,n)-matrix, B : (n,m)-matrix en C : (p,n)-matrix.

De oplossing x van (1) met beginwaarde $x(0) = x_0$ geven we aan met $x(t,x_0,u)$ en analoog voor y : $y(t,x_0,u)$. Volgens stelling 2.1 is dan

$$x(t,x_0,u) = e^{At} x_0 + \int_0^t e^{A(t-\tau)} Bu(\sigma)d\sigma$$

en

$$y(t,x_0,u) = C e^{At} x_0 + C \int_0^t e^{A(t-\tau)} B u(\sigma)d\sigma.$$

DEFINITIE 3.1.
Het systeem (1) heet *regelbaar* (of *bestuurbaar*) als er voor elke x_0 en x_1 een $t_1 > 0$ en $u \in \mathcal{U}$ bestaat zodanig dat $x(t_1,x_0,u) = x_1$.

Een systeem is dus regelbaar als we vanuit een willekeurig punt x_0 elk willekeurig ander punt x_1 door middel van een geschikte besturing kunnen bereiken.

OPMERKING.

Soms wordt regelbaarheid gedefiniëerd als de mogelijkheid om vanuit een willekeurig punt x_0 de oorsprong te bereiken (dit met het oog op het feit dat dit vaak voor regelsystemen hetgeen is wat men wil bereiken). Men spreekt dan wel van *nulbestuurbaarheid*. Het "omgekeerde" begrip - vanuit de oorsprong kan men ieder willekeurig punt bereiken - noemt men *bereikbaarheid*. Het blijkt dat voor kontinue systemen deze drie begrippen volledig ekwivalent zijn. Voor tijdsdiskrete systemen zijn regelbaarheid en bereikbaarheid ekwivalent maar nulbestuurbaarheid is een zwakkere eigenschap.

DEFINITIE 3.2.

Het systeem (1) is *waarneembaar* (observeerbaar) als er een $t_1 > 0$ bestaat zodat voor willekeurige besturingsfunktie $u \in \mathcal{U}$, uit $y(t,x_0,u) = y(t,x_1,u)$ voor $0 \leq t \leq t_1$, volgt dat $x_0 = x_1$.

Een systeem is dus waarneembaar als we uit kennis van de ingang u en de uitgang y over een zeker tijdsinterval de toestand x_0 eenduidig kunnen bepalen.

OPMERKING.

Ook waarneembaarheid kan iets anders gedefiniëerd worden. Beschouw in plaats van het systeem beschreven door (1) het volgende systeem

$$(2) \qquad \begin{cases} \dot{x} = Ax \\ y = Cx \end{cases}$$

dat uit (1) ontstaat door de ingangsfunktie identiek nul te nemen. Het systeem (2) is waarneembaar als we uit kennis van de uitgangsfunktie $y(t)$, $0 \leq t \leq t_1$, de begintoestand x_0 uniek kunnen bepalen. Ofwel als $y(t,x_0) = y(t,x_1)$ voor $0 \leq t \leq t_1$ dan geldt $x_0 = x_1$. Hoewel niet direkt evident, is dit begrip van waarneembaarheid ekwivalent met dat van definitie 3.2!

Zoals te verwachten valt, kunnen we regelbaarheid en waarneembaarheid van (1) volledig uitdrukken in termen van de systeem-matrices (A,B,C). (Een lineair systeem van de vorm (1) wordt daarom ook wel aangeduid met (A,B,C).)

Definieer: $R := [B \vdots AB \vdots \ldots \vdots A^{n-1}B]$, dus R is een (n,mn)-matrix samengesteld uit de n blokken A^jB $j = 0,\ldots,n-1$. R wordt de *regelbaarheidsmatrix* genoemd (zie stelling 3.4).

LEMMA 3.3. $\forall n_0 \geq n$ is $\text{rang}[B \vdots AB \vdots \ldots \vdots A^{n_0-1}B] = \text{rang } R$

BEWIJS. Laat $R^k := [B \vdots \ldots A^{k-1}B]$, dan geldt $\text{rang}(R^{k+1}) = \text{rang}(R^k)$ of $\text{rang}(R^{k+1}) > \text{rang}(R^k)$. Stel $k \in \mathbb{N}$ is z.d.d. $\text{rang}(R^{k+1}) = \text{rang}(R^k) \Rightarrow$ de kolommen van A^kB zijn lineair afhankelijk van de kolommen van $R^k \Rightarrow \exists (m,m)$-matrices $D_0,\ldots,D_{k-1}$ zodat $A^kB = BD_0 + ABD_1 + \ldots + A^{k-1}BD_{k-1}$.
Dus $A^{k+1}B = ABD_0 + \ldots + A^kBD_{k-1}$, met andere woorden de kolommen van $A^{k+1}B$ zijn afhankelijk van de kolommen van R^{k+1}.
Dus $\text{rang}(R^k) = \text{rang}(R^{k+1}) \Rightarrow \text{rang}(R^{k+1}) = \text{rang}(R^{k+2})$.
Hieruit, en uit het feit dat de rang van R ten hoogste gelijk is aan n, volgt dat $\forall n_0 \geq n$ $\text{rang}[B \vdots AB \vdots \ldots \vdots A^{n_0-1}B] = \text{rang } R$. $\square$

STELLING 3.4. *Het systeem* (1) *is regelbaar* $\Leftrightarrow$ *de regelbaarheidsmatrix* $R = [B \vdots AB \vdots \ldots A^{n-1}B]$ *heeft rang* n.

BEWIJS. We zullen eerst aantonen dat de stelling geldt voor het geval dat we nemen $x_0 = 0$ in definitie 3.1.
Veronderstel dat de rangkonditie niet geldt. Voor elke ingang $u(t)$, $0 \leq t \leq t_1$, geldt

$$x(t_1,0,u) = \int_0^{t_1} e^{A(t_1-\sigma)} B\, u(\sigma)d\sigma$$

$$= \int_0^{t_1} \{I + A(t_1-\sigma) + \frac{A^2}{2!}(t_1-\sigma)^2 + \ldots\} B\, u(\sigma)d\sigma$$

$$= B\int_0^{t_1} u(\sigma)d\sigma + AB\int_0^{t_1}(t_1-\sigma)u(\sigma)d\sigma + A^2B\int_0^{t_1}\frac{(t_1-\sigma)}{2!}u(\sigma)d\sigma$$

$$+ \ldots$$

dus $x(t_1,0,u)$ is een lineaire kombinatie van de kolommen van $B, AB, A^2B, \ldots$.
Op basis van lemma 3.3 zien we dat, als $\text{rang}(R) = \text{rang}[B \vdots AB \ldots \vdots A^{n-1}B] < n$, er een vektor $x_1 \in \mathbb{R}^n$ bestaat die onafhankelijk is van de kolommen van R en dus ook nooit bereikt kan worden.

Laten we nu aannemen dat de rangkonditie wel vervuld is. We zullen laten zien dat we vanuit $x_0 = 0$ in willekeurig kleine tijd t_1 een arbitrair punt x_1 kunnen bereiken.
Laat $t_1 > 0$, definiëer:

$$W = \int_0^{t_1} e^{-A\sigma}\, BB^T\, e^{-A^T\sigma} d\sigma.$$

We gaan eerst aantonen dat W niet singulier is.
Stel $\exists \underline{a} \in \mathbb{R}^n$ met $W\underline{a} = 0$, dus

$$\underline{a}^T W \underline{a} = 0$$

$$\Rightarrow \int_0^{t_1} \underline{a}^T e^{-A\sigma} BB^T e^{-A^T\sigma} \underline{a} d\sigma = 0$$

$$\Rightarrow \quad \underline{a}^T e^{-At} B = 0 \ , \ \forall\ 0 \le t \le t_1 \ .$$

Dus

$$\begin{array}{ll} \underline{a}^T B = 0 & (t = 0) \\ \underline{a}^T AB = 0 & (\text{differentiëren, } t = 0) \\ \underline{a}^T A^2 B = 0 & (\text{nogmaals differentiëren, } t = 0) \\ \vdots & \vdots \\ \underline{a}^T A^{n-1} B = 0 & \end{array}$$

, $\text{rang}(R) = n \quad \Rightarrow \quad \underline{a} = 0 \quad \Rightarrow \quad W$ niet-singulier.
Voor gegeven x_1 en willekeurige $t_1 > 0$ definiëren we

$$\bar{u}(t) = B^T e^{-A^T t} W^{-1} e^{-At_1} x_1$$

$$\Rightarrow \quad x(t_1,0,\bar{u}) = \int_0^{t_1} e^{A(t_1-\sigma)} BB^T e^{-A^T\sigma} W^{-1} e^{-At_1} x_1 \, d\sigma$$

$$= e^{At_1} WW^{-1} e^{-At_1} x_1 = x_1 .$$

Tenslotte merken we op dat de keuze van $x_0 = \underline{0}$ in definitie 3.1 geen beperking is. Immers in het geval $x_0 \neq \underline{0}$ kunnen we altijd een ingangsfunktie $u(t)$, $0 \le t \le t_1$, vinden zodat

$$x_1 - e^{At_1} x_0 = \int_0^{t_1} e^{A(t_1-\sigma)} B\, u(\sigma) d\sigma$$

en dus $\quad x(t_1,x_0,u) = x_1 .$ $\square$

OPMERKING.
Het is voor de hand liggend dat regelbaarheid van het systeem (1) niet beïnvloed wordt door de output (outputmatrix C); men spreekt wel van regelbaarheid van het systeem (A,B).

VOORBEELD 3.5.

Een bank heeft een n-tal filialen, die we nummeren van 1 tot n, verspreid over de wereld. Alleen filiaal n heeft een economische binding met z'n omgeving. De overige filialen betrekken (leveren) geld aan elkaar volgens de volgende regel:

$$\begin{aligned} \dot{x}_1 &= \alpha_1 x_2 \\ \dot{x}_2 &= \alpha_2 x_3 \\ &\vdots \\ \dot{x}_{n-1} &= \alpha_{n-1} x_n \qquad\qquad \alpha_1, \ldots, \alpha_n \neq 0. \end{aligned}$$

(We hebben de hoeveelheid geld in filiaal i ook x_i genoemd.) Tenslotte kunnen we het rendement van het n^e filiaal voorstellen door $\dot{x}_n = u$

$$\Rightarrow \text{systeem} \quad \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix}' = \begin{pmatrix} 0 & \alpha_1 & & \\ & \ddots & \ddots & \bigcirc \\ & & \ddots & \alpha_{n-1} \\ \bigcirc & & & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} u$$

we zien dat $[B \vdots AB \vdots \ldots \vdots A^{n-1}B] = \begin{bmatrix} 0 & 0 & \alpha_1 \cdots \alpha_{n-1} \\ \cdot & & 0 \\ \cdot & & \\ \cdot & & \\ \cdot & 0 & \\ 0 & \alpha_{n-1} & \\ 1 & 0 & 0 \end{bmatrix}$

en dus is het systeem regelbaar. Voor de bankonderneming betekent dit, dat als ze zelf de besturingsfunktie ("rendement") volledig kan bepalen, dat ze dan op een willekeurig tijdstip $t_1 > 0$ een willekeurig bedrag $\bar{x}_i$ op filiaal i kan bezitten. □

OPMERKING.

Het bovenstaande voorbeeld is kunstmatig maar illustreert het begrip regelbaarheid goed. Vele ekonomische systemen hebben als extra eis dat i.p.v. $x \in \mathbb{R}^n$ er moet gelden $x \in \mathbb{R}^n_+$ (ook eventueel voor $u \in \mathbb{R}^m$ en $y \in \mathbb{R}^p$), waardoor verschillende begrippen aanzienlijk lastiger kunnen worden.

We zullen vervolgens een analoog resultaat af gaan leiden met betrekking tot waarneembaarheid. (Zie ook opmerking na definitie 3.2.)

We definiëren $W := \text{Ker}\left(\begin{bmatrix} C \\ CA \\ \vdots \\ \vdots \\ CA^{n-1} \end{bmatrix}\right)$ de *niet-waarneembare ruimte* van (1).

Analoog aan stelling 3.4 valt nu de volgende stelling te bewijzen. We zullen dat hier achterwege laten.

STELLING 3.6. *Het systeem* (1) *is waarneembaar* $\Longleftrightarrow$ *De niet-waarneembare ruimte* $\equiv \{\underline{0}\}$.

OPMERKING.
We zien dat waarneembaarheid van (A,B,C) niet afhangt van B. We spreken dan ook wel van waarneembaarheid van het paar (C,A). Hierbij hoort dan eigenlijk het systeem:

$$\begin{cases} \dot{x} = Ax \\ y = Cx \end{cases},$$

Gevolg: (A,B) is regelbaar $\Longleftrightarrow$ (B^T, A^T) is waarneembaar.

We willen dit hoofdstuk besluiten met enige opmerkingen over diskrete-tijd systemen. De definities van regelbaarheid en waarneembaarheid van diskrete-tijd systemen zijn analoog aan die van kontinue systemen. Ook de voorwaarden voor regelbaarheid en waarneembaarheid zijn gelijk aan die voor kontinue systemen:

STELLING 3.7. *Het systeem* $x_{k+1} = Ax_k + Bu_k$ *is regelbaar* $\Longleftrightarrow$ rang $[B \vdots AB \vdots \ldots A^{n-1}B] = n$.

STELLING 3.8. *Het systeem* $x_{k+1} = Ax_k + Bu_k$, $y_k = Cx_k$ *is waarneembaar*

$$\Longleftrightarrow \text{Ker}\left(\begin{bmatrix} C \\ CA \\ \vdots \\ \vdots \\ CA^{n-1} \end{bmatrix}\right) = \{\underline{0}\}.$$

VOORBEELD 3.9.

Beschouw het systeem

$$\begin{cases} x_{k+1} = Ax_k \\ y_k = Cx_k \end{cases}$$

en veronderstel dat x_0 en $\bar{x}_0$ ($\bar{x}_0 \neq x_0$) dezelfde uitgang genereren. Dan geldt

dus $\quad y_0 = Cx_0 = C\bar{x}_0$

dus $\quad x_0 - \bar{x}_0 \in \operatorname{Ker} C$

$\quad y_1 = CAx_0 = CA\bar{x}_0$

dus $\quad x_0 - \bar{x}_0 \in \operatorname{Ker} CA \qquad$ en in het algemeen

$\quad y_k = CA^k x_0 = CA^k \bar{x}_0$

ofwel $\quad x_0 - \bar{x}_0 \in \operatorname{Ker} CA^k$.

Samenvattend krijgen we dan dat

$$x_0 - \bar{x}_0 \in \operatorname{Ker} \left(\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \right),$$

ofwel de waarneembaarheidskonditie is inderdaad niet vervuld. □

OPGAVEN.

1. Laat $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

   Bepaal een ééndimensionale b-vektor zodanig dat (A,b) regelbaar is.
2. Bewijs dat R de kleinste A-invariante deelruimte is die de kolommen van B bevat. (Een lineaire deelruimte V is A-invariant als geldt $AV \subset V$.)
3. Kontroleer of de voorbeelden 1.1 en 1.2 regelbaar en waarneembaar zijn.
4. Konstrueer een discreet systeem dat wel nulbestuurbaar maar niet regelbaar is. N.B. Een triviaal voorbeeld heeft de gewenste eigenschap.
5. Bewijs dat (C,A) waarneembaar is d.e.s.d.

   als
$$\operatorname{rang} \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = n.$$

HOOFDSTUK 4

LINEAIRE INGANGS-UITGANGSSYSTEMEN

Er bestaan een aantal verschillende manieren om een lineair dynamisch systeem in Input-Output vorm voor te stellen. We zullen dit weer afzonderlijk doen voor tijdskontinue en tijdsdiskrete systemen.

4.1. Het tijdsdiskrete geval

Ons uitgangspunt is in dit geval (zie definitie 1.4) een lineair dynamisch systeem in I/O-vorm $\sum_{I/O} = \{T, U, \mathcal{U}, Y, \mathcal{Y}, F\}$ waarbij

$T = \mathbb{Z}$,

$U = \mathbb{R}^m$,

$Y = \mathbb{R}^p$,

$\mathcal{U}$ = alle rijtjes $\{u_k\}_{-\infty}^{\infty}$ met $u_k \in \mathbb{R}^m$ die nul zijn op $-\infty$ (d.w.z. iedere ingangsrij $\{u_k\}$ heeft de eigenschap dat er een $k_0 \in \mathbb{Z}$ bestaat z.d.d. $u_k = 0$ voor alle $k \leq k_0$),

$\mathcal{Y}$ = alle rijtjes $\{y_k\}_{-\infty}^{\infty}$ met $y_k \in \mathbb{R}^p$ die nul zijn op $-\infty$,

$F : \mathcal{U} \to \mathcal{Y}$ een lineaire niet-anticiperende afbeelding. Verder zullen we weer aannemen dat $\sum_{I/O}$ stationair is.

We gaan eerst de uitgang berekenen die wordt veroorzaakt door de ingang $\{u\,\delta_{kl}\}_{k \in \mathbb{Z}}$ $(\ell \in \mathbb{Z}\,,\ u \in \mathbb{R}^m)$.
Hierbij is

$$\delta_{kl} = 1 \quad \text{voor } k = \ell$$
$$\quad = 0 \quad\quad\quad k \neq \ell.$$

Op grond van de lineariteit weten we dat

$$y_k = W_{kl}\, u,$$

waarbij W_{kl} een (p,m)-matrix is, die afhangt van het tijdstip k waarop we de uitgang berekenen èn van het tijdstip ℓ waarop de ingangswaarde niet nul is. Uiteraard geldt $W_{k\ell} = 0$ voor $k < \ell$. (F niet-anticiperend.) Vanwege de lineariteit krijgen we dat een willekeurige ingangsrij $\{u_k\}_{k\in\mathbb{Z}}$ een uitgangsrij $\{y_k\}_{k\in\mathbb{Z}}$ genereert waarbij:

$$y_k = \sum_{\ell=-\infty}^{k} W_{k\ell}\, u_\ell.$$

N.B. Deze som is altijd gedefiniëerd! (Waarom?)
We zien dus dat de afbeelding F bepaald wordt door de matrices $W_{k\ell}$. Tot nu toe hebben we het stationair zijn van F nog niet gebruikt. Stel

$F : \{u_k\} \mapsto \{y_k\}$, dan geldt $\forall N \in \mathbb{Z}$

$F : \{u_{k+N}\} \mapsto \{y_{k+N}\}$

$$\Rightarrow \quad \sum_{-\infty}^{k} W_{k\ell}\, u_{\ell+N} = \sum_{\ell=-\infty}^{k+N} W_{k+N,\ell} u_\ell$$

$$= \sum_{\ell'=-\infty}^{k} W_{k+N,\ell'+N}\, u_{\ell'+N}$$

$$\Rightarrow \quad W_{k\ell} = W_{k+N,\ell+N}$$

$$\Rightarrow \quad W_{k,\ell} = W_{k-\ell,0}.$$

Notatie $\quad G_k := W_{k,0}$.
Het rijtje matrices $\{G_k\}_0^\infty$ bepalen het systeem volledig; de bijbehorende systeemvergelijking reduceert tot:

$$y_k = \sum_{\ell=-\infty}^{k} G_{k-\ell}\, u_\ell.$$

OPMERKING.
De matrices $\{G_k\}_0^\infty$ worden de *Markov parameters* van het systeem genoemd.

VOORBEELD 4.1.
Een standaardvoorbeeld van een systeem in Ingangs-Uitgangsvorm is een produktieproces (in een fabriek of een chemisch produktieproces). We introduceren 2 (1-dimensionale) grootheden :

u : de ingang (grondstoffen),

y : de uitgang (produkten).

We geven een ingangsrij weer aan met $\{u_k\}_{k\in\mathbb{Z}}$ en de uitgangsrij weer met $\{y_k\}_{k\in\mathbb{Z}}$.
Definieer

$$\begin{cases} G_{2k} = 2 & k = 1,2,\ldots \qquad G_0 = 0. \\ G_{2k-1} = 1 & k = 1,2,\ldots \end{cases}$$

De (1,1)matrices $\{G_k\}_{k \in \mathbb{N}}$ bepalen het produktieproces volledig. □

Er bestaat nog een geheel andere methode om een diskreet systeem te beschrijven (zie Appendix B). In plaats van te werken met Markov-parameters $\{G_k\}_0^\infty$ kijken we dan naar de z-getransformeerde van deze rij matrices. We nemen daarom aan dat $\{G_k\}_0^\infty$ slechts geometrische groei heeft, dus

$$\sum_{k=0}^{\infty} \|G_k\| r^{-k} < \infty \text{ voor een } r \in \mathbb{R}^+ .$$

We definiëren dan de *overdrachtsfunktie*

$$(*) \quad \hat{G}(z) := \sum_{k=0}^{\infty} G_k z^{-k}.$$

<u>OPMERKING.</u>
De z-getransformeerde van een rij matrices is de matrix die verkregen wordt door voor elk der matrix-elementen de bijbehorende z-transformatie uit te voeren.

De vergelijking (*) geeft de systeembeschrijving in het *frekwentiedomein*, dit in tegenstelling tot de beschrijving door de rij $\{G_k\}_0^\infty$ die zich afspeelt in het tijdsdomein.

Zonder bewijs vermelden we het belangrijkste resultaat voor een diskreet systeem in het frekwentiedomein.

<u>STELLING 4.2.</u> *Als voor een diskreet systeem* $\Sigma_{I/O}$ *de Markov parameters* $\{G_k\}_0^\infty$ *en de ingang* $\{u_k\}$ *geometrische groei hebben, dan geldt dit ook voor de uitgangsrij* $\{y_k\}$ *en bovendien*

$$\hat{y}(z) = \hat{G}(z)\hat{u}(z).$$

<u>OPMERKING.</u>
Bovenstaande formule geldt natuurlijk alleen op het gebied waar zowel $\{G_k\}_0^\infty$ en $\{u_k\}_0^\infty$ konvergeren (zie ook Appendix B).

<u>VOORBEELD 4.3.</u> (= Voorbeeld 4.1)
We gaan de overdrachtsfunktie van voorbeeld 4.1 bepalen. Het konvergentiegebied zullen we daarbij en passant bepalen.

$$\begin{aligned}\hat{G}(z) &= \sum_{k=0}^{\infty} G_k\, z^{-k}\\ &= \frac{1}{z} + \frac{2}{z^2} + \frac{1}{z^3} + \frac{2}{z^4} + \dots\\ &= \frac{1}{z}\left(1 + \frac{1}{z^2} + \frac{1}{z^4} + \dots\right) + \frac{2}{z^2}\left(1 + \frac{1}{z^2} + \frac{1}{z^4} + \dots\right).\end{aligned}$$

Nu geldt voor $|z^2| > 1$ dat

$$1 + \frac{1}{z^2} + \frac{1}{z^4} + \dots = \frac{1}{1 + \frac{1}{z^2}} = \frac{z^2}{z^2-1}\,.$$

Dus voor $|z^2| > 1$ is

$$\hat{G}(z) = \left(\frac{1}{z} + \frac{2}{z^2}\right)\frac{z^2}{z^2-1} = \frac{z+2}{z^2-1}\,. \qquad \square$$

Tenslotte zullen we een derde, ekwivalente, beschrijving geven voor een diskreet systeem $\sum_{I/O}$, nl. die welke beschreven worden door een differentievergelijking.
Beschouw:

$$(**) \qquad P_r\, y_{k+r} + P_{r-1}\, y_{k+r-1} + \dots + P_0 \cdot y_k = Q_r\, u_{k+r} + \dots + Q_0\, u_k,$$

waarbij

$P_0, P_1, \dots, P_r$: (p,p)-matrices,
$Q_0, Q_1, \dots, Q_r$: (p,m)-matrices,

en de extra voorwaarde $\det(P_r) \neq 0$.

<u>STELLING 4.4.</u> *De differentievergelijking* (**) *definieert een Ingangs-Uitgangs-systeem met Overdrachtsfunktie*

$$\hat{G}(z) = (P_r\, z^r + P_{r-1}\, z^{r-1} + \dots + P_0)^{-1}\,(Q_r\, z^r + \dots + Q_0).$$

(*Uiteraard moet* $|z| > r_a$ *voor e.e.a.* $r_a \in \mathbb{R}^+$.)

VOORBEELD 4.5. (= Voorbeeld 4.3 = Voorbeeld 4.1)
We kunnen nu eenvoudig bepalen wat een bijbehorende differentievergelijking is voor het produktieproces dat beschreven is in voorbeeld 4.1. Uit voorgaande stelling en uit voorbeeld 4.3 volgt meteen dat $y_{k+3} - y_k = u_{k+1} + 2u_k$. □

OPMERKINGEN.
(i) Een basiseigenschap van de overdrachtsfunktie voor een lineair systeem is, dat $\hat{G}(z)$ een rationale funktie is. Dit wil zeggen dat ieder van de matrixelementen uit $\hat{G}(z)$ te schrijven als een quotient $p(z)/q(z)$, waarbij p en q polynomen in z. Bovendien geldt dat de graad van p kleiner of gelijk is aan de graad van q.
(ii) In de ekonometrie wordt een Ingangs-Uitgangs-systeem beschreven door (**) een *rekursief* model genoemd als $P_r = I$. Als P_r niet singulier is (en eventueel $P_{r-1} = \ldots = P_0 = 0$) spreekt men wel van een *simultaan* model. Het model is *overgeïdentificeerd* als P_r een singuliere matrix is.

OPGAVEN.
1. Laat
$$\hat{G}(z) = \begin{pmatrix} \frac{z}{z^3+1} & 0 \\ 0 & \frac{1}{z} \end{pmatrix}.$$
Bepaal de Markov-parameters die bij $\hat{G}(z)$ horen.
Wat is de bijbehorende differentievergelijking?
2. Beschouw de differentievergelijking
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}(k+2) + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}(k+1) + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}(k) =$$
$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}(k+2) + \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}(k+1).$$
Bepaal de overdrachtsfunktie.

4.2. Het tijdskontinue geval

Laten we nu aannemen dat $\sum_{I/O}$ gemodelleerd is op de volgende wijze:

$T = \mathbb{R}$,

$U = \mathbb{R}^m$,

$Y = \mathbb{R}^p$,

$\mathcal{U}$ = alle stuksgewijs kontinue funkties $u : \mathbb{R} \to \mathbb{R}^m$ die nul zijn op $-\infty$
(Dus voor iedere ingangsfunktie $u(\cdot)$ bestaat er een t_0 zodat $u(t) = 0$ voor alle $t \leq t_0$).

$\mathcal{Y}$ = Analoog gedefiniëerd.

$F : \mathcal{U} \to \mathcal{Y}$ wordt gedefiniëerd door

$$y(t) = \sum_{k=0}^{\infty} W_k(t)u(t-t_k) + \int_{-\infty}^{t} W(t,\tau)u(\tau)d\tau.$$

(Merk op dat F lineair is).

Met $0 = t_0 < t_1 < \ldots < t_k < \ldots \quad \lim_{k\to\infty} t_k = \infty$,

$W_k(\cdot) : \mathbb{R} \to \mathbb{R}^{p\times m}$ stuksgewijs kontinu en

$W(\cdot,\cdot) : \mathbb{R}_+^2 \to \mathbb{R}^{p\times m}$ stuksgewijs kontinu.

OPMERKING.

Elke term $W_k(t)u(t-t_k)$ $(k > 0)$ geeft een tijdsvertraging van de ingangsfunktie naar de uitgangsfunktie weer.

Uiteraard vereenvoudigt de afbeelding F aanzienlijk als we weer de stationariteitsvoorwaarde aannemen. Er volgt dan, na een eenvoudige berekening, dat $W_k(\cdot)$ onafhankelijk is van t en $W(t,\tau) = W(t+T,\tau+T)$ $\forall t,\tau,T$ ofwel $W(t,\tau) = W(t-\tau,0)$. Stel $G_k := W_k(t)$ en $G(t) := W(t,0)$. Dan is

$$y(t) = \sum_{k=0}^{\infty} G_k\, u(t-t_k) + \int_{-\infty}^{t} G(t-\tau)u(\tau)d\tau.$$

OPMERKINGEN.

(i) Voorbeeld 1.3 is een illustratie van een dergelijk systeem.

(ii) Laat $\delta(t)$ de *Dirac deltafunktie* zijn, d.w.z. $\delta(t) = 0$ voor $t \neq 0$ en $\int_{-\infty}^{\infty} \delta(t)dt = 1$. (Deze wat slordige definitie van Dirac deltafunktie kan wiskundig zonder veel problemen geformaliseerd worden.)
Dan wordt $\sum_{k=0}^{\infty} G_k\, \delta(t-t_k) + G(t)$, $t \geq 0$ de *impulsresponsie* van $\sum_{I/O}$ genoemd. (Hierbij is $G(t)$ een (p,m)-matrix.)

We zullen van nu af aan verder aannemen dat de Ingangs-Uitgangsafbeelding F geen tijdsvertragingen bevat; dus $W_k \equiv 0$. De Ingangs-Uitgangsrepresentatie wordt dan:

$$y(t) = G_0\, u(t) + \int_{-\infty}^{t} G(t-\tau)u(\tau)d\tau$$

en $\sum_{I/O}$ heeft als impulsresponsie $G_0\delta(t) + G(t)$.

Inplaats van de representatie van $\sum_{I/O}$ in het tijdsdomein, kunnen we analoog aan het diskrete geval weer een andere representatie maken. Hiervoor gaan we over op de Laplace getransformeerde van de impulsresponsie. We veronderstellen daarom dat de impulsresponsie slechts exponentiële groei heeft (zie Appendix B), dus $\exists\sigma \in \mathbb{R}$ met

$$\int_0^{\infty} \|G(t)\| e^{-\sigma t}\, dt < \infty.$$

De *overdrachtsfunktie* van $\sum_{I/O}$ is dan

$$\hat{G}(s) := G_0 + \int_0^{\infty} G(t)e^{-st}\, dt.$$

OPMERKING.
De Laplace getransformeerde van een matrixfunktie is weer komponentsgewijs gedefiniëerd.

VOORBEELD 4.6. (= Voorbeeld 1.2)
Laten we aannemen dat de overlevingskans $P(x,t)$ niet afhangt van t. In dat geval is het systeem beschreven door

$$N(t) = \int_{-\infty}^{t} P(t-\tau)B(\tau)d\tau$$

stationair. $(P(x) := P(x,t))$.
Laten we nu aannemen dat bij benadering $P(t) = e^{-t}$. Dan is de overdrachtsfunktie van dit systeem

$$\int_0^{\infty} P(t)e^{-st}\, dt = \int_0^{\infty} e^{-t-st}\, dt = \frac{1}{-1-s}\, e^{-t-st}\Big|_0^{\infty} = \frac{1}{1+s}\,,$$

waarbij we voor de laatste gelijkheid aannemen dat $Re(s) \geq -1$; dit definieert meteen het konvergentiegebied van de overdrachtsfunktie. (Merk overigens op dat deze overdrachtsfunktie een zogeheten analytische voortzetting bezit, waardoor hij bijna overal gedefiniëerd is.) □

De overdrachtsfunktie $\hat{G}(s)$ van een lineair systeem $\sum_{I/O}$ geeft de systeembeschrijving in het *frekwentiedomein*. (Let op de volledig analoge terminologie die er bestaat tussen tijdsdiskrete en tijdskontinue systemen.) Een belangrijk resultaat voor de overdrachtsfunktie is het volgende:

STELLING 4.7. *Als de impulsresponsies van* $\sum_{I/O}$ *en de ingang exponentiële groei hebben dan is dat ook het geval voor de uitgang. Bovendien geldt* $\hat{y}(s) = \hat{G}(s)\hat{u}(s)$.

OPMERKING.
Bovenstaande formule geldt alleen op het gebied waarop zowel $\hat{G}(s)$ als $\hat{u}(s)$ goed gedefiniëerd zijn.

BEWIJS.

$$y(t) = G_0 u(t) + \int_{-\infty}^{t} G(t-\tau)u(\tau)d\tau$$

$$\Rightarrow \quad \hat{y}(s) = \int_{-\infty}^{\infty} G_0 u(t)\ e^{-st}\ dt + \int_{-\infty}^{\infty}\int_{-\infty}^{t} G(t-\tau)u(\tau)e^{-st}d\tau dt$$

$$\Rightarrow \quad \hat{y}(s) = \int_{-\infty}^{\infty} G_0\ u(t)e^{-st}\ dt + \int_{-\infty}^{\infty}\int_{-\infty}^{t} G(t-\tau)e^{-s(t-\tau)}u(\tau)e^{-s\tau}d\tau dt$$

$$\Rightarrow \quad \hat{y}(s) = G(s)\hat{u}(s).$$

□

Tenslotte zullen we laten zien dat we ook de lineaire systemen die beschreven worden door differentiaalvergelijkingen op soortgelijke wijze kunnen representeren.

Laat gegeven zijn het systeem beschreven door de differentiaalvergelijkingen:

$$(*) \qquad P_r \frac{d^r y}{dt^r} + P_{r-1}\frac{d^{r-1}y}{dt^{r-1}} + \dots + P_0 y = Q_r \frac{d^r u}{dt^r} + \dots + Q_0\ u$$

waarbij $P_0, P_1 \dots P_r$: (p,p)-matrices,
$Q_0, Q_1 \dots Q_r$: (p,m)-matrices,
en $\det(P_r) \neq 0$.

STELLING 4.8. *De differentiaalvergelijking* (*) *definieert een Ingangs-Uitgangssysteem met als overdrachtsfunktie*

$$\hat{G}(s) = (P_r\ s^r + P_{r-1}\ s^{r-1} + \dots + P_0)^{-1}\ (Q_r\ s^r + \dots + Q_0).$$

(Met Re s $\geq \sigma$ *voor een* $\sigma \in \mathbb{R}$, *zie echter ook de opmerking aan het eind van voorbeeld* 4.6.)

OPMERKING.
De overdrachtsfunktie is ook nu weer een rationale funktie in s. Het zal duidelijk zijn dat er zekere ekwivalenties tussen de drie representaties bestaan. We zullen daar in dit diktaat verder niet op in gaan.

OPGAVEN.
1. Bepaal een differentiaalvergelijking, die aanleiding geeft tot dezelfde overdrachtsfunktie als in voorbeeld 4.6.
2. Stel

$$G(t) = \begin{pmatrix} e^{\alpha t} & e^{2\alpha t} \\ 0 & e^{t} \end{pmatrix}.$$ Bepaal $\hat{G}(s)$.

   Kunt u ook een differentiaalvergelijking bij deze inpulsresponsie konstrueren?

HOOFDSTUK 5

REALISATIETHEORIE

Eén van de belangrijkste elementen van de systeemtheorie - en overigens niet alleen van belang voor de systeemtheorie - is de realisatietheorie. Zoals we in voorgaande hoofdstukken al hebben gezien geeft de beschrijving van een systeem met toestandsruimte automatisch aanleiding tot een Ingangs-Uitgangssysteem. (Schematisch $\sum_M \Rightarrow \sum_{I/O}$.) Realisatietheorie houdt zich bezig met de omgekeerde vraag: Kunnen we een (lineair) Ingangs-Uitgangssysteem beschrijven door een (lineair) systeem met toestandruimte? (Ofwel: $\sum_{I/O} \Rightarrow \sum_M$?) Het belang hiervan is dat men voor een systeem in Ingangs-Uitgangsvorm om de uitgang $y(t)$ ten gevolge van een ingangsfunktie $u(t)$ te berekenen de volledige informatie van de funktie $u(t)$, $t \le t_1$, moet gebruiken. Daarentegen wordt deze procedure voor een systeem met toestandsruimte veel eenvoudiger als men de toestand op een tijdstip t_0, $t_0 < t_1$, weet; dan namelijk heeft men slechts de informatie van de ingangsfunktie $u(t)$ voor $t_0 \le t \le t_1$ nodig om $y(t_1)$ te berekenen. "Het verleden van de ingangsfunktie $u(t)$, $t < t_0$, wordt volledig samengevat in de (begin)toestand $x(t_0)$".
We zullen in dit hoofdstuk zien dat precies de Ingangs-Uitgangssystemen uit het vorige hoofdstuk een *realisatie* tot een systeem met toestandsruimte kunnen hebben. Het zal zelfs blijken dat deze I/O systemen een *minimale realisatie* hebben, d.w.z. er is een realisatie $\sum_M$ voor $\sum_{I/O}$, waarvan de toestandsruimte minimale dimensie heeft. Intuïtief betekent dit dat informatie van een niet-minimaal systeem $\sum_M$ ook verkregen kan worden uit een "kleiner" systeem $\sum_{M'}$ (met een toestandsruimte $\mathcal{X}'$ waarvoor geldt dat $\dim \mathcal{X}' < \dim \mathcal{X}$).

OPMERKING.
Men kan op eenvoudige wijze een lineair Input-Output systeem realiseren via de *triviale realisatie* . In dat geval neemt men als toestandsruimte de verledens van de ingangen - d.w.z. $\mathcal{X} = \mathcal{U}^- = \{u : (-\infty,0) \to U \mid \exists v \in \mathcal{U} \text{ met } u(t) = v(t)\}$. Het zal duidelijk zijn dat dit een systeem met toestandsruimte (∞-dimensionaal) genereert. Een minimale realisatie krijgt men door die

(verleden) ingangen u_1 en u_2 uit U^- te identificeren welke eenzelfde toekomstige uitgang voortbrengen. We zullen hier verder niet op deze abstrakte realisatietheorie ingaan.

5.1. Het tijdsdiskrete geval

We gaan uit van een lineair Ingangs-Uitgangssysteem zoals beschreven in hoofdstuk 4.2:

$$\Sigma_{I/O} : y_k = \sum_{\ell=-\infty}^{k} G_{k-\ell}\, u_\ell \qquad G_k : (p,m)\text{-matrix.}$$

We zoeken een realisatie van de vorm

$$\Sigma_M : \begin{cases} x_{k+1} = Ax_k + Bu_k \\ y_k = Cx_k + Du_k \end{cases}$$

De Markov-parameters van Σ_M zijn eenvoudig te bepalen met behulp van de toestandsevolutiefunktie voor Σ_M zoals we dat in hoofdstuk 2 hebben gezien. Willen we dat Σ_M en $\Sigma_{I/O}$ hetzelfde Ingangs-Uitgangsgedrag hebben dan zullen de Markov-parameters van $\Sigma_{I/O}$ en Σ_M gelijk moeten zijn. (Waarom?) We vinden dan dat

$$D = G_0$$
$$CA^{k-1}B = G_k \qquad k = 1,2,\dots .$$

Het realisatieprobleem bestaat dus hieruit:
Zoek bij gegeven Markov-parameters $\{G_k\}_0^\infty$ matrices A, B, C en D zodanig dat aan bovenstaande vergelijkingen is voldaan. Merk op dat de dimensie van de toestandsruimte nog niet bekend is! We zullen nu een procedure geven welke de matrices A, B, C en D genereert, en wel zodanig dat (A,B) regelbaar en (C,A) waarneembaar is.
Definieer de volgende (∞,∞)-matrix:

$$H := \begin{bmatrix} G_1 & G_2 & G_3 & \cdots\cdots & G_\ell & \cdots\cdots \\ G_2 & G_3 & G_4 & \cdots\cdots & G_{\ell+1} & \cdots \\ G_3 & \cdot & \cdot & \cdot\;\cdot\;\cdot & \vdots & \\ \vdots & & & & \vdots & \\ G_k & G_{k+1} & \cdots\cdots & & G_{k+\ell-1} & \cdots \\ \vdots & \vdots & & & \vdots & \end{bmatrix}$$

Hierbij zijn de elementen van H de (p,m)-matrices G_k. H wordt de *Hankel matrix* van $\Sigma_{I/O}$ genoemd.

Laat verder

$$H_{k,\ell} =: \begin{bmatrix} G_1 & G_2 & \cdots\cdots & G_\ell \\ G_2 & G_3 & \cdots\cdots & G_{\ell+1} \\ \vdots & \vdots & & \vdots \\ G_k & G_{k+1} & \cdots\cdots & G_{k+\ell-1} \end{bmatrix}$$

We definiëren tenslotte:

$$\text{rang } H = \sup_{k,\ell} \text{ rang } H_{k,\ell} = \lim_{k,\ell\to\infty} \text{ rang } H_{k,\ell}.$$

We nemen aan dat rang $H = n < \infty$. Dan geldt:

<u>STELLING 5.1.</u> *Laat* $\{G_k\}_0^\infty$ *de rij Markov-parameters van een lineair Ingangs-Uitgangssysteem* $\sum_{I/O}$ *zijn. Dan bestaat er een lineair systeem* $\sum_M$, *beschreven door matrices* A, B, C *en* D, *dat* $\sum_{I/O}$ *realiseert met de volgende eigenschappen*

(i) dim X = rang H = n.

De realisatie is <u>minimaal</u> , d.w.z.

(ii) (A,B) *is regelbaar.*

(iii) (C,A) *is waarneembaar.*

We zullen geen bewijs van deze stelling geven; we vermelden slechts dat het bewijs is gebaseerd op de faktorisatie:

$$H_{k,\ell} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix} [B \; AB \; \ldots \; A^{\ell-1}B]$$

Dat een realisatie van een rij Markov-parameters $\{G_k\}_0^\infty$ minimaal gekozen kan worden - mits er tenminste een realisatie bestaat - kunnen we als volgt inzien. Bij de relatie tussen $\sum_{I/O}$ en de realisatie $\sum_M$ wordt er vanuit gegaan dat de begintoestand gelijk is aan nul. Daaruit volgt dat alleen de toestanden, die vanuit nul bereikbaar zijn - de bereikbare verzameling - van belang zijn voor $\sum_{I/O}$. Als $\sum_M$ niet bereikbaar is, dan is de toestandsruimte kennelijk te groot voor het doel waarvoor zij is ingevoerd, namelijk een geheugen om de relevante informatie uit het verleden te onthouden. Een soortgelijke redenering geldt ook met betrekking tot waarneembaarheid. Als de realisatie $\sum_M$ niet waarneembaar is, dan bestaan er toestanden $x_0 \neq 0$ die bij een willekeurige ingangsfunktie $u(\cdot)$ dezelfde uitgangsfunktie $y(\cdot)$ genereren als de begintoestand 0. Dergelijke toestanden bevatten informatie

over de verleden ingangen die niet relevant is voor de toekomst. We zien de minimaliteit van $\sum_M$ fraai terug in de hierboven gegeven faktorisatie van $H_{k,\ell}$: regelbaarheid impliceert dat de matrix $[B\vdots AB\vdots \ldots \vdots A^nB]$ surjektief is, en uit waarneembaarheid volgt dat $[C\vdots CA \ldots \vdots CA^{-1}]^T$ injektief is.
We weten nu dat, als de matrices $\{A,B,C,D\}$ de rij $\{G_k\}_0^\infty$ realiseren, dat dan geldt: $D = G_0$ en $CA^{k-1}B = G_k$ $(k = 1,2,\ldots)$. Rest ons tenslotte de vraag hoe we de matrices A, B, C en D kunnen bepalen. Veronderstel dat de Markov-parameters $\{G_k\}_0^\infty$ gegeven zijn en dat de bijbehorende Hankel-matrix eindige rang heeft. (In principe is dit een oneindige test!) Dan hebben we het volgende realisatie-algoritme. (N.B. Er bestaan vele realisatiealgoritmen.)

STELLING 5.2. *Silverman's algoritme.*

Stap 1: Bepaal k en ℓ zodat rang H = rang $H_{k,\ell} = n$.
Stap 2: Bepaal een niet-singuliere (n,n)-submatrix F van $H_{k,\ell}$, die bestaat uit de elementen van de $r_1^e,\ldots,r_n^e$ rij en $s_1^e,\ldots,s_n^e$ kolom van $H_{k,\ell}$.
Stap 3: Bepaal

$$\bar{H}_{k,\ell} := \begin{bmatrix} G_2 & G_3 & \cdots\cdots G_{\ell+1} \\ G_3 & G_4 & \cdots\cdots G_{\ell+2} \\ \vdots & \vdots & \vdots \\ G_{k+1} & \cdots\cdots & G_{k+\ell+1} \end{bmatrix}$$

en bepaal $\bar{F}$: de (n,n)-submatrix van $\bar{H}_{k,\ell}$ die bestaat uit de elementen van de $r_1^e,\ldots,r_n^e$ rij en $s_1^e,\ldots,s_n^e$ kolom van $\bar{H}_{k,\ell}$.
Bepaal de (n,m)-submatrix F_1 van $H_{k,\ell}$ die bestaat uit de elementen van de $r_1^e,\ldots,r_n^e$ rij en de $1^e,\ldots,m^e$ kolom van $H_{k,\ell}$.
Bepaal de (p,n)-submatrix F_2 van $H_{1,\ell}$ die bestaat uit de $s_1^e,\ldots,s_n^e$ kolommen van $H_{1,\ell}$.
Dan is

$$A = \bar{F}F^{-1}$$
$$B = F_1$$
$$C = F_2\,F^{-1}$$
$$D = G_0,$$

een realisatie die voldoet aan de eisen van stelling 5.1. □

We zullen bovenstaande procedure toepassen op het voorbeeld uit hoofdstuk 4.

<u>VOORBEELD 5.3.</u> (= Voorbeeld 4.1)

Laat $\{G_k\}_{k \in \mathbb{N}}$ gegeven zijn door

$$\begin{cases} G_0 = 0 & \\ G_{2k} = 2 & k = 1,2,\dots \\ G_{2k-1} = 1 & k = 1,2,\dots \qquad \text{dus } p = m = 1. \end{cases}$$

Dan is

$$H = \begin{pmatrix} 1 & 2 & 1 & 2 & 1 & \cdots \\ 2 & 1 & 2 & 1 & 2 & \cdots \\ 1 & 2 & 1 & 2 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

rang H = rang $H_{2,2} = 2$

$$F = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \qquad (\text{dus } r_1 = 1,\ r_2 = 2,\ s_1 = 1,\ s_2 = 2)$$

$$\overline{H}_{2,2} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \qquad \overline{F} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$F_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$F_2 = (1\ \ 2).$$

Dus is

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \cdot \frac{1}{3} \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$C = (1\ \ 2) \cdot \frac{1}{3} \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix} = (1\ \ 0)$$

$$D = 0.$$

□

We hebben hiermee het realisatieprobleem opgelost. Men zou kunnen denken dat stelling 5.1 een unieke realisatie definieert. Hiermee bedoelen we dat de matrices A, B, C en D uniek zijn als aan de eisen (i), (ii) en (iii) voldaan is. Dit is *niet* het geval. (Dit werd al gesuggereerd met de opmerking dat er verschillende realisatie-algoritmen bestaan.) Toch is er mooi verband aan te geven tussen twee minimale realisaties:

STELLING 5.4. *Stel* $\Sigma_M = \{A,B,C,D\}$ *en* $\tilde{\Sigma}_M = \{\bar{A},\bar{B},\bar{C},\bar{D}\}$ *zijn twee minimale realisaties van hetzelfde lineaire Ingangs-Uitgangssysteem* $\Sigma_{I/O}$. *Dan bestaat er een inverteerbare matrix* S *zodanig dat*

$$A = S\bar{A}S^{-1}$$
$$B = S\bar{B}$$
$$C = \bar{C}S^{-1}$$
$$D = \bar{D}\ .$$

OPMERKING.

Men spreekt in dit geval van isomorfe realisaties. Als we uitgaan van het systeem

$$\begin{cases} x_{k+1} = \bar{A}x_k + \bar{B}u_k \\ y_k = \bar{C}x_k + \bar{D}u_k \end{cases}$$

en we passen de transformatie $\tilde{x} = Sx$ toe dan krijgen we voor $\tilde{x}$ de volgende vergelijking

$$\begin{cases} \tilde{x}_{k+1} = S\bar{A}S^{-1}\tilde{x}_k + S\bar{B}u_k \\ y_k = \bar{C}S^{-1}\tilde{x}_k + Du_k. \end{cases}$$

Dus voor de isomorfe realisaties hebben we een koördinatentransformatie op de toestandsruimte toegepast.

OPGAVEN.

1. Stel we hebben een lineair Ingangs-Uitgangssysteem beschreven door de differentievergelijking $y_{k+2} - y_k = u_{k+1} + 2u_k$.
   Construeer een minimale realisatie.
2. (zie ook hoofdstuk 4.1, opgave 1.)
   Bepaal een minimale realisatie voor $\Sigma_{I/O}$, welke als overdrachtsfunktie heeft

$$\hat{G}(z) = \begin{pmatrix} \frac{z}{z^3+1} & 0 \\ 0 & \frac{1}{z} \end{pmatrix}.$$

5.2. Het tijdskontinue geval

Aangezien dit vrijwel analoog verloopt zullen we dit slechts kort beschrijven. We gaan nu uit van een lineair Ingangs-Uitgangssysteem $\Sigma_{I/O}$ zoals beschreven in hoofdstuk 4.2:

$$\Sigma_{I/O} : y(t) = \sum_{k=0}^{\infty} G_k\, u(t-t_k) + \int_{-\infty}^{t} G(t-\tau)u(\tau)d\tau.$$

We zoeken een realisatie van de vorm

$$\Sigma_M : \begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases}$$

Analoog aan het tijdsdiskrete geval moeten we, opdat Σ_M hetzelfde ingangs-uitgangsgedrag zal vertonen, er voor zorgen dat de impulsresponsie van Σ_M en $\Sigma_{I/O}$ gelijk zijn. Met behulp van stelling 2.1 volgt daaruit dat

$G_k = 0, \quad k \geq 1$ en

$$D = G_0$$
$$C\, e^{At}B = G(t) \qquad t \geq 0.$$

A priori ziet dit er nog wat lastiger uit dan in het tijdsdiskrete geval. Echter definieer de *geassocieerde Markov-parameters*

$$\tilde{G}_k := \frac{d^{k-1}}{dt^{k-1}} G(t) \Big|_{t=0} \qquad k = 1,2,\ldots .$$

Het realisatieprobleem reduceert dan tot het realisatieprobleem voor tijdsdiskrete systemen met als Markov-parameters de geassocieerde Markov-parameters $\{\tilde{G}_k\}_0^\infty$.

Stelling 5.1 blijft dan ook voor tijdskontinue systemen volledig gelden.

OPMERKINGEN.

(i) Wezenlijk is dat $G_k = 0$, $k \geq 1$. (Niet $G_0 = 0$!)

(ii) De realisatie die we op deze manier krijgen is dus weer minimaal.

(iii) Voor een praktisch probleem levert de procedure nogal wat problemen daar we te maken krijgen met afgeleiden van experimenteel vastgestelde grootheden.

Tenslotte merken we nog op dat stelling 5.4 direkt analoog waar is voor kontinue Ingangs-Uitgangssystemen.

HOOFDSTUK 6

DISKRETE STOCHASTISCHE SYSTEMEN EN HET KALMANFILTER

De systemen die we tot nog toe behandeld hebben zijn alle *deterministisch* van aard; bijvoorbeeld voor een lineair Ingangs-Uitgangssysteem kunnen we, als we een bepaalde ingangsrij precies weten, de korresponderende uitgangswaarden precies bepalen. Een probleem hierbij is echter dat vanuit een praktisch oogpunt we nooit de waarde van variabelen precies kunnen bepalen. (meetfouten!) Vaak kan men volstaan met de deterministische beschrijving; de uitkomsten geven een voldoend nauwkeurige beschrijving van de werkelijkheid. (Men moet hierbij wel bedenken dat een systeem nooit de werkelijkheid weergeeft; het kan slechts dienen als een mathematisch model voor de realiteit.) Er bestaat echter ook een methode om de onzekerheid in de systeembeschrijving formeel te definiëren: de *stochastische systeemtheorie*. Slordig gezegd komt deze methode erop neer dat alle grootheden in stochastische zin volledig bepaald zijn. (Zie Appendix C voor een beknopte inleiding op begrippen en notaties uit de waarschijnlijkheidsrekening en de stochastiek.)

6.1. Stochastische systemen in diskrete tijd

Beschouw de differentievergelijking

(*) $x_{k+1} = A_k x_k + G_k v_k \qquad k \in \mathbb{N}$, met beginwaarde x_0,

waarbij: A_k : (n,n)-matrix, $\forall k \in \mathbb{N}$,

G_k : (n,m)-matrix, $\forall k \in \mathbb{N}$,

$\{v_k\}$ is een vektorwaardige witte ruis proces (op $\mathbb{R}^m$).

$\{v_k\}$ is een Gaussisch proces, met als kovariantie $E[v_k v_\ell^T] = Q_k \, \delta_{k\ell}$, $\forall k,\ell \in \mathbb{N}$;

x_0 is een Gaussische stochastische variabele met verwachtingswaarde $E[x_0] = \bar{x}_0$ en kovariantie $E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = P_0$ en x_0 is onafhankelijk van v_k, $k \in \mathbb{N}$.

We zien onmiddellijk dat we op deze manier een stochastisch proces $\{x_k\}$ gedefinieerd hebben. Er geldt echter meer; de oplossing $\{x_k\}$ van de differentievergelijking is in stochastische zin volledig bepaald.

Er geldt:

(i) x_k *is een Gaussische stochastische variabele,*
immers (zie hoofdstuk 2)

$$x_k = A_{k-1}\cdot\ldots\cdot A_0 x_0 + A_{k-1}\cdot\ldots\cdot A_1 G_0 v_0 + A_{k-1}\cdot\ldots\cdot A_2 G_1 v_1 + \ldots + A_{k-1} G_{k-2} v_{k-2} + G_{k-1} v_{k-1},$$

dus is x_k een lineaire kombinatie van de Gaussische variabelen $x_0, v_0, \ldots, v_{k-1}$. Uit de eigenschappen van $x_0, v_0, \ldots, v_{k-1}$ (Gaussisch, onderling onafhankelijk) volgt de bewering direkt. (Merk op dat een lineaire transformatie van Gaussische stochastische variabelen weer een Gaussische stochastische variabele is.)

(ii) Uit (i) volgt direkt dat
$\{x_k\}$ *is een Gaussisch stochastisch proces.*

(iii) $\{x_k\}$ *is een Markov-proces.*
Dit volgt uit het feit dat $\{v_k\}$ een witte ruis proces is en uit de causaliteit van het systeem (zie ook definitie 1.4).
Voor $k_1 < k_2 < \ldots < k_m < k$ geldt immers

$$x_k = A_{k-1} \ldots A_{k_m} x_{k_m} + A_{k-1} \ldots A_{k_m+1} G_{k_m} v_{k_m} + \ldots + G_{k-1} v_{k-1}$$

en $v_{k_m}, \ldots, v_{k-1}$ zijn alle onafhankelijk van $x_{k_1}, x_{k_2}, \ldots, x_{k_m}$. Dus

$$p(x_k \mid x_{k_1}, \ldots, x_{k_m}) = p(x_k \mid x_{k_m}).$$

We kunnen van het stochastische proces $\{x_k\}$ ook de verwachtingswaarde en kovariantie precies bepalen. Allereerst hebben we dat

$$E[x_k] = A_{k-1}\cdot\ldots\cdot A_0 \bar{x}_0 \quad (\text{ofwel } E[x_k] = A_{k-1} E[x_{k-1}])$$

en voor de kovariantie

$$P_{k,\ell} = E[(x_k - \bar{x}_k)(x_\ell - \bar{x}_\ell)^T]$$

krijgen we, na gebruikmaking van bovenstaande uitdrukking en de onafhankelijkheid van $x_0, v_0, \ldots, v_{k-1}$, dat

$$P_{k,\ell} = A_{k-1} \ldots A_\ell \{A_{\ell-1} \ldots A_0 P_0 (A_{\ell-1} \ldots A_0)^T + A_{\ell-1} \ldots A_1 G_0 Q_0 G_0^T (A_{\ell-1} \ldots A_1)^T + \ldots + A_{\ell-1} G_{\ell-2} Q_{\ell-2} G_{\ell-2}^T A_{\ell-1}^T + G_{\ell-1} Q_{\ell-1} G_{\ell-1}^T\}.$$

Hieruit volgt dat $P_{k,k}$ voldoet aan de differentievergelijking

$$P_{k+1,k+1} = A_k P_k A_k^T + G_k Q_k G_k^T,$$

en daaruit volgt dan omgekeerd weer

$$P_{k,\ell} = A_{k-1} \cdots A_\ell P_{\ell,\ell}.$$

We hebben dus gezien dat de oplossing $\{x_k\}$ van de differentievergelijking (*) als stochastisch proces uniek te bepalen is. Aldus hebben we een stochastische versie gegeven van een dynamisch systeem met toestand in rekursieve vorm. (Vergelijk met definitie 1.6.) Een verschil met het deterministische geval is dat de ingang in (*) (respektievelijk de begintoestand x_0) een stochastisch proces (respektievelijk stochastische variabele) is.

OPMERKINGEN.

(i) In de differentievergelijking (*) kunnen we ook nog een deterministische ingang toevoegen, $x_{k+1} = A_k x_k + B_k u_k + G_k v_k$, beginwaarde x_0, waarbij $A_k, G_k, \{v_k\}$ en x_0 als voorheen en
B_k : (n,s)-matrix, $\forall k \in \mathbb{N}$.
u_k : een (deterministisch) gegeven rij vektoren in $\mathbb{R}^s$.
De voorgaande beschouwing kan zonder problemen tot dit geval worden uitgebreid.

(ii) Het zal duidelijk zijn dat de stationaire versie van de gegeven stochastische differentievergelijking de beschrijving aanzienlijk vereenvoudigt. Het is echter in het algemeen niet zo dat de oplossing $\{x_k\}$ in dat geval een stationair Gaussisch proces is! (Zie opgave 1.) Aangezien we het Kalmanfilter ook voor niet stationaire systemen willen behandelen, hebben we ook hier de stationariteit niet verondersteld.

(iii) De gehele opzet verandert niet als we $k \in \mathbb{Z}$ nemen en een beginwaarde x_{k_0} vastleggen.

We zullen tenslotte ook een *uitgangsproces* associëren met de stochastische differentievergelijking

$$x_{k+1} = A_k x_k + G_k v_k, \qquad x_0$$

uitgang: $\qquad y_k = C_k x_k + w_k$

waarbij:

$A_k, G_k, \{v_k\}$ en x_0 als voorheen en
C_k : (p,n)-matrix , $\forall k \in \mathbb{N}$,
$\{w_k\}$ vektorwaardig ($\mathbb{R}^p$) witte ruis proces.
$\{w_k\}$ is Gaussisch met als kovariantie $E[w_k w_\ell^T] = R_k\, \delta_{k\ell}$, $\quad \forall k,\ell \in \mathbb{N}$,
en $\{v_k\}$ en $\{w_k\}$ zijn onderling onafhankelijke processen; x_0 onafhankelijk van v_k en w_k , $\forall k \in \mathbb{N}$.

Het is eenvoudig in te zien dat het stochastische uitgangsproces $\{y_k\}$ weer een Gaussisch proces is. $\{y_k\}$ is echter geen Markov-proces (in het algemeen). Ook van het stochastische proces $\{y_k\}$ kunnen we weer de verwachtingswaarde en kovariantie bepalen:

$$\bar{y}_k = E[y_k] = C_k\, E[x_k]$$

en analoog aan de kovariantie van $\{x_k\}$ vinden we dat

$$E[(y_k - \bar{y}_k)(y_k - \bar{y}_k)^T] = \begin{cases} C_k^T\, A_{k-1} \cdots A_\ell P_\ell C_\ell + R_k \delta_{k\ell}, & k \geq \ell \\ C_k^T\, P_k (A_{\ell-1} \cdots A_k)^T\, C_\ell + R_k \delta_{k\ell}, & k \leq \ell. \end{cases}$$

Hierbij is $P_{\ell,\ell}$ weer gegeven door

$$P_{\ell,\ell} = E[(x_\ell - \bar{x}_\ell)(x_\ell - \bar{x}_\ell)^T].$$

OPMERKINGEN.

(i) Voor het stationaire geval wordt de uitdrukking voor de kovariantie van $\{y_k\}$ weer vereenvoudigd, maar we zijn er opnieuw niet van verzekerd dat $\{y_k\}$ een stationair proces is.

(ii) De aanname dat de stochastische processen $\{v_k\}$ en $\{w_k\}$ onderling onafhankelijk zijn is niet essentieel. Ook zonder deze voorwaarde kunnen we de stochastische eigenschappen van het uitgangsproces afleiden.

We besluiten deze sektie met het geven van het stochastisch analogon van definitie 1.6.

DEFINITIE 6.1. *Een Gaussische stochastische dynamische systeemrepresentatie* (zonder ingangen) wordt gedefinieerd door een stochastische differentievergelijking en een daarbij behorend uitgangsproces. (Met uiteraard de passende vooronderstellingen.)

6.2. Het Kalmanfilter

We gaan uit van het stochastische dynamische systeem:

$$\begin{cases} x_{k+1} = A_k x_k + B_k v_k, & x_0 \\ y_k = C_k x_k + w_k \end{cases}$$

waarbij, zoals gebruikelijk:

A_k : (n,n)-matrix , $\forall k \in \mathbb{N}$,
B_k : (n,m)-matrix , $\forall k \in \mathbb{N}$,
C_k : (p,m)-matrix , $\forall k \in \mathbb{N}$,
$\{v_k\}$ en $\{w_k\}$ zijn vektorwaardige witte ruis processen op $\mathbb{R}^m$ respektievelijk $\mathbb{R}^p$.
$\{v_k\}$ en $\{w_k\}$ zijn onderling onafhankelijke Gaussische processen met $E[v_k v_\ell^T] = Q_k\ \delta_{k\ell}$ en $E[w_k w_\ell^T] = R_k\ \delta_{k\ell}$, $\forall k,\ell \in \mathbb{N}$,
x_0 is een Gaussische stochastische variabele met verwachtingswaarde $E[x_0] = \bar{x}_0$ en kovariantie $E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = P_0$.
x_0 is onafhankelijk van v_k en w_k, $k \in \mathbb{N}$.

Een voor de hand liggende vraag is nu: kunnen we op basis van de waarnemingen - het uitgangsproces - een schatting maken van de toestand? Bij het preciseren van deze vraag onstaat het volgende probleem:

Filterprobleem (één-staps voorspelprobleem)
Schat (= voorspel) de grootheid x_k op basis van de uitgangswaarden tot op tijdstip $k-1$; dus bepaal $\hat{x}_{k|k-1} := E[x_k | y_0, y_1, \ldots, y_{k-1}]$.

Uiteraard willen we, afgezien van het bepalen van een mogelijke oplossing van dit probleem, ook een uitspraak kunnen doen over de aard van de schatter. We voeren daartoe eerst de volgende notatie in:

$$Y_k := \{y_0, \ldots, y_k\}$$

$$\Sigma_{k|k-1} := E[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T | Y_{k-1}].$$

$\Sigma_{k|k-1}$ wordt de *foutkovariantiematrix* genoemd. Nu is het voor de hand liggend om te eisen dat we, indien mogelijk, een schatter $\hat{x}_{k|k-1}$ zoeken waarvoor de foutkovariantiematrix minimaal is. Het nu volgende resultaat toont aan dat het filterprobleem inderdaad een oplossing heeft, die de foutkovariantie minimaliseert! De oplossing wordt het *Kalmanfilter* genoemd en wordt gegeven door

STELLING 6.2. *De oplossing van het bovengenoemde filterprobleem wordt gegeven door:*

$$\begin{cases} \hat{x}_{k+1|k} = [A_k - K_k C_k]\, \hat{x}_{k|k-1} + K_k y_k \\ \hat{x}_{0|-1} = \bar{x}_0 \end{cases}$$

waarbij de matrix K_k *gedefinieerd wordt door*

$$K_k = A_k \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T + R_k]^{-1}$$

de inverse

$$[C_k \Sigma_{k|k-1} C_k^T + R_k]^{-1}$$

bestaat en de symmetrische foutkovariantiematrix $\Sigma_{k|k-1}$ *voldoet aan de zogenoemde matrix-Riccati vergelijking:*

$$\Sigma_{k+1|k} = A_k[\Sigma_{k|k-1} - \Sigma_{k|k-1} C_k^T (C_k \Sigma_{k|k-1} C_k^T + R_k)^{-1} C_k \Sigma_{k|k-1}]\, A_k^T + B_k Q_k B_k^T$$

$$\Sigma_{0|-1} = P_0.$$

OPMERKING.

In plaats van het hierboven beschouwde filterprobleem kunnen we ook een *vereffeningsprobleem* beschouwen. In dat geval wordt gevraagd de toestand x_k te schatten op basis van de uitgangswaarden tot aan het tijdstip k; ofwel bepaal $\hat{x}_{k|k} = E[x_k | y_0, \ldots, y_k]$. De oplossing verloopt analoog. Laat namelijk $\Sigma_{k|k} := E[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T | Y_k]$ dan wordt het bijbehorende filter gegeven door:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \Sigma_{k|k-1} C_k^T (C_k \Sigma_{k|k-1} C_k^T + R_k)^{-1} (y_k - C_k \hat{x}_{k|k-1})$$

waarbij

$$\Sigma_{k|k} = \Sigma_{k|k-1} - \Sigma_{k|k-1} C_k^T (C_k \Sigma_{k|k-1} C_k^T + R_k)^{-1} C_k \Sigma_{k|k-1}.$$

We zullen hier verder alleen nog maar de oplossing van het één-staps voorspelprobleem beschouwen; alles geldt mutatis mutandis ook voor het vereffeningsprobleem.

We zullen de belangrijkste eigenschappen van het Kalmanfilter geven:

1. Het Kalmanfilter is een lineair niet-stationair tijdsdiskreet dynamisch systeem met een eindig dimensionale toestandsruimte. De ingang van het

filter is het proces $\{y_k\}$, de uitgang is $\{\hat{x}_{k|k-1}\}$. De foutkovariantie-matrix is onafhankelijk van de inputs $Y_{k-1} = \{y_0, y_1, \ldots, y_{k-1}\}$, dus $\Sigma_{k|k-1} = E[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T]$.

2. De voorwaardelijke verwachtingswaarde van x_k gegeven Y_{k-1} is gelijk aan $\hat{x}_{k|k-1}$ met als voorwaardelijke kovariantie $\Sigma_{k|k-1}$. Het Kalmanfilter levert dus een procedure om de voorwaardelijke dichtheid van x_k te bepalen.
3. Het Kalmanfilter is niet tijdsinvariant; zelfs al was het oorspronkelijke systeem dat wel. (Dit is een van de redenen om het filter voor niet-stationaire systemen te geven.)
4. Het Kalmanfilter genereert de beste *lineaire* schatter $\hat{x}_{k|k-1}$. Dit betekent dat als we een andere schatter $\tilde{x}_{k|k-1}$ gekozen hadden, waarvoor geldt $\tilde{x}_{k+1|k} = (A_k - \tilde{K}_k C_k)\tilde{x}_{k|k-1} + \tilde{K}_k y_k$, voor willekeurige matrices $\tilde{K}_k$ dan voldoet de foutkovariantiematrix $\tilde{\Sigma}_{k|k-1}$ van $x_k - \tilde{x}_{k|k-1}$ ($\tilde{\Sigma}_{k|k-1} := E[(x_k - \tilde{x}_{k|k-1})(x_k - \tilde{x}_{k|k-1})^T]$) aan de eigenschap $\tilde{\Sigma}_{k|k-1} - \Sigma_{k|k-1}$ is semi-definiet positief. Dus het Kalmanfilter minimaliseert de foutkovariantie (tenminste als we ons beperken tot lineaire filters). Er geldt overigens wel $E[x_k - \tilde{x}_{k|k-1}] = 0$.

6.3. Het stationaire Kalmanfilter

We hebben tot nog toe geen aanname betreffende de stationariteit van het stochastische systeem gedaan, De belangrijkste reden hiervoor is dat, zoals we al gezien hebben, de oplossing van een stationair stochastisch dynamisch systeem in het algemeen niet stationair is. (Evenmin als de bijbehorende Kalmanfilter-schatter.) Daarnaast is de opbouw van het niet-stationaire Kalmanfilter volstrekt analoog aan de stationaire versie. We zullen nu een versie van het Kalmanfilter geven waarvoor de oplossing wel stationair is.

We gaan uit van het stochastische systeem

$$\begin{cases} x_{k+1} = Ax_k + Bv_k \\ y_k = Cx_k + w_k \end{cases}$$

waarbij:

$k \in \mathbb{Z}$, gegeven beginwaarde x_{k_0} ,

A : (n,n)-matrix,

B : (n,m)-matrix,

C : (p,m)-matrix,

$\{v_k\}$ is een stationair Gaussisch witte ruis proces met kovariantie $E[v_k v_\ell^T] = Q\ \delta_{k\ell},\ k,\ell \in \mathbb{Z},\ Q > 0$.
$\{w_k\}$ is een stationair Gaussisch witte ruis proces met kovariantie $E[w_k w_\ell^T] = R\ \delta_{k\ell},\ \forall k,\ell \in \mathbb{Z}$.
x_{k_0} is Gaussisch stochastische variabele met $E[x_{k_0}] = \bar{x}_{k_0}$ en $E[(x_{k_0} - \bar{x}_{k_0})(x_{k_0} - \bar{x}_{k_0})^T] = P_0$.
$\{v_k\}$ en $\{w_k\}$ zijn onderling onafhankelijk.
x_{k_0} en v_k (respectievelijk w_k) zijn onafhankelijk, $\forall k \in \mathbb{Z}$.
We veronderstellen bovendien dat (A,B,C) minimaal is; dus (A,B) is regelbaar en (C,A) is waarneembaar.

Dan geldt:

STELLING 6.3. (*Het stationaire Kalmanfilter*)
Onder bovenstaande voorwaarden geldt voor $k_0 \to -\infty$ *dat het bijbehorende Kalmanfilter stationair wordt. Dit filter wordt gegeven door*

$$\hat{x}_{k+1|k} = [A - KC]\ \hat{x}_{k|k+1} + Ky_k,$$

waarbij

$$K = A\ \Sigma\ C^T[C\ \Sigma\ C^T + R]^{-1}$$

en de foutkovariantie Σ *gegeven wordt door de enige, positief definiete oplossing van de algebraïsche matrix-Riccati vergelijking*

$$\Sigma = A\,[\Sigma - \Sigma\ C^T(C\ \Sigma\ C^T + R)^{-1}\ C\ \Sigma]\ A^T + BQB^T.$$

We zullen geen bewijs van deze stelling geven, maar een aantal eigenschappen van het stationaire Kalmanfilter geven.

1. In het "gewone" Kalmanfilter voldoet de foutkovariantiematrix $\Sigma_{k|k-1}$ aan de matrix-Riccati vergelijking

$$\begin{cases} \Sigma_{k+1|k} = A[\Sigma_{k|k-1} - \Sigma_{k|k-1}\ C^T(C\ \Sigma_{k|k-1}\ C^T + R)^{-1}\ C\ \Sigma_{k|k-1}]A^T + BQB^T \\ \Sigma_{k_0|k_0-1} = P_0 \end{cases}$$

De algebraïsche matrix-Riccati vergelijking volgt hieruit door een konstante matrix, namelijk Σ, te bepalen die aan deze vergelijking voldoet. Het verband tussen de rij $\{\Sigma_{k|k-1}\}$ en de matrix Σ wordt gegeven door

$$\lim_{k\to\infty} \Sigma_{k|k-1} = \Sigma$$

met andere woorden, de parameters van het niet-stationaire Kalmanfilter konvergeren naar die van het stationaire Kalmanfilter.

2. Een belangrijk voordeel van het stationaire Kalmanfilter is dat we van de stochastische grootheid x_{k_0}, het gemiddelde $\bar{x}_{k_0}$ en de kovariantie P_0 niet hoeven te kennen. In de praktijk is P_0 in het algemeen onbekend.
3. Minimaliteit van het drietal (A,B,C) is geen noodzakelijke voorwaarde voor het tijdsinvariante Kalmanfilter. Het is bijvoorbeeld al voldoende dat elke eigenwaarde λ van de matrix A voldoet aan $|\lambda| < 1$ (stabiliteit).
4. De matrix A − KC is stabiel: elke eigenwaarde λ van deze matrix voldoet aan $|\lambda| < 1$.

VOORBEELD 6.4.
Beschouw het stochastische dynamische systeem

$$x_{k+1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x_k + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} v_k$$

$$y_k = (1 \quad 0)\, x_k + w_k$$

met $E[v_k v_\ell^T] = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \delta_{k\ell}$, $\quad E[w_k w_\ell] = \delta_{k\ell}$.

Het bijbehorende stationaire Kalmanfilter kunnen we nu eenvoudig bepalen. De foutkovariantiematrix

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$$

voldoet dan aan

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \left\{ \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} - \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \left[(1 \quad 0) \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \right]^{-1} (1 \quad 0) \, . \right.$$

$$\left. \cdot \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \right\} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$$

ofwel

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} \sigma_{22} - \dfrac{\sigma_{12}^2}{1+\sigma_{11}} + 1 & 1 \\ 1 & 4 \end{pmatrix} .$$

Dus $\sigma_{12} = 1$

$\sigma_{22} = 4$

$\sigma_{11} = 2 + 2\sqrt{2}$ (Σ is positief definiet).

Het filter wordt dan gegeven door

$$K = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2+2\sqrt{2} & 1 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\left[(1\ \ 0)\begin{pmatrix} 2+2\sqrt{2} & 1 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1\right]^{-1} = \begin{pmatrix} \frac{1}{3+2\sqrt{2}} \\ 0 \end{pmatrix}$$

$$\hat{x}_{k+1|k} = \begin{pmatrix} \frac{-1}{3+2\sqrt{2}} & 1 \\ 0 & 0 \end{pmatrix} \hat{x}_{k|k-1} + \begin{pmatrix} \frac{1}{3+2\sqrt{2}} \\ 0 \end{pmatrix} y_k .$$

Het is bovendien eenvoudig in te zien dat het gewone Kalmanfilter inderdaad naar deze oplossing konvergeert. □

In het volgende hoofdstuk zullen we een andere toepassing geven van het Kalmanfilter. Daarbij zal het filter gebruikt worden om de systeem-matrices A,B en C optimaal te schatten.

OPGAVEN.

1. Beschouw het 1-dimensionale systeem

$$x_{k+1} = x_k + v_k \text{ met}$$
$$E[v_k v_\ell] = \delta_{k\ell} \text{ en } x_o = 0.$$

Bewijs dat $\{x_k\}$ *niet* stationair is.

2. Probeer een Kalmanfilter te bepalen voor het systeem

$$\begin{cases} x_{k+1} = A_k x_k + B_k v_k + G_k u_k \\ y_k = C_k x_k + w_k \end{cases}$$

waarbij A_k, B_k, C_k en $\{v_k\}$ en $\{w_k\}$ als in het filterprobleem en G_k is een (n,ℓ)-matrix voor elke k en $\{u_k\}$ is een deterministische ingangsrij.

3. Beschouw het 1-dimensionale systeem

$$x_{k+1} = ax_k$$
$$y_k = x_k + w_k, \qquad E[w_k w_\ell] = \delta_{k\ell}.$$

Bewijs dat voor $|a| < 1$ en voor willekeurige $\Sigma_{0|-1}$, de oplossing $\Sigma_{k+1|k}$ van de Riccati-vergelijking konvergeert naar de oplossing van de algebraïsche Riccati-vergelijking.

HOOFDSTUK 7

SYSTEEMIDENTIFIKATIE

In het vorige hoofdstuk hebben we gezien hoe we van een lineair Ingangs-Uitgangssysteem een systeem met toestandsruimte hebben kunnen maken. Van verzenlijk belang hierbij was de kennis van de Markov-parameters. Voor verschillende praktische problemen is het uitermate moeilijk om deze I/O - systeembeschrijving precies te krijgen. Zo kunnen we voor een ekonomisch/ekonometrisch model er niet vanuit gaan dat we ongehinderd de besturing kunnen manipuleren. Wat ons rest is een beperkte hoeveelheid informatie (bijvoorbeeld kwartaalcijfers van het Centraal Planburo) tussen uitgangswaarden. Bovendien kunnen we er niet vanuit gaan dat we bij het modelleren voldoende hebben aan deze gegevens (mogelijk spelen ook andere onbekende ingangen een rol), terwijl we uiteraard te maken hebben met meetfouten in ingangs- en uitgangsfunktie. Kunnen we nu toch op basis van de bestaande informatie tot een acceptabele systeembeschrijving komen? *Systeemidentifikatie* houdt zich bezig met dit probleem. (Een van de belangrijke vragen die hierbij uiteraard een grote rol speelt is de vraag: wat is een akseptabele systeembeschrijving?) We zullen ons hier kort met dit juist voor de ekonometrie zo belangrijke onderwerp bezig houden en zullen ons bovendien beperken tot tijdsdiskrete systemen. We zullen aandacht besteden aan de relatief eenvoudige aanpak van het identifikatieprobleem door middel van de (lineaire) kleinste kwadratenmethode. Bovendien zullen we het identifikatieprobleem omzetten in een filterprobleem en vervolgens het bijbehorende Kalmanfilter bepalen.

7.1. Kleinste kwadraten

Stel we hebben een stochastisch proces $\{Y_k, k \in (1,\ldots,N)\}$ waarvan de gemiddelde waarde $E[Y_k]$, een lineaire funktie is van een parameter vektor θ:

$$E[Y_k] = x_k^T \theta$$

waarbij x_k een bekende vektorwaardige grootheid ($k = 1,\dots,N$). Het doel is om een goede schatting van θ te doen (goed in de zin van de kleinste kwadraten) op basis van een realisatie van het stochastische proces $\{y_k,\ k = 1,\dots,N\}$. De oplossing is in feite heel simpel: Bepaal de waarde $\hat{\theta}$, die de som van de kwadraten van de deviaties tussen y_k en de gemiddelde waarde $x_k^T\,\theta$ minimaliseert. Dus minimaliseer

$$S = \sum_{k=1}^{N} (y_k - x_k^T\,\theta)^2.$$

Schrijf

$$Y := \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} \qquad X := \begin{pmatrix} x_1^T \\ \vdots \\ x_N^T \end{pmatrix}$$

Dan is $\quad S = (Y - X\,\theta)^T(Y - X\,\theta).$

Het minimum van S vinden we door naar θ te differentiëren; de waarde van θ waarvoor S minimaal voldoet dan aan

$$(X^TX)\ \hat{\theta} = X^TY$$

Als X^TX inverteerbaar is dan geldt dus

$$\hat{\theta} = [X^TX]^{-1}\ X^TY.$$

$\hat{\theta}$ wordt de *kleinste kwadratenschatter* genoemd. De kleinste kwadratenschatter voldoet aan een aantal mooie eigenschappen, die we zullen samenvatten in een stelling.

STELLING 7.1. *Beschouw een stochastisch proces* Y, *met gemiddelde waarde* $E(Y) = X\,\theta$ *en kovariantie* $\sigma^2 I$. *Dan heeft de kleinste kwadratenschatter* $\hat{\theta} = [X^TX]^{-1}\ X^TY$ *de volgende eigenschappen:*

(i) $\hat{\theta}$ *is lineair in de data*

(ii) $E_y[\hat{\theta}] = \theta$, *d.w.z.* $\hat{\theta}$ *is een zuivere schatter*

(iii) $E_y[(\hat{\theta}-\theta)(\hat{\theta}-\theta)^T] \le E_y[(\tilde{\theta}-\theta)(\tilde{\theta}-\theta)^T]$ *voor iedere lineare schatter* $\tilde{\theta}$ *die voldoet aan* $E_y[\tilde{\theta}] = \theta$.

7.2. Stochastische Ingangs-Uitgangssystemen

In het vorige hoofdstuk hebben we stochastische dynamische systemen gedefinieerd. We zullen nu een soortgelijke uitbreiding geven aan de lineaire

Ingangs-Uitgangssystemen. We gaan hierbij uit van een lineaire differentievergelijking (vergelijk met hoofdstuk 4)

$$y_k + A_1 y_{k-1} + \dots + A_n y_{k-n} = B_1 u_{k-1} + \dots + B_n u_{k-n},$$

waarbij

$A_1,\dots,A_n$ zijn (p,p)-matrices,

$B_1,\dots,B_n$ zijn (p,m)-matrices.

(In hoofdstuk 4 lieten we ook nog een term $B_0 u_k$ toe.)

De bijbehorende overdrachtsfunktie is dan

$$\hat{G}(z) = \frac{B_1 z^{-1} + \dots + B_n z^{-n}}{I + A_1 z^{-1} + \dots + A_n z^{-n}} := \frac{B(z)}{A(z)} .$$

We voegen nu een stochastische komponent toe:
Laat $\{\xi_k\}_{k=0}^{\infty}$ een vektorwaardig stationair Gaussisch witte ruis proces op $\mathbb{R}^q$ zijn, met als kovariantiematrix $E[\xi_k \xi_\ell^T] = Q\,\delta_{k\ell}$.
Beschouw de *stochastische differentievergelijking*:

$$y_k + A_1 y_{k-1} + \dots + A_n y_{k-n} = B_1 u_{k-1} + \dots + B_n u_{k-n} + C_0 \xi_k + C_1 \xi_{k-1} + \\ + \dots C_n \xi_{k-n},$$

waarbij $C_0,\dots,C_n$ (p,q)-matrices zijn.
Het zal duidelijk zijn dat $\{y_k\}$ hierdoor - bij gegeven deterministische ingangsrij u_k - een uniek bepaald stochastisch proces wordt, waarvan de eigenschappen analoog als in hoofdstuk 6 bepaald kunnen worden.
De overdrachtsfunktie van een stochastisch systeem kunnen we weer op de bekende manier bepalen; we vinden:

$$\hat{G}(z) = \frac{B_1 z^{-1} + \dots + B_n z^{-n}}{I + A_1 z^{-1} + \dots + A_n z^{-n}} + \frac{C_0 + C_1 z^{-1} + \dots + C_n z^{-n}}{I + A_1 z^{-1} + \dots + A_n z^{-n}} := \frac{B(z)}{A(z)} + \frac{C(z)}{A(z)} .$$

<u>OPMERKING.</u>
We zullen ons verder beperken tot systemen met één ingang en één uitgang. De koëfficienten in de differentievergelijking zullen dan aangeduid worden met $a_1,\dots,a_n$, $b_1,\dots,b_n$, $c_0,\dots,c_n$.

7.3. Identifikatie van stochastische systemen

Voor het identificeren van het stochastisch model, zoals beschreven in de vorige sektie, moeten we nu dus op één of andere manier de systeemparameters $a_1,\dots,a_n$, $b_1,\dots,b_n$ en $c_1,\dots,c_n$ zien te schatten. We zullen dat voor een tweetal zeer eenvoudige gevallen doen. In het algemeen kan men voor een schatter $(\hat{a}_1,\dots,\hat{a}_n,\hat{b}_1,\dots,\hat{b}_n,\hat{c}_1,\dots,\hat{c}_n)$ *niet* bewijzen dat deze *sterk konsistent* is. We zullen op dat probleem hier verder niet ingaan.

a) $C(z) = A(z)$ (Dit komt overeen met meetruis op de waarnemingen).

Definieer: $\theta^T = (a_1,\dots,a_n,b_1,\dots,b_n)$

$$x_k^T = (-y_{k-1},-y_{k-2},\dots,-y_{k-n},u_{k-1},\dots,u_{k-n}),\ k = 1,\dots,N.$$

Dan geldt voor het korresponderende stochastische proces $\{Y_k,\ k = 1,\dots,N\}$ dat door het stochastische systeem gedefinieerd wordt:

$$E[Y_k] = x_k^T\theta.$$

We kunnen nu rechtstreeks de kleinste kwadraten procedure, zoals in 7.1 beschreven, toepassen!

b) $C(z) = 1$.

Het model ziet er dus als volgt uit:

$$y_k = -a_1y_{k-1} + \dots - a_ny_{k-n} + b_1u_{k-1} + \dots + b_nu_{k-n} + \xi_k$$

$$= x_k^T\theta + \xi_k \qquad k = 1,\dots,N$$

waarbij x_k en θ weer gedefinieerd worden door

$$x_k^T := (-y_{k-1},\dots,-y_{k-n},u_{k-1},\dots,u_{k-n})$$

$$\theta^T := (a_1,\dots,a_n,b_1,\dots,b_n).$$

Laat

$$Y^T = [y_1,\dots,y_N]$$
$$X^T = [x_1,\dots,x_N]$$
$$\Sigma^T = [\xi_1,\dots,\xi_N].$$

Dan is

$$Y = X\theta + \Sigma\,.$$

We kunnen nu opnieuw een kleinste kwadraten procedure op deze vergelijking loslaten. We vinden dan de schatter

$$\hat{\theta} = [X^TX]^{-1}\,X^TY.$$

Aangezien we aangenomen hebben dat $\{\xi_k\}$ een witte ruis proces is weten we dat $E\,[\xi_k\xi_\ell] = 0$ voor $k \neq \ell$. Uit deze eigenschap volgt dat de bovengenoemde kleinste kwadratenschatter $\hat{\theta}$ *sterk konsistent* is. Dit wil zeggen dat als we N naar oneindig laten gaan (dus oneindig veel opeenvolgende waarnemingen doen), de schatter $\hat{\theta}$ naar de werkelijke waarde van θ konvergeert.

7.4. Rekursieve kleinste kwadratenmethode

Van de kleinste kwadratenmethode die we in hoofdstuk 7.1 hebben beschreven en die we hebben toegepast in hoofdstuk 7.2 bestaat ook een *rekursieve* versie. Voor praktische identifikatieproblemen is deze methode van groot belang. (Mits natuurlijk de konsistentie van de kleinste kwadratenschatter verzekerd is, zie 7.2) Bij de rekursieve kleinste kwadratenmethode gaat men als volgt te werk. Stel we hebben, zoals in 7.1, de kleinste kwadratenschatter van θ bepaald op basis van de waarnemingen $y_1,\ldots,y_N$:

$$\hat{\theta}_N = [X_N^T X_N]^{-1} X_N^T Y_N, \qquad \text{waarbij}$$

$$X_N = \begin{bmatrix} x_1^T \\ \vdots \\ x_N^T \end{bmatrix}, \qquad Y_N = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}.$$

Veronderstel dat we nu een extra observatie y_{N+1} hebben gedaan. Dan kunnen we uiteraard opnieuw $\hat{\theta}_{N+1}$ gaan bepalen op bovenstaande manier, maar dit geeft voor grote N (en afhankelijk van de grootte van de vektoren x_i) erg veel rekenwerk en dus sneller onnauwkeuriger schattingen. We kunnen $\hat{\theta}_{N+1}$ echter ook bepalen aan de hand van $\hat{\theta}_N$, x_{N+1} en y_{N+1}, en wel als volgt.

Definieer:

$$\Sigma_N = [X_N^T X_N]^{-1}$$

Dan kunnen we $\hat{\theta}$ rekursief bepalen met behulp van de volgende vergelijkingen:

$$\hat{\theta}_{N+1} = \hat{\theta}_N + K_{N+1}(y_{N+1} - x_{N+1}^T \hat{\theta}_N)$$

waarbij

$$K_{N+1} = \Sigma_N x_{N+1} / (1 + x_{N+1}^T \Sigma_N x_{N+1}).$$

Bovendien kunnen we hieruit Σ_N rekursief bepalen:

$$\Sigma_{N+1} = \left[I - \Sigma_N \frac{x_{N+1} x_{N+1}^T}{1 + x_N^T \Sigma_N x_{N+1}} \right] \Sigma_N,$$

mits er een geschikte beginwaarde Σ_0 bekend is.

7.5. Identifikatie door middel van het Kalmanfilter

We kunnen het identifikatieprobleem, zoals geschetst in 7.2, ook als een filterprobleem interpreteren. Het niet-stationaire Kalmanfilter uit het vorige hoofdstuk levert ons dan een rekursieve procedure om de systeemparameters te bepalen.

We gaan weer uit van de stochastische differentievergelijking (één ingang - één uitgang):

$$y_k + a^{(1)} y_{k-1} + \dots + a^{(n)} y_{k-n} = b^{(1)} u_{k-1} + \dots + b^{(n)} u_{k-n} + \xi_k,$$

waarbij $\{\xi_k\}$ een Gaussisch witte ruis proces met als kovariantie $E[\xi_k \xi_\ell] = R_k \delta_{k\ell}$. We veronderstellen nu bovendien dat de $a^{(i)}$ en $b^{(i)}$ aan storingen onderhevig zijn en stellen daarom

$$a_{k+1}^{(i)} = a_k^{(i)} + w_k^{(i)}, \qquad i = 1,\dots,n,$$
$$b_{k+1}^{(i)} = b_k^{(i)} + \bar{w}_k^{(i)}, \qquad i = 1,\dots,n,$$

met $\{w_k^{(i)}\}$ en $\{\bar{w}_k^{(i)}\}$ onderling onafhankelijke Gaussische witte ruis processen, die bovendien onafhankelijk zijn van ξ_k.

Definieer:

$$\theta_k^T = (a_k^{(1)}, a_k^{(2)}, \dots, a_k^{(n)}, b_k^{(1)}, \dots, b_k^{(n)})$$
$$\omega_k = (w_k^{(1)}, \dots, w_k^{(n)}, \bar{w}_k^{(1)}, \dots, \bar{w}_k^{(n)})$$
$$C_k = (-y_{k-1}, \dots, -y_{k-n}, u_{k-1}, \dots, u_{k-n})$$

en laat

$$E[\omega_k \omega_\ell^T] = Q_k \delta_{k\ell}.$$

Dan is:

$$\theta_{k+1} = \theta_k + \omega_k$$
$$y_k = C_k \theta_k + \xi_k.$$

We hebben hier een stochastisch dynamisch systeem gekregen waarvoor we het filterprobleem weer kunnen oplossen, namelijk we kunnen $\hat{\theta}_{k+1|k}$ weer met een speciale versie van het Kalmanfilter bepalen. (C_k hangt van $y_{k-1}, \dots$ $\dots, y_{k-n}$ af! Merk overigens op dat θ hier de rol van toestand vervult.)
Het filter wordt:

$$\hat{\theta}_{k+1|k} = [I - K_k C_k] \cdot \hat{\theta}_{k|k-1} + K_k y_k$$

met

$$K_k = \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T + R_k]^{-1}$$

en

$$\Sigma_{k+1|k} = \Sigma_{k|k-1} - \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T + R_k]^{-1} C_k \Sigma_{k|k-1} + Q_k.$$

Als beginwaarde voor $\hat{\theta}_{k+1|k}$, $\hat{\theta}_{0|-1}$, kiezen we een a priori schatting van de systeemparameters en als "begin"-foutkovariantiematrix kiezen we de a priori kovariantiematrix van de systeemparameters. Eigenschappen over $\hat{\theta}_{k+1|k}$ kunnen we met behulp van hoofdstuk 6 afleiden.

<u>VOORBEELD 7.1.</u>
Veronderstellen we bovendien in bovenstaande modellering dat $Q_k = 0$ en $R_k = 1$ dan krijgen we

$$\hat{\theta}_{k+1|k} = [I - K_k C_k]\, \hat{\theta}_{k|k-1} + K_k y_k$$
$$K_k = \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T + 1]^{-1}$$

$$\Sigma_{k+1|k} = \left[I - \Sigma_{k|k-1} \frac{C_{k+1} C_{k+1}^T}{1 + C_{k+1} \Sigma_{k|k-1} C_{k+1}^T} \right] \Sigma_{k|k-1}$$

hetgeen precies overeen komt met de rekursieve kleinste kwadratenmethode! Met andere woorden, de rekursieve kleinste kwadratenmethode is een speciaal Kalmanfilter. De eigenschappen voor deze schatter volgen nu dus uit hoofdstuk 6. □

<u>OPGAVEN.</u>

1. Gebruik de kleinste kwadratenmethode om een kwadratische funktie van de vorm $E[y_t] = at^2 + bt + c$ te vinden, waarbij de volgende gegevens bekend zijn

t :	1	2	3	4	5	6	7	8	9	10
y_t :	9,6	1,3	0	0,4	4,1	3,8	9	1,8	0,7	0,1

2. Beschouw het systeem $y_t = ay_{t-1} + \xi$ Hierbij is $\{\xi_t\}$ een ongekorreleerd stochastisch proces dat de waarden +1 en −1 kan aannemen.

$$P(\xi_k = 1) = P(\xi_k = -1) = \tfrac{1}{2}$$

Bewijs dat de kleinste kwadratenschatter voor a een zuivere schatter is als $N = 2$ en y_0 bekend is.

HOOFDSTUK 8

TERUGKOPPELING EN OPTIMALE BESTURING

Tot nog toe hebben we ons alleen maar bezig gehouden met het analyseren van een gegeven systeem. (Ingangs-Uitgangssysteem of systeem met toestandsruimte.) Een tweede, even belangrijk facet van de systeemtheorie is het kunnen "regelen" van een systeem. We zullen eerst een paar regelwetten introduceren en vervolgens een tweetal regelproblemen behandelen.

We gaan uit van een lineair stationair systeem in diskrete tijd

$$\begin{cases} x_{k+1} = Ax_k + Bu_k \\ y_k = Cx_k \end{cases},$$

waarbij

A een (n,n)-matrix,
B een (n,m)-matrix,
C een (p,n)-matrix.

Bij een regelwet wordt de ingangsfunctie bepaald op basis van de toestand of de uitgang. Bij de eerste mogelijkheid - de zogeheten *toestandsterugkoppeling* - veronderstelt men het volgende verband: $u = Fx$, waarbij F een (m,n)-matrix is. Hierdoor verkrijgen we het (autonome) systeem:

$$\begin{cases} x_{k+1} = (A + BF)x_k \\ y_k = Cx_k. \end{cases}$$

(Merk op dat de toestandsterugkoppeling *geen* tijdsfunktie is.)

We kunnen de input ook bepalen op grond van de uitgang - *uitgangsterugkoppeling* - door aan te nemen dat $u = Ky$ voor een (n,p)-matrix K. Het resultaat is dan

$$\begin{cases} x_{k+1} = (A + BKC)x_k \\ y_k = Cx_k \end{cases}$$

(Ook hier is de uitgangsterugkoppeling onafhankelijk van de tijd.)

OPMERKINGEN.

(i) Men kan uiteraard ook ingewikkelder konstrukties voor een terugkoppeling bedenken. Zo is het bijvoorbeeld mogelijk om de uitgang van ons oorspronkelijke systeem op te vatten als de ingang van een nieuw systeem en dan de nieuwe uitgang "terug te koppelen" naar de oorspronkelijke ingang.

(ii) Voor systemen in kontinue tijd kunnen we op analoge manier toestands- en uitgangsterugkoppeling definiëren. We zullen daar verder niet op in gaan.

8.1. Stabiliteit door toestandsterugkoppeling

Veronderstel dat we een autonoom systeem $x_{k+1} = Ax_k$, $x_0 = \bar{x}_0$ gegeven hebben. Dus $x_k = A^k \bar{x}_0$. Het is eenvoudig in te zien (zie voorbeeld 2.1) dat, als de eigenwaarden van de matrix A alle in absolute waarde kleiner dan 1 zijn, dat x_k naar de oorsprong konvergeert. (onafhankelijk van $\bar{x}_0$) Het autonome systeem wordt dan *stabiel* genoemd en er geldt:

PROPOSITIE 8.1. *Het autonome systeem* $x_{k+1} = Ax_k$ *is stabiel als en slechts als alle eigenwaarden van de matrix* A *in absolute waarde kleiner zijn dan* 1.

Veronderstel nu dat we een dynamisch systeem hebben van de vorm

$$\begin{cases} x_{k+1} = Ax_k + Bu_k \\ y_k = x_k \end{cases}$$

Aangezien stabiliteit vaak een gewenste eigenschap van een systeem is, kunnen we ons afvragen of we dit systeem ook kunnen stabiliseren. Beter gezegd: bestaat er een toestandsterugkoppeling u = Fx zodanig dat de uitgang (= toestand) naar nul konvergeert? De volgende stelling geeft een gedeeltelijk antwoord op deze vraag.

STELLING 8.2. *Laat gegeven zijn*

$$\begin{cases} x_{k+1} = Ax_k + Bu_k \\ y_k = x_k \end{cases}$$

dan geldt: als (A,B) *regelbaar is dan bestaat er een toestandsterugkoppeling* u = Fx *zodanig dat de uitgang* y_k *naar nul konvergeert* (*voor* $k \to \infty$).

OPMERKINGEN.

(i) Merk op dat in dit geval de toestandsterugkoppeling ook een uitgangs-terugkoppeling is.

(ii) Voor tijdskontinue systemen geldt een analoge stelling. In dat geval is het autonome systeem $\dot{x} = Ax$ stabiel als van elke eigenwaarde van de matrix A het reële deel kleiner is dan nul (vergelijk met opgave 3 van hoofdstuk 2).

VOORBEELD 8.3.

Laat gegeven zijn het systeem

$$\begin{cases} x_{k+1} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} x_k + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u_k \\ y_k = x_k \end{cases}$$

Dit systeem is regelbaar want

$$\dim\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \vdots \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right] = \dim\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \vdots \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right] = 2.$$

We gaan nu een toestandsterugkoppeling $u = Fx$ konstrueren die dit systeem stabiel maakt. We schrijven $F = (f_1 \ f_2)$, dan is

$$A + BF = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}(f_1 \ f_2) = \begin{pmatrix} 2 & 1 \\ f_1 & 1+f_2 \end{pmatrix}.$$

We dienen er voor te zorgen dat de beide eigenwaarden van deze matrix, δ_1 en δ_2, in absolute waarde kleiner dan 1 zijn. Kies daartoe $f_1 = \frac{7}{4}$ en $f_2 = -2$. Dan is $\lambda_1 = \lambda_2 = \frac{1}{2} < 1$, en dus konvergeert elke oplossing van

$$x_{k+1} = \begin{pmatrix} 2 & 1 \\ \frac{7}{4} & -1 \end{pmatrix} x_u$$

naar nul. □

8.2. Optimale besturing door toestandsterugkoppeling

We gaan uit van het lineaire systeem Σ

$$\begin{cases} x_{k+1} = Ax_k + Bu_k \\ y_k = Cx_k, \end{cases}$$

waarbij A weer een (n,n)-matrix,

B een (n,m)-matrix,

C een (p,n)-matrix.

We veronderstellen dat er bovendien een *kostenfunctie* J gegeven is

$$J = \sum_{k=0}^{\infty} (u_k^T R u_k + x_k^T C^T C x_k),$$

met R een symmetrische positief definiete (m,m)-matrix.

PROBLEEM. Bepaal, indien mogelijk, een besturing u die, voor een gegeven beginvoorwaarde x_0, J minimaliseert.

Onder de extra aanname dat $\sum$ minimaal is zullen we de oplossing van dit probleem geven. De optimale besturing u blijkt dan weer gevonden te kunnen worden door toestandsterugkoppeling.

STELLING 8.4. *De optimale besturingsfunktie voor* $\sum$ *die* J *voor een willekeurige beginvoorwaarde* x_0 *minimaliseert wordt gegeven door* $u^* = -Fx$, *waarbij*

$$F = \{R + B^T[C^TC + K]B\}^{-1} B^T[C^TC + K]A$$

en de symmetrische (n,n)*-matrix* K *is de positief definiete oplossing van de algebraïsche Riccati-vergelijking:*

$$K = A^T[C^TC + K][A - B(\{R + B^T[C^TC + K]B\}^{-1}B^T[C^TC + K]A)].$$

De minimale kosten zijn dan gelijk aan

$$J^* = x_0^T K x_0.$$

In plaats van te optimaliseren over een oneindig aantal tijdstippen kunnen we ook een eindig optimaliseringsprobleem formuleren. We nemen dan als kostenfunktie

$$\bar{J} = \sum_{k=0}^{k=k_0-1} (u_k^T R u_k + x_k^T C^T C x_k) + x_{k_0}^T M x_{k_0}.$$

Ook in dit geval kunnen we de optimale besturingsfunktie weer vinden als een niet-stationaire toestandsterugkoppeling.

STELLING 8.5. *De optimale besturingsfunktie voor het systeem* $\sum$ *die* $\bar{J}$ *voor een willekeurige beginvoorwaarde* x_0 *minimaliseert, wordt gegeven door* $u_k = -F_k x_k$, $k = 0,\dots,k_0-1$, *waarbij*

$$F_k = \{R + B^T[C^TC + K_{k+1}]B\}^{-1} B^T[C^TC + K_{k+1}] A.$$

De symmetrische (n,n)*-matrices* K_k *voldoen aan de vergelijking* $(k = 0,\ldots,k_0)$:

$$\begin{cases} K_k = A^T[C^TC + K_{k+1}][A - B(\{R + B^T[C^TC + K_{k+1}]B\}^{-1}B^T[C^TC + K_{k+1}]A)] \\ K_{k_0} = M \end{cases}$$

De minimale kosten worden dan gegeven door $\bar{J}^* = x_0^T K_0 x_0$.

OPMERKINGEN.

(i) Vergelijk de oplossing van stelling 8.5 en stelling 8.4 met de oplossing van het filterprobleem uit hoofdstuk 6. Het optimaliseren over oneindig veel tijdstippen geeft een gelijksoortige algebraïsche Riccati vergelijking als het stationaire Kalmanfilter. Daarentegen verschillen stelling 8.5 en stelling 6.2 (het niet-stationaire Kalmanfilter) wel van elkaar, namelijk voor het Kalmanfilter is de beginvoorwaarde van de Riccati-vergelijking gegeven terwijl in stelling 8.5 de eindvoorwaarde gegeven is.

(ii) Analoog aan wat er voor het stationaire Kalmanfilter geldt, geldt ook hier weer dat de oplossing K van de algebraïsche Riccati-vergelijking voldoet aan $\lim_{k\to\infty} K_k = K$, waarbij K_k de oplossing is van de Riccati-vergelijking uit stelling 8.5.

(iii) Het systeem $x_{k+1} = (A - BF)x_k$ is stabiel; dus onafhankelijk van de beginvoorwaarde konvergeert iedere oplossing naar nul. (en dus konvergeert ook de uitgangsvariabele $y_k = Cx_k$ naar nul.)

(iv) Voor tijdskontinue systemen kan een analoge optimale besturingstheorie worden ontwikkeld.

VOORBEELD 8.6.

Beschouw het 1-dimensionale systeem

$$x_{k+1} = x_k + u_k$$
$$y_k = x_k$$

met als kostenfunktie

$$J = \sum_{k=0}^{\infty} (u_k^2 + x_k^2).$$

Voor de algebraïsche Riccati-vergelijking vinden we dan

$$k = 1(1+k)\left(1 - \frac{1+k}{2+k}\right)$$

dus

$$k^2 + k - 1 = 0$$

met als positieve oplossing $-\frac{1}{2}+\frac{1}{2}\sqrt{5}$.

De toestandsterugkoppeling wordt dan gegeven door

$$u = -Fx = -\left(\frac{\frac{1}{2}+\frac{1}{2}\sqrt{5}}{\frac{3}{2}+\frac{1}{2}\sqrt{5}}\right)x = (\tfrac{1}{2} - \tfrac{1}{2}\sqrt{5})\ x\,.$$

De minimale kosten zijn gelijk aan $(-\frac{1}{2}+\frac{1}{2}\sqrt{5})x_0^2$.

Het systeem $x_{k+1} = (A + BF)x_k = (\frac{3}{2} - \frac{1}{2}\sqrt{5})x_k$ is stabiel. □

We willen dit hoofdstuk besluiten met het geven van een stochastische versie van het optimaliseringsprobleem.

We gaan uit van het systeem $\sum$

$$\begin{cases} x_{k+1} = Ax_k + Bu_k + v_k \\ y_k = Cx_k, \end{cases}$$

hierbij zijn A, B en C als voorheen en v_k is een Gaussisch witte ruis proces met als kovariantie $E[v_k v_\ell^T] = Q_k\ \delta_{k\ell}$. De kostenfunktie wordt gegeven door

$$\bar{J} = E[\sum_{k=0}^{k_0-1} (u_k^T\ R\ u_k + x_k^T\ C^TC\ x_k) + x_{k_0}\ M\ x_{k_0}].$$

Dan geldt:

<u>STELLING 8.7.</u> *De optimale besturingsfunktie voor het systeem* $\sum$, *dat minimaal wordt verondersteld, die de kostenfunktie* $\bar{J}$ *voor elke willekeurige beginvoorwaarde* x_0 *minimaliseert wordt gegeven door* $u_k = -F_k x_k$, $k = 0,\dots, k_0-1$, *waarbij*

$$F_k = \{R + B^T[C^TC + K_{k+1}]\ B\}^{-1}B^T[C^TC + K_{k+1}]\ A\,.$$

De matrices K_k *worden bepaald door:*

$$\begin{cases} K_k = A^T[C^TC + K_{k+1}][A - BF_k] \\ K_{k_0} = M. \end{cases}$$

De minimale kosten zijn dan gegeven door:

$$\bar{J}^* = x_0^T\ K_0\ x_0 + \sum_{k=1}^{k_0} \mathrm{spoor}\{Q_{k-1}[K_k + C^TC]\}.$$

Ofwel

$$\begin{cases} K_k = A^T[C^TC + K_{k+1}][A - B(\{R + B^T(C^TC + K_{k+1}] B\}^{-1}B^T[C^TC + K_{k+1}]A)] \\ K_{k_0} = M\ . \end{cases}$$

De minimale kosten worden dan gegeven door

$$\bar{J}^* = x_0^T K_0 x_0 + \sum_{k=1}^{k_0} \text{spoor}\{Q_{k-1}[K_k + C^TC]\}.$$

OPMERKINGEN.

(i) De optimale besturingsfunktie wordt niet door de ruisterm beïnvloed, de minimale kosten daarentegen wel. (vergelijk met stelling 8.5)

(ii) Zonder problemen laat deze stelling zich uitbreiden tot het analogon van stelling 8.4.

OPGAVEN.

1. Is het voorbeeld 1.1 stabiliseerbaar door middel van toestandsterugkoppeling? Zo ja, hoe luidt een mogelijke terugkoppelingswet?
2. We kunnen ook een alternatieve versie van stelling 8.2 geven. Veronderstel dat het systeem $x_{k+1} = Ax_k + Bu_k$ regelbaar is. Dan bestaat er een toestandsterugkoppeling $u = Fx$, zodanig dat voor het gevormde autonome systeem $x_{k+1} = (A + BF)x_k$ elke beginwaarde x_0 in ten hoogste n (= dimensie toestandsruimte) stappen naar nul gebracht wordt. Probeer voor voorbeeld 8.3 zo'n terugkoppeling te vinden.
3. Werk opmerking (ii) na stelling 8.7 uit.
4. Beschouw het systeem

$$\begin{cases} x_{k+1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x_k + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u_k, \quad x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ y_k = (1 \quad 0)x_k \end{cases}$$

met als kostenfunktie $J = \sum_{k=0}^{\infty} \{u_k^2 + x_k^T \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} x_k\}$.

Bereken J^*.

APPENDIX A

LINEAIRE DIFFERENTIAALVERGELIJKINGEN

Laat M een reële (n,n)-matrix zijn. Analoog aan het geval dat M een (1,1)-matrix is (d.w.z. een getal) definiëren we

(1) $$e^M := \sum_{k=0}^{\infty} \frac{M^k}{k!}$$

en we zien direkt dat deze oneindige som absoluut konvergeert:

$$\left\| \sum_{k=0}^{\infty} \frac{M^k}{k!} \right\| \le \sum_{k=0}^{\infty} \frac{\|M\|^k}{k!} = e^{\|M\|}.$$

Beschouwen we nu de differentiaalvergelijking

(2) $\dot{x}(t) = A\, x(t), \quad x(0) = x_0, \quad A : (n,n)\text{-matrix},$

dan gaan we met behulp van (1) de oplossing van (2) afleiden.
Uit (1) volgt

$$e^{At} = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!}$$

en

$$\frac{d}{dt}\left(e^{At}\right) = \sum_{k=0}^{\infty} \frac{k\, t^{k-1} A^k}{k!} = A\, e^{At}.$$

Bovendien geldt $e^{A.0} = I_{n,n}$ (de (n,n)-eenheidsmatrix) zodat we nu als oplossing van (2) krijgen

$$x(t) = e^{At} x_0.$$

Om de algemene oplossing van (2) te bepalen kunnen we ook nog iets anders te werk gaan. Beschouw de bij (2) behorende matrix differentiaalvergelijking

(3) $$\begin{cases} \dot{\Phi}(t,\tau) = A\, \Phi(t,\tau), \text{ hierbij is } \dot{\phi}(t,\tau) = \frac{\partial\phi}{\partial t}(t,\tau) \\ \Phi(\tau,\tau) = I_{n,n} \end{cases}$$

dan geldt:

(4) $$\Phi(t,\tau) = e^{A(t-\tau)}.$$

Dit volgt uit de limiet voor $k \to \infty$ van de rekursie

$$M_0(t) = I$$
$$M_{k+1}(t) = I + \int_\tau^t A\, M_k(\sigma)d\sigma;$$
$$\Phi(t,\tau) = \lim_{k\to\infty} M_k(t);$$

de gewone manier om oplossingen van een differentiaalvergelijk te bepalen!

$\Phi(t,\tau)$ wordt wel de *transitiematrix* van (2) genoemd. De oplossing van de vergelijking (2) vinden we m.b.v. (4) door te stellen $\tau = 0$ (beginwaarde was gegeven voor $t = 0$!) en

$$x(t) = \Phi(t,0)x_0.$$

Een bekende veel gebruikte methode om e^{At} te bepalen is de volgende. (Er bestaan meerdere methoden waarvan uiteraard ook (1) numeriek toepassing vindt.) Zoals we bekend zullen veronderstellen kunnen we A in *Jordan kanonieke vorm* brengen, d.w.z.

$$J = SAS^{-1} \qquad (A = S^{-1}JS)$$

waarbij J een *blokdiagonaal* matrix is.

$$J = \begin{bmatrix} J_1 & & & & \bigcirc \\ & J_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ \bigcirc & & & & J_k \end{bmatrix} \qquad \begin{array}{l} J_i : n_i \times n_i\text{-matrix} \\ \sum_{i=1}^{k} n_i = n \end{array}$$

Nu zijn een aantal J_i's $- J_1,\dots,J_\ell -$ van de vorm

$$J_i = \lambda_i I + N_i; \qquad \begin{array}{l} I : n_i \times n_i\text{-eenheidsmatrix} \\ N_{ij} : n_{ij} \times n_{ij} \text{ matrix} \\ \sum_{j=1}^{k_i} n_{ij} = n_i \end{array}$$

$$N_i = \begin{bmatrix} N_{i1} & & & & \bigcirc \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ \bigcirc & & & & N_{ik_i} \end{bmatrix}$$

en

$$N_{ij} = \begin{bmatrix} 0 & 1 & & \bigcirc \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \bigcirc & & & 0 \end{bmatrix}$$

terwijl de overige J_i de vorm hebben $(i = \ell+1,\dots,k)$

$$J_i = \begin{bmatrix} M_{i1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & M_{ik_i} \end{bmatrix} \qquad \begin{gathered} M_{ij} = 2n_{ij} \times 2n_{ij} \\ \sum_{i=1}^{k_i} 2n_{ij} = n_i \end{gathered}$$

$$M_{ij} = \begin{bmatrix} \sigma_i & \omega_i & 1 & 0 & \bigcirc & \bigcirc \\ -\omega_i & \sigma_i & 0 & 1 & & \\ \bigcirc & & \sigma_i & \omega_i & 1 \quad 0 & \bigcirc \\ & & -\omega_i & \sigma_i & 0 \quad 1 & \\ \bigcirc & & \bigcirc & & \vdots \quad \vdots & \vdots \quad \vdots \end{bmatrix}$$

Nu is het bepalen van e^{Jt} gemakkelijk en

$$e^{At} = S^{-1} e^{Jt} S,$$

namelijk

$$e^{\lambda_i t} = \begin{bmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \vdots \\ & \ddots & \ddots & \ddots & \frac{t^2}{2!} \\ & & \ddots & \ddots & t \\ \bigcirc & & & & 1 \end{bmatrix} \qquad i = 1,\dots,\ell$$

$$e^{M_{ij}t} = e^{\sigma_i t} \begin{bmatrix} \Omega_i & t\Omega_i & \cdots & \frac{t^{n_{ij}-1}\,\Omega_i}{(n_{ij})!} \\ 0 & \Omega_i & \cdots & \vdots \\ 0 & 0 & \ddots & \vdots \\ \cdots & \cdots & \ddots & t\Omega_i \\ 0 & \cdots & 0 & \Omega_i \end{bmatrix}$$

met

$$\Omega = \begin{bmatrix} \cos \omega_i t & \sin \omega_i t \\ -\sin \omega_i t & \cos \omega_i t \end{bmatrix}.$$

Door kombinatie van deze oplossingen krijgen we e^{Jt} en vervolgens $S^{-1}\ e^{Jt}\ S$.

APPENDIX B

LAPLACE TRANSFORMATIE EN Z-TRANSFORMATIE

(i) De z-transformatie

Definieer:

$$\ell_+ = \{\{f_k\}_0^\infty,\ f_k \in \mathbb{C} \mid \sum_0^\infty |f_k| < \infty\}.$$

Stel nu dat er een rij $\{f_k\}_0^\infty$ gegeven is en laat

$$G_a(f) = \{z \in \mathbb{C} \mid \{f_k\, z^{-k}\} \in \ell_+\}.$$

Voor $z \in G_a$ definiëren we dan

$$\hat{f}(z) = \sum_{k=0}^\infty f_k\, z^{-k}.$$

$\hat{f}(z)$ heet de *z-getransformeerde* van $\{f_k\}_0^\infty$.

OPMERKING.

Als een rij $\{f_k\}_0^\infty$ slechts *geometrische groei* heeft - d.w.z. $\exists r \in \mathbb{R}^+$ zodat $\sum_0^\infty |f_k| r^{-k} < \infty$ - is het konvergentiegebied $G_a(f)$ van de vorm $|z| > r_a$ (of $|z| \geq r_a$) voor een $r_a \in \mathbb{R}^+$.

We zullen in het kort een aantal eigenschappen voor de z-getransformeerde formuleren. We zullen ons daarbij beperken tot het geval dat de rijen, die we hier bekijken, geometrische groei hebben.

1. $f_0 = \lim\limits_{z\to\infty} \hat{f}(z)$.

   Als $\lim\limits_{k\to\infty} f_k$ bestaat, dan is $|z| > 1$ en

   $$\lim_{|z|\downarrow 1} (z-1)\, \hat{f}(z) = f_\infty.$$

2. Laat gegeven een rij $\{f_k\}_0^\infty$ en definieer de rij $\{g_k\}_0^\infty$ door

   $g_0 = 0$
   $g_1 = f_0$
   $g_2 = f_1 \quad \ldots$

   dan is

   $$\hat{g}(z) = \frac{1}{z} \cdot \hat{f}(z).$$

3. Stel gegeven twee rijen $\{f_k\}_0^\infty$ en $\{g_k\}_0^\infty$ en een rij $\{h_k\}_0^\infty$ gedefinieerd door

$$h_k = \sum_{\ell=0}^{k} f_{k-\ell} \cdot g_\ell .$$

Dan is

$$\hat{h}(x) = \hat{f}(z) \; . \; \hat{g}(z).$$

4. $$f_k = \frac{1}{2\,i} \oint \hat{f}(z)\; z^{k-1}\; dz, \quad k \in \mathbb{N},$$

waarbij $\oint$ wil zeggen dat de integratie over een cirkel - in het konvergentiegebied G_a - moet worden uitgevoerd. Deze formule geeft dus een procedé om uit de z-getransformeerde van een rij - dat is dus de funktie $\hat{f}(z)$ - de oorspronkelijke rij terug te vinden.

(ii) Laplace getransformeerde

Definieer:

$$L_+ := \{f : \mathbb{R}^+ \to \mathbb{C} \mid \int_0^\infty |f(t)|dt < \infty\},$$

$$\mathbb{R}^+ = [0,\infty),$$

en evenzo

$$L_+^{loc} = \{f : \mathbb{R}^+ \to \mathbb{C} \mid \int_0^T |f(t)|dt < \infty \mid \forall 0 \le T < \infty\}.$$

Voor een funktie $f \in L_+^{loc}$ definiëren we:

$$G_a(f) := \{s \in \mathbb{C} \mid f(t)\, e^{-st} \in L_+\}.$$

We zullen ons hier beperken tot het geval dat de hier beschouwde funkties slechts *exponentiële groei* hebben, namelijk $\exists \sigma_a \in \mathbb{R}$ zodat

$$G_a = \{s \in \mathbb{C} \mid \text{Re } s \ge \sigma_a\}$$

of

$$G_a = \{s \in \mathbb{C} \mid \text{Re } s > \sigma_a\}.$$

Voor $s \in G_a$ definiëren we

$$\hat{f}(s) = \int_0^\infty f(t)\, e^{-st}\, dt.$$

$\hat{f}(s)$ heet de *Laplace getransformeerde* van f.

We zullen weer enige eigenschappen over Laplace getransformeerden geven:

1. Laat $f \in L_+^{loc}$ en $g \in L_+^{loc}$ met $g(t) = f(t-a)$, $a \geq 0$.
   Dan is $\hat{g}(s) = e^{-as}\hat{f}(s)$.

2. Laat $f \in L_+^{loc}$ en $g \in L_+^{loc}$ met $g(t) = e^{-\alpha t}f(t)$

   $\Rightarrow \hat{g}(s) = \hat{f}(s+\alpha)$.

3. $\hat{\bar{f}}(s) = \bar{\hat{f}}(\bar{s})$; $^-$ geeft de komplex toegevoegde aan.

4. $\hat{f}(s)$ is een analytische funktie.

5. Als f differentieerbaar op $[0,\infty)$ dan geldt

   $$\hat{\dot{f}}(s) = s\,\hat{f}(s) - f(0).$$

6. $\lim\limits_{t\to 0} f(t) = \lim\limits_{s\to\infty} s\,\hat{f}(s)$,

   $\lim\limits_{t\to\infty} f(t) = \lim\limits_{s\to 0} s\,\hat{f}(s)$ mits $\lim\limits_{t\to\infty} f(t)$ bestaat.

7. Als $f,g \in L_+^{loc}$ dan wordt de *konvolutie* van f en g gedefinieerd door

   $$h(t) = \int_0^t f(t-\tau)g(\tau)d\tau, \quad t \geq 0.$$

   Notatie: $h = f * g$;

   er geldt:

   $$\hat{h}(s) = \hat{f}(s).\hat{g}(s).$$

8. We merken tenslotte nog op dat uit de Laplace getransformeerde $\hat{f}(s)$ we "zo ongeveer" de oorspronkelijke afbeelding $f(t)$ weer terug kunnen halen.

APPENDIX C

WAARSCHIJNLIJKHEIDSREKENING EN STOCHASTISCHE PROCESSEN

1. Waarschijnlijkheidsrekening

Laat Ω een verzameling zijn en A een familie van deelverzamelingen van Ω waarvoor geldt:

1) $A_1, A_2 \in A \Rightarrow A_1 \cup A_2 \in A$
2) $A \in A \Rightarrow \bar{A} \in A$ ($\bar{A}$ is het complement van A)
3) $\Omega \in A$
4) $A_i \in A$ $(i \in \mathbb{N}) \Rightarrow \bigcup_{i \in \mathbb{N}} A_i \in A$.

De familie A heet in zo'n geval een *σ-algebra* en (Ω, A) wordt een *meetbare ruimte* genoemd. Ω heet de *basisruimte*, terwijl de elementen van A *gebeurtenissen* genoemd worden. Een *waarschijnlijkheidsruimte* (Ω, A, P) is een meetbare ruimte (Ω, A) voorzien van een afbeelding, de *waarschijnlijkheidsmaat* $P : A \to [0,1]$ die voldoet aan

$$P(A) \geq 0, \quad \forall A \in A$$

$$P(\Omega) = 1$$

$$P\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i \in \mathbb{N}} P(A_i)$$

voor elke aftelbare verzameling van gebeurtenissen $A_i (i \in \mathbb{N})$, die paarsgewijs disjunkt zijn ($A_i \cap A_j = \emptyset$, $i \neq j$).

Laat A en B twee gebeurtenissen zijn en veronderstel dat $P(B) > 0$. De *voorwaardelijke waarschijnlijkheid* $P(A|B)$ van A, gegeven B, is dan

$$P(A|B) = \frac{P(A \cap B)}{P(B)} .$$

Twee gebeurtenissen A en B heten *onafhankelijk* als geldt dan $P(A \cap B) = P(A) \,.\, P(B)$ (ofwel $P(A|B) = P(A)$ voor $P(B) > 0$). Op een analoge manier zijn gebeurtenissen $A_1, \ldots A_n$ *onderling onafhankelijk* als geldt

$$P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}).P(A_{i_2}). \; \ldots \; P(A_{i_k}),$$

voor alle rijtjes $1 \leq i_1 < i_2 < \ldots < i_k \leq n$.

Een *stochastische variabele* op een waarschijnlijkheidsruimte (Ω, Q, P) is een afbeelding $X : \Omega \to \mathbb{R}$ die voldoet aan de volgende eigenschap $\forall a \in \mathbb{R}$ geldt $\{\omega \mid X(\omega) \leq a\} \in A$, dus $P(X \leq a) := \{\omega \mid X(\omega) \leq a\}$ is goed gedefinieerd.

Met een gegeven stochastische variabele X kunnen we een *distributiefunktie* $F_X : \mathbb{R} \to [0,1]$ associëren. Deze wordt gegeven door

$$F_X(x) = P(X \le x) \qquad (= P(\{\omega \in \Omega \mid X(\omega) \le x\}).$$

In het geval dat F_X een differentieerbare funktie is, wordt de *dichtheidsfunktie* $P_X : \mathbb{R} \to \mathbb{R}$ gedefinieerd als

$$P_X(x) = \frac{dF_X(x)}{dx} .$$

Voor stochastische variabelen X en Y wordt de *gemeenschappelijke* distributiefunktie $F_{X,Y} : \mathbb{R}^2 \to [0,1]$ gedefinieerd door

$$F_{X,Y}(x,y) = P\{(X \le x) \cap (Y \le y)\}$$

en evenzo is de gemeenschappelijke dichtheidsfunktie

$$P_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$$

(mits deze afgeleide bestaat).
Voor stochastische variabelen X en Y wordt de *voorwaardelijke* dichtheidsfunktie $P_{X|Y} : \mathbb{R}^2 \to \mathbb{R}$ gedefinieerd door

$$P_{X|Y}(x|y) = \frac{P_{X,Y}(x,y)}{P_Y(y)}$$

als $P_Y(y) > 0$.
Een *stochastische vektor* is een vektor $X = (X_1,\dots,X_n)$, X_i stochastische variabele, met als distributiefunktie

$$F_X(x) = F_X(x_1,\dots,x_n) = P\{(X_1 \le x_1) \cap (X_2 \le x_2) \cap \dots \cap (X_n \le x_n)\}$$

en als dichtheidsfunktie

$$P_X(x) = P_X(x_1,\dots,x_n) = \frac{\partial^n}{\partial x_1 \dots \partial x_n} F_X(x_1,\dots,x_n).$$

Twee stochastische variabelen (vektoren) X en Y heten *onafhankelijk* als geldt dat $F_{X,Y}(x,y) = F_X(x) \, . \, F_Y(y)$.

We zullen verder alleen maar het begrip stochastische variabele gebruiken, hoewel we daarmee ook stochastische vektoren bedoelen.

De *verwachtingswaarde* van een stochastische variabelen X wordt gedefinieerd als $E[X] = \int_{-\infty}^{\infty} x P_X(x) dx$.

Laat X een stochastische variabele zijn en laat $g : \mathbb{R} \to \mathbb{R}$ een kontinu differentieerbare funktie zijn. Dan is $Y = g(X)$ weer een stochastische variabele met als distributiefunktie $F_Y(y) = P(Y \le y) = P(\{x | g(x) \le y\})$. De *kovari-*

antie van een stochastische variabele X is gegeven door $E\{(X-E[X])(X-E[X])^T\}$.

De voorwaardelijke verwachtingswaarde van een stochastische variabele X, gegeven een gebeurtenis A wordt gedefinieerd als $E[X|A] = \int_{-\infty}^{+\infty} xP_X(x|A)dx$. Evenzo worde de *voorwaardelijke verwachtingswaarde van* X, *gegeven* Y, voor twee stochastische variabelen X en Y gedefinieerd door

$$E[X|Y = y] = \int_{-\infty}^{\infty} x\ P_{X|Y}(x|y)dx.$$

2. Stochastische processen

Laat gegeven zijn een waarschijnlijkheidsruimte (Ω,A,P). Een *stochastisch proces* $\{\xi_k\}$ is een afbeelding $\xi : \Omega \times N \to \mathbb{R}$ waarvoor $\forall k \in \mathbb{N}, \xi_k := \xi(\cdot,k)$ een stochastische variabele is.

Een stochastische variabele X is *Gaussisch* als geldt

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right].$$

In dat geval is

$$E[X] = \mu, \quad E[(X - E[X])^2] = \sigma^2.$$

Een stochastische vektor X is *Gaussisch* als geldt

$$P_X(x) = \frac{1}{(2\pi)^{n/2}} \frac{1}{|\Sigma|^{1/2}} \exp[-\tfrac{1}{2}(x-m)^T\ \Sigma(x-m)].$$

Dan geldt $E[X] = m$

$$E[(X-m)(x-m)^T] = \Sigma.$$

(Opmerking: Hierbij is noodzakelijk dat de kovariantiematrix Σ niet singulier is.)

De *gemiddelde waardefunktie (kortweg het gemiddelde)* m_k van een stochastisch proces $\{\xi_k\}$ wordt - analoog aan de definitie voor een stochastische variabele - gedefinieerd als de tijdsfunktie

$$m_k = E[\xi_k], \quad k \in \mathbb{N}.$$

De *kovariantie* is de verzameling

$$E[(\xi_k - m_k)(\xi_\ell - m_\ell)^T], \quad k,\ell \in \mathbb{N}.$$

Twee stochastische processen $\{\xi_k\}$ en $\{\eta_k\}$ heten *ongekorreleerd* als geldt

$$E[\xi_k\eta_\ell^T] = E[\xi_k] \cdot E[\eta_\ell^T],\ \forall k,\ell \in \mathbb{N}.$$

$\{\xi_k\}$ en $\{\eta_k\}$ heten *onafhankelijk* als voor alle $\{k_1,\dots,k_n\}$ en $\{\ell_1,\dots,\ell_m\}$ geldt dat $[\xi_{k_1}^T,\dots,\xi_{k_n}^T]^T$ is onafhankelijke van $[\eta_{\ell_1}^T,\dots,\eta_{\ell_m}^T]^T$.

Een stochastisch proces $\{\xi_k\}$ heet een *Markov-proces* als $\forall k_1,\dots,k_n \in \mathbb{N}$, met $k_1 > k_2 > \dots > k_n$ geldt dat

$$P_{X_{k_1}|X_{k_2}\dots X_{k_n}}(x_{k_1}|x_{k_2},\dots,x_{k_n}) = P_{X_{k_1}|X_{k_2}}(x_{k_1}|x_{k_2}).$$

Een stochastisch proces $\{\xi_k\}$ is *stationair* als:

$\forall m \in \mathbb{N}$, $\forall k_1,\dots,k_m \in \mathbb{N}$, $\forall n_0 \in \mathbb{N}$ geldt

$$P_{\xi_{k_1}\dots\xi_{k_m}}(x_{k_1},\dots,x_{k_m}) = P_{\xi_{k_1+n_0}\dots\xi_{k_m+n_0}}(x_{k_1+n_0},\dots,x_{k_m+n_0}),$$

d.w.z. de geïnduceerde dichtsheidsfunkties zijn niet afhankelijk van translaties in de tijd-as.

Een *witte ruis proces* is een stochastisch proces $\{\xi_k\}$ dat aan de volgende eigenschappen voldoet:

(i) De verwachtingswaarde is gelijk aan nul ($E[\xi_k] = 0$, $k \in \mathbb{N}$)

(ii) $E[\xi_k \xi_\ell^T] = C_k \delta_{k\ell}$.

Een stochastisch proces $\{\xi_k\}$ heet *Gaussisch* als geldt dat voor elke verzameling $k_1,\dots,k_n \in \mathbb{N}$, $n \in \mathbb{N}$, de stochastische variabele $[\xi_{k_1}^T,\dots,\xi_{k_n}^T]^T$ Gaussisch is.