

De huizenprijs voorspeld!

Een spectraalanalyse over de schuifdynamiek van leidende
indicatoren

Scriptie voor het doctoraalexamen
algemene economie van Ard den Reijer
begeleiding dr. A. Hoogstrate
Tilburg, 27 oktober 2000

Voorwoord

Een afstudeerproject is bij uitstek het moment in de studie waar de theorie in de praktijk kan worden toegepast. Bij mij werden het de huizenprijzen. Van de Rabobank kreeg ik de gelegenheid om gedurende een stage de uitdaging aan te gaan om de huizenprijs te voorspellen. Graag wil ik hierbij alle collega's van de Stafgroep Economisch Onderzoek bedanken voor de prettige samenwerking. In het bijzonder mijn stagebegeleider Wim van Assenbergh en de mede-stagiair Jitske Jettinghoff. De begeleiding van het academische gehalte van de scriptie was handen van André Hoogstrate. De ontmoetingen met hem waren altijd zo verfrissend dat ik na afloop weer volop goede ideeën had. André, bedankt daarvoor. Professor Van der Genugten wil ik bedanken voor zijn bereidheid om als tweede begeleider plaats te nemen in de examencommissie. Tot slot wil ik alle mensen hartelijk bedanken zonder wiens steun ik niet herboren zou zijn, nadat ik de laatste deadline niet overleefde.

Ard den Reijer,

HOOFDSTUK1: INLEIDING	3
HOOFDSTUK 2: TIJDREEKSDECOMPOSITIE	6
2.1: Stochastische differentievergelijkingen	6
2.2: Stationariteit	9
2.3: De trendcomponent	9
2.3.1: Deterministische trends	10
2.3.2: Stochastische trends	10
2.3.3: Stochastisch en deterministisch: random walk plus drift	12
2.4: Toetsen voor eenheidswortels	12
HOOFDSTUK 3: SPECTRAALANALYSE	14
3.1: Fourier Analyse: een introductie	14
3.1.1: Systemen van orthogonale functies	15
3.1.2: Fourier-reeksen	18
3.2: Univariate spectraalanalyse	21
3.2.1: Stationaire processen	21
3.2.2: Het spectrum: interpretatie en toepassing	23
3.2.3: Filters	26
3.3: Multivariate spectraalanalyse	30
3.3.1: Multivariate generalisatie van de spectraaldefinities	30
3.3.2: Multivariate extensie: samenhang en verschuiving	33
3.3.3: Bivariate spectraalanalyse toegepast op het algemene lineaire model	36
3.3.4: Het multivariate algemene lineaire model en partiële maatstaven	40
3.4: De spectraalschatter	41
3.4.1: Het periodogram en de impliciete inconsistentie	42
3.4.2: De kernelschatter: een consistente gewogen integraal van het periodogram	44
HOOFDSTUK 4: DE CONJUNCTUUR VOORSPELD	49
4.1: Beschrijving data	49
4.2: Conjunctuurbepaling	51
4.2.1: Conjunctuurbepaling met deterministische trend, de barometrische methode	51
4.2.1.1: Trendalgoritme	51
4.2.1.2: Conjunctuurbepaling en interpretatie	53
4.2.1.2.1: Huizenprijsconjunctuur	53
4.2.1.2.2: Interpretatie van de invloed van standaardisatie op de huizenprijsconjunctuur	54
4.2.1.2.3: Conjunctuur van de overige reeksen	56
4.2.2: Conjunctuurbepaling met stochastische trend, de differentiemethode	57
4.3: Voorspellen van de conjunctuur	58
4.3.1: Barometrische voorspeller van de conjunctuur	59
4.3.2: Spectraalvoorspeller van de conjunctuur	60
4.3.2.1: De bandbreedte van de spectraalschatter	61
4.3.2.2: De selectie van de reeksen	62
4.3.2.3: Het schatten van het leidende cyclische indicatorenmodel	63

HOOFDSTUK 5: DE HUIZENPRIJS VOORSPELD, EEN EVALUATIE	65
5.1: Van conjunctuurvoorspelling naar puntvoorspelling	65
5.1.1: De barometrische voorspeller; deterministische trend	65
5.1.2: De spectraalvoorspeller; stochastische trend	65
5.2: De kwaliteit van de voorspellers, een terugtestevaluatie	66
BIBLIOGRAFIE	69
APPENDIX A: UITWERKINGEN EN BEWIJZEN SPECTRAALANALYSE	70
APPENDIX B: DOCUMENTATIE PROCEDURES EN RESULTATEN	82
B.1: Documentatie procedure en resultaten deterministische werkwijze	82
B.1.1: Conjunctuur bepalen met deterministische trend	82
B.1.2: Conjunctuur voorspellen	88
B.2: Documentatie procedure en resultaten stochastische werkwijze	88
B.2.1: Conjunctuur bepalen met stochastische trend	88
B.2.2: Selectie van de leidende cyclische indicatoren	91
APPENDIX C: MATLAB-APPLICATIE	93
APPENDIX D: ESB-ARTIKEL "PRIJS JE RIJK"	98

Hoofdstuk1: Inleiding

De economie is een dynamisch proces met allerlei variabelen die elkaar onderling beïnvloeden. Hierbij valt te denken aan macro-economische grootheden als rente, wisselkoers, inflatie, beurskoers, bedrijfswinsten, nationale productie, geldhoeveelheid, etc. Er bestaan talloze theorieën over de onderlinge samenhang en causaliteitsrelaties tussen de variabelen, die tevens vaak het gedrag van een variabele in de loop van de tijd als gevolg van een impuls in een van de variabelen beschrijven. Een gebied waar de postulering van deze 'lead-lag'-relaties veel gebezigd wordt is bijvoorbeeld de monetaire economie, die stelt dat een vergroting van de geldhoeveelheid op dit moment tot inflatie in de (nabije) toekomst kan leiden. Voor beleidsmakers en commerciële instellingen is het nuttig om te weten hoe sterk relaties zijn en met welke timing ze elkaar beïnvloeden.

Deze studie beperkt zich tot interacties in een woningmarktspecifieke context van macro-economische en huizenmarkt gerelateerde variabelen. De bestaande ideeën omtrent het functioneren van de woningmarkt beschrijven een marktdynamiek, die wordt veroorzaakt door interactie van macro-economisch georiënteerde elastische vraagfactoren en starre, vertraagd reagerende aanbodfactoren. De prijsvorming is het resultaat van vraag en aanbod en reageert als gevolg eveneens vertraagd op de ontwikkelingen van voornamelijk vraaggeoriënteerde variabelen. Deze constatering impliceert het bestaan van variabelen die een voorspellende waarde hebben voor het toekomstige prijspeil van woningen. Dergelijke variabelen worden leidende indicatoren genoemd. Cruciaal voor een leidende indicator is in hoeverre zijn beweging samenhang vertoont met de beweging van de huizenprijs en in hoeverre zijn ontwikkeling voorloopt op de ontwikkeling van de huizenprijs. Het is bijvoorbeeld interessant te weten hoe lang het duurt voordat een verandering in de hypotheekrente doorwerkt in de huizenprijs en in welke mate dat gebeurt.

Deze studie beoogt de constructie van een huizenprijsvoorspeller voor de korte tot middellange termijn op basis van het voorlopende karakter van met de woningmarkt samenhangende variabelen. In plaats van het postuleren en verifiëren van modelmatige economische theorie wordt het onderzoek naar relaties gedaan met behulp van statistische technieken. De leidende indicatoren zijn de bouwstenen van de huizenprijsvoorspeller en worden behalve als determinanten van de huizenprijs tevens geselecteerd op hun voorspellende waarde.

De cyclische beweging van een tijdreeks is de beweging op korte tot middellange termijn en als voorspeller voor deze periode is de huizenprijsvoorspeller geconstrueerd op basis van de cyclische bewegingen van de leidende indicatoren. In hoofdstuk 2 wordt de klassieke ontleding van tijdreeksen in vier componenten beschreven. Met de focus van de huizenprijsvoorspeller op de cyclische component van een tijdreeks is de kernvraag op welke wijze deze component uit een tijdreeks geïsoleerd kan worden. Hoofdstuk 2 concentreert zich op de eliminatie van de structurele trendmatige beweging van een tijdreeks op lange termijn. Er worden twee typen van trendvormen behandeld, inclusief de onderliggende concepten van de verschillende typen en de wijze waarop een trendvorm uit een reeks gefilterd kan worden. Het eerste type is de deterministische trendvorm, die een functie is van de tijd en waarvan de lineaire en de polynomiale trend de meest bekende uitingen zijn. Het tweede type trendvorm is de stochastische trend en is feitelijk een sommatie van gerealiseerde stochasten. De resterende beweging van een voor de trend gecorrigeerde tijdreeks bestaat uit de overige drie componenten en beschikt afhankelijk van de gekozen trendvorm over bepaalde statistische eigenschappen. De mede overgebleven cyclische component van een tijdreeks is afhankelijk van de gekozen trendvorm al dan niet economisch interpreteerbaar.

De gehanteerde methodiek bestaat uit de constructie van een huizenprijsvoorspeller aan de hand van een lineaire combinatie van cyclische leidende indicatoren. Nu in hoofdstuk 2 de cyclische beweging van een tijdreeks bepaald is, wordt er in hoofdstuk 3 gereedschap aangedragen om invulling te geven aan de methodiek. De spectraalanalyse levert het gereedschap in de vorm van maatstaven, die zowel de samenhang als de mate van voorlooptijd van een tijdreeks ten opzichte van de huizenprijsreeks

kwantificeren. Bovendien biedt de spectraalanalyse een raamwerk om het lineaire leidende indicatorenmodel integraal te schatten. In hoofdstuk 3 wordt de spectraalanalyse gedefinieerd, opgebouwd en aan de hand van voorbeelden geïllustreerd. Het kernidee achter de spectraalanalyse is de ontleding van een tijdreeks in periodieke golf functies, waarbij een tijdreeksvariabele gerepresenteerd kan worden als een gewogen som van golven met verschillende golflengtes. Het spectrum beschrijft een tijdreeks als functie van de frequentie en biedt hiermee een raamwerk om het gedrag van een tijdreeks te ontleden naar intervallen, die representatief zijn voor de korte, middellange en lange beweging. Deze ontleding naar intervallen is tevens van toepassing op de gedefinieerde spectraalmaatstaven, waardoor het mogelijk is de schuifdynamiek van leidende indicatoren te analyseren.

Met de methodiek en het gereedschap is het theoriegedeelte afgerond en deze worden in hoofdstuk 4 toegepast voor de constructie van een voorspeller voor de cyclische beweging van de huizenprijs. Voortbouwend op de twee typen van trendmatige beweging worden er twee typen huizenprijsvoorspellers geconstrueerd en met elkaar vergeleken.

Het eerste type voorspeller wordt geconstrueerd aan de hand van voor stochastische trendvorm gecorrigeerde tijdreeksvariabelen. Met spectraalanalyse worden de leidende indicatoren geïdentificeerd en het leidende indicatorenmodel voor de cyclische beweging geschat. Op deze wijze wordt er een spectraalvoorspeller voor de cyclische beweging van de huizenprijs verkregen.

Het tweede type voorspeller wordt geconstrueerd aan de hand van voor deterministische trendvorm gecorrigeerde tijdreeksvariabelen en staat in de traditie van de conjunctuurindicatoren, die een barometrische indicatie geven voor de conjuncturele toestand van de economie. De barometrische methode genereert een conjunctuurdiagnose van de algemeen economische toestand op basis van een niet-causale, mechanische wijze, die geen specifiek economisch theoretische grondslag kent. De methodiek van de barometrische traditie is succesvol gebleken in een algemeen economische omgeving en de deterministische voorspeller van de huizenprijsconjunctuur is een toepassing van de barometrische methodiek in een marktspecifieke context. Hierbij is conjunctuur synoniem aan cyclische beweging. De potentie van de marktspecifieke toepassing wordt ingegeven door de vaststelling dat de Nederlandse huizenmarkt door zijn omvang een macro-economische factor van betekenis is. De barometrische methodiek bestaat uit toepassing van een aantal standaard statistische technieken. Door een deterministische trendvorm te berekenen is het mogelijk een conjunctuurreeks uit een tijdreeks te isoleren, waarvan de beweging interpreteerbaar is in termen van economische en marktspecifieke ontwikkelingen. De samenhang en de verschuiving van een conjunctuurreeks ten opzichte van de huizenprijsconjunctuurreeks wordt bepaald door de kruislingse correlaties en de visuele, kwalitatieve overeenkomst van conjuncturele omslagpunten. Aan de hand van criteria voor samenhang en verschuiving worden er leidende conjunctuurindicatoren geselecteerd, die de met regressieanalyse gewogen componenten vormen van de voorspeller voor de huizenprijsconjunctuur.

De uiteindelijke onderzoeksvraag is het genereren van een puntvoorspelling voor de huizenprijs in de nabije toekomst. Deze studie beoogt de constructie van een huizenprijsvoorspeller voor de korte tot middellange termijn. De bestaande ideeën omtrent het functioneren van de woningmarkt bieden ruimte voor het bestaan van leidende indicatoren voor de cyclische beweging van de huizenprijs. Door zowel de barometrische methodiek als door toepassing van de spectraalanalyse is met een leidende cyclische indicatorenmodel de huizenprijsconjunctuur voorspeld. In hoofdstuk 5 worden uiteindelijk een barometrische voorspeller en een spectraalvoorspeller van de oorspronkelijke huizenprijs verkregen door respectievelijk de conjunctuurvoorspellers met de bijbehorende trendvoorspellers terug te koppelen naar de oorspronkelijke reeks. De trendmatige component van een tijdreeks kent per definitie een starre en relatief onveranderlijke ontwikkeling en de trendvoorspeller is een extrapolatie van deze starre en relatief onveranderlijke ontwikkeling. Een meer precieze duiding van de trendvoorspeller hangt af van de keuze voor een deterministische of een stochastische trendvorm.

Ter evaluatie wordt de relatieve kwaliteit van beide puntvoorspellers voor de huizenprijs verkregen door voorspellingen voor het nabije verleden te genereren en te berekenen in hoeverre de voorspelde en de gerealiseerde huizenprijs verschillen. Dit wordt ook wel 'out-of-sample' terugtesten genoemd.

De constructie, beschrijving en de resultaten van de conjunctuurvoorspeller en de puntvoorspeller van de huizenprijs volgens de barometrische methodiek zijn gebaseerd op een praktijkstage. De stage is uitgevoerd bij de *Stafgroep Economisch Onderzoek* van *Rabobank Nederland* te Utrecht. De Stafgroep is de macro-economische denktank binnen de Rabobankorganisatie en verricht onderzoek op het gebied van financieel-economische onderwerpen ten behoeve van de eigen organisatie, de lokale banken en hun klanten. Het conjunctuurcluster binnen de Stafgroep richt zich op de macro-economische ontwikkeling in de industrielanden. Mede op basis van de resultaten van deze stage is een artikel verschenen in de *Economisch Statistische Berichten* (ESB) van 28-7-2000, met als titel 'Prijs je rijk'. Dit artikel is opgenomen in appendix D van deze scriptie.

Hoofdstuk 2: Tijdreeksdecompositie

De doelstelling van het voorspellen van een tijdreeksvariabele op korte tot middellange termijn op basis van synchrone en leidende cyclische, ofwel conjuncturele bewegingen van aanverwante tijdreeksvariabelen vereist een precieze duiding van de cyclische en overige componenten waaruit een tijdreeks is opgebouwd. De klassieke methodologie¹ van de tijdreeks-econometrie bestaat uit de decompositie van een reeks in vier componenten: trend, cyclisch, seizoen en transitoir. De waarnemingen van een tijdreeks zijn realisaties van het onderliggende datagenererende economische proces, dat in de praktijk meestal onbekend is. Deze onbekendheid van het onderliggende proces wordt ondervangen door het genererende proces stochastisch te modelleren. De tijdreeksvariabele is dan een functie van zijn realisaties uit het verleden, tijd en een stochastische component. De trendcomponent is vaak deterministisch van aard en wordt in een tijdreeksvergelijking uitgedrukt als functie van de tijd.

In dit hoofdstuk zal echter worden aangetoond dat zelfs de trendmatige beweging, die een permanente, niet-wegebbende invloed op een tijdreeks uitoefent, vaak stochastisch beschreven kan worden. Het wegfilteren van de trendmatige component uit een tijdreeks resulteert in een reeks, die statistisch gezien een stabiel gedrag vertoont en een beschrijving geeft op basis van de overige drie componenten van het gedrag van de tijdreeks op korte en middellange termijn. Uiteindelijk wordt er een transformatie en een statistische testprocedure beschreven, waarmee zowel stochastische als deterministische trends geëlimineerd kunnen worden.

2.1: Stochastische differentievergelijkingen

Beschouw de functie $y=f(t)$. Als de functie beschouwd wordt op het moment dat de onafhankelijke variabele t de specifieke waarde t^* aanneemt genereert deze functie een specifieke waarde voor de afhankelijke variabele y , dus $y_{t^*}=f(t^*)$. Op dezelfde wijze kan de volgende specifieke waarde worden gedefinieerd: $y_{t^*+h}=f(t^*+h)$. Als praktische kwestie worden de meeste economische tijdreeksdata voor discrete tijdsperioden verzameld. Hiervoor worden de equidistante intervallen t^* , t^*+h , t^*+2h , t^*+3h ,... beschouwd, waarbij h genormaliseerd wordt tot 1. Een discrete tijdreeks impliceert dat t discreet is en niet noodzakelijk dat de onderliggende variabele ook een discreet karakter kent. Zo is de regenneerslag in Nederland een continue variabele, maar de op jaarbasis gecumuleerde neerslagtotalen voor de jaren 1 tot t een discrete tijdreeks. In veel economische toepassingen is de variabele t een representatie van 'tijd', zodat h een verandering in tijd, ofwel de lengte van een tijdsperiode vertegenwoordigt. De variabele t hoeft echter niet noodzakelijk gerelateerd te zijn aan het type tijdsinterval zoals dat wordt gemeten door een klok of een kalender. In plaats van jaren, maanden, dagen of minuten kan t ook gerefereerd worden als een ordinale maat aan een geordende gebeurtenis. Zo kan bijvoorbeeld y_t de uitkomst zijn van de t draaiing van een rouletterad.

Het is gebruikelijk de elementen van een waargenomen tijdreeks $\{\dots, y_{t-2}, y_{t-1}, y_t, y_{t+1}, y_{t+2}, \dots\}$ uit te drukken als $\{y_t\}$. De uitdrukking van een specifiek element in de reeks is dan y_t . Tenzij anders is gespecificeerd loopt de index t van $-\infty$ tot ∞ .

Het eerste orde verschil van y_t is gedefinieerd als de waarde van de functie geëvalueerd op tijdstip $t=t^*+1$ minus de waarde van de functie gedefinieerd op tijdstip t^* , dat wil zeggen:

$$\Delta y_t \equiv y_t - Ly_t = y_t - y_{t-1}$$

$$\Delta y_{t+1} \equiv y_{t+1} - Ly_{t+1} = y_{t+1} - y_t$$

Hierbij is de vertragingsoperator L een zodanige lineaire bewerker dat voor elke waarde y_t geldt:

$$Ly_t = y_{t-1} \quad (2.1.1)$$

¹ Dit is de zogenaamde 'Harvard method of time series analysis', Boschan & Ebanks (1970)

De $(L)^2$ voorafgaand aan y_t betekent simpelweg het vertragen van y_t met t perioden. De differentiebewerker Δ is aan de vertragsingsbewerker gerelateerd en gedefinieerd als:

$$\Delta \equiv (1-L). \quad (2.1.2)$$

Het tweede orde verschil kan op analoge wijze als volgt worden aangeduid:

$$\Delta^2 y_t \equiv \Delta(y_t - L(y_t)) = y_t - 2L(y_t) + L^2(y_t) = y_t - 2y_{t-1} + y_{t-2}$$

$$\Delta^2 y_{t-1} \equiv \Delta(y_{t-1} - L(y_{t-1})) = y_{t-1} - 2L(y_{t-1}) + L^2(y_{t-1}) = y_{t-1} - 2y_{t-2} + y_{t-3}$$

Het n^e -orde verschil is analoog gedefinieerd³, maar in de praktijk van economische tijdreeksen wordt er zelden gebruik gemaakt van n^e -orde verschillen, waarbij $n > 2$.

In de bovenstaande introductie is de variabele y louter beschouwd als een realisatie van een functie op een bepaald tijdstip. De functie f wordt nu zodanig uitgebreid dat in zijn meest algemene vorm de waarde van een variabele uitgedrukt wordt als een functie van de vertraagde waarden van deze variabelen, het tijdstip t en andere variabelen. Als f een lineaire functie is en als er sprake is van p vertraagde variabelen wordt er als volgt een p^e orde lineaire differentievergelijking verkregen:

$$y_t = f(y_{t-1}, \dots, y_{t-p}, t, x) \equiv a_0 + \sum_{i=1}^p a_i y_{t-i} + x_t = a_0 + \sum_{i=1}^p a_i L^i y_t + x_t \quad (2.1.3)$$

Economische theorie is in verscheidene situaties in staat om bepaalde a_i 's te dicteren als functie van variabelen in de economie. Zolang de a_i 's echter geen functie zijn van y_t of x_t , zoals in deze studie, worden ze beschouwd als parameters. Indien x_t een enkelvoudige stochast is en voor periode t uitgedrukt kan worden als een storingsvariabele wordt vergelijking (2.1.3) ook wel autoregressief van orde p , AR(p), genoemd, waarbij $p \in \{0, 1, 2, \dots, \infty\}$.

Het witte ruis-proces is een model dat stochastiek in de differentievergelijkingen introduceert en is daarmee een basisbouwsteen voor discrete stochastische tijdreeksmodellen. Een reeks $\{e_t\}$ is witte ruis als elk element in de reeks een verwachte waarde van nul heeft, een constante variantie kent en ongecorrleerd is met zijn vertraging. Als $E\{x\}$ de verwachte theoretische gemiddelde waarde van x is geldt formeel dat de reeks $\{e_t\}$ een witte ruis-proces is als voor elke waarde van t en s :

$$\begin{aligned} i) & E\{e_t\} = E\{e_{t-s}\} = 0 \\ ii) & E\{e_t\}^2 = E\{e_{t-s}\}^2 = \sigma^2 & (\text{Var}\{e_t\} = \text{Var}\{e_{t-s}\} = \sigma^2) \\ iii) & E\{e_t e_{t-s}\} = E\{e_{t-j} e_{t-j-s}\} = 0 \quad \forall j & (\text{Cov}\{e_t, e_{t-s}\} = \text{Cov}\{e_{t-j}, e_{t-j-s}\} = 0) \end{aligned}$$

² De vertragsingsbewerker heeft onder andere de volgende eigenschappen:

- 1 $Lc = c$
- 2 distributiviteit: $(L^i + L^j)y_t = L^i y_t + L^j y_t = y_{t-i} + y_{t-j}$
- 3 associativiteit: $L^i L^j y_t = L^{i+j} y_t = y_{t-i-j}$
- 4 L tot een negatieve waarde is feitelijk een voorspellingsbewerker, namelijk:
 $j = -i, L y_t = y_{t+j} = y_{t+i}$
- 5 voor $|a| < 1$ geldt voor de oneindige som $(1 + aL + a^2 L^2 + \dots) y_t = y_t / (1 - aL)$.
Dit is eenvoudig in te zien. Vermenigvuldig beide uitdrukkingen met $(1 - aL)$.
 $(1 - aL)(1 + aL + a^2 L^2 + \dots) y_t = y_t \Leftrightarrow (1 - aL + aL - a^2 L^2 - a^2 L^2 + \dots) y_t = y_t$.
Omdat $a^n L^n y_t$ naar nul convergeert als n naar oneindig gaat is bovenstaande vergelijking een identiteit.
- 6 voor $|a| > 1$ geldt voor de oneindige som $(1 + (aL)^{-1} + (aL)^{-2} + \dots) y_t = -aL y_t / (1 - aL) \Leftrightarrow y_t / (1 - aL) = -(aL)^{-1} \sum_i (aL)^i y_t$. Dit is eenvoudig in te zien. Vermenigvuldig beide uitdrukkingen met $(1 - aL)$.
 $(1 - aL)(1 + (aL)^{-1} + (aL)^{-2} + \dots) y_t = -aL y_t \Leftrightarrow (1 - aL + (aL)^{-1} - 1 + (aL)^{-2} - (aL)^{-2} + \dots) y_t = -aL y_t$.
Omdat $a^n L^n y_t$ naar nul convergeert als n naar oneindig gaat is bovenstaande vergelijking een identiteit.

³ De exacte definitie luidt als volgt:

$$\Delta^n y_t = \Delta^{n-1} y_t - \Delta^{n-1} y_{t-1} = \sum_{r=0}^n (-1)^r \binom{n}{r} y_{t-r}, \quad t = n, n+1, \dots$$

waarbij $\binom{n}{r} = \frac{n!}{r!(n-r)!}$

de binomiaalcoëfficiënten zijn.

In het vervolg zal een witte ruis-proces aangegeven worden in de volgende notatie: $e_t \sim (0, \sigma^2)$.

De term x_t in (2.1.3) is het genererende proces en zijn vorm kan veel algemener zijn dan slechts een enkelvoudige witte ruis-stochast. Het is mogelijk om huidige en vertraagde waarden van andere variabelen op te nemen, de parameter tijd om een deterministische trendbeweging te beschrijven, dummyvariabelen om belangrijke gebeurtenissen op te nemen en vertraagde stochastische storingen. Door een geschikte keuze van het genererende proces kan een uitgebreid scala van dynamische macro-economische modellen gecreëerd worden.

Een belangrijk speciaal geval van de $\{x_t\}$ -reeks is het volgende:

$$x_t = \sum_{i=0}^q b_i e_{t-i} = \sum_{i=0}^q b_i L^i e_t \quad (2.1.4)$$

waarbij $q \in \{0, 1, 2, \dots, \infty\}$, de b_i 's constanten zijn en de $\{e_t\}$ -reeks een witte ruis proces. De reeks (2.1.4) wordt ook wel voortschrijdend gemiddelde van orde q genoemd, ofwel $MA(q)$ ⁴.

Substitutie van (2.1.4) in (2.1.3) combineert tot het $ARMA(p, q)$ -model.

$$y_t = a_0 + \sum_{i=1}^p a_i L^i y_t + \sum_{i=0}^q b_i L^i e_t \quad (2.1.5)$$

Een oplossing van de differentievergelijking (2.1.5) drukt de waarde van y_t uit als een functie van de elementen uit de $\{x_t\}$ -reeks, de tijdsparameter t en eventueel de initiële beginwaarde van de afhankelijke variabele y_0 .

In het algemeen worden aan een oplossing tevens stabiliteitscondities opgelegd, die ervoor zorgdragen dat het systeem voor $t \rightarrow \infty$ niet-explosief is. Vanuit een puur mathematisch oogpunt is er geen reden om convergentie-eisen aan het systeem op te leggen. Economische theorie suggereert echter een begrenzing in de zin dat de limietwaarde van een variabele in een tijdreeks eindig is. Reële rente, reël inkomen per hoofd van de bevolking en andere economische variabelen worden nauwelijks verondersteld de waarde plus of min oneindig aan te nemen.

Onder de voorwaarde dat het $ARMA(p, q)$ -model aan de stabiliteitscondities⁵ voldoet resulteert, gebruikmakend van de vertragsbewerker, de oplossing van een $ARMA(p, q)$ -model in:

$$y_t = \frac{a_0 + \sum_{i=0}^q b_i e_{t-i}}{1 - \sum_{i=1}^p a_i L^i} \quad (2.1.6)$$

In veel gevallen is het niet noodzakelijk de oplossing verder uit te werken om de specifieke coëfficiënt voor elk element in $\{e_t\}$ te verkrijgen. Belangrijk is echter wel de constatering dat de uitwerking van de oplossing feitelijk een $MA(\infty)$ -proces⁶ oplevert.

⁴ de Engelse term voor voortschrijdend gemiddelde is moving average, vandaar de naamgeving $MA(q)$.

⁵ de stabiliteitscondities garanderen de convergentie (dat wil zeggen de niet-explosiviteit) van het dynamische systeem. De condities luiden als volgt: zij

$$(1 - a_1 L - \dots - a_n L^n) = \prod_{i=1}^n (1 - c_i L)$$

Voor een n^e - orde vergelijking is een noodzakelijke voorwaarde dat: $\sum_{i=1}^n c_i < 1$,

een voldoende voorwaarde is, omdat de waarden van c_i zowel positief als negatief kunnen zijn: $\sum_{i=1}^n |c_i| < 1$,

Zie voor een uitputtende analyse Enders (1995) hoofdstuk 1 of Fuller (1996) paragraaf 2.4.

⁶ de Wold-decompositie stelt zelfs dat elke stationaire tijdreeks beschreven kan worden als de som van een voorspelbare (constante) component en een $MA(\infty)$ -proces.

Stabiliteit is echter een noodzakelijke, maar geen voldoende voorwaarde voor stationariteit van een $ARMA$ -proces. Voor stationariteit geldt namelijk de additionele eis dat voor iedere $s \geq 0$ moet gelden $(b_s + b_1 b_{s+1} + b_2 b_{s+2} + \dots) < \infty$. Zie hiervoor Enders (1995) paragraaf 2.4.

2.2: Stationariteit

Stationariteit wordt grosso modo geïnterpreteerd als trendloosheid en is een voorwaarde voor toepassing van de spectraalanalyse. De klassieke tijdreeksdecompositie stelt dat er in een tijdreeks potentieel een trendcomponent aanwezig is. De toepassing van een geschikt filter op een tijdreeks elimineert deze trendcomponent en resulteert in een reeks, die statistisch gezien stationair is. In deze paragraaf wordt de definitie van stationariteit gegeven.

Stationaire stochastische processen zijn een klasse van processen die gekarakteriseerd kunnen worden door het verschijnsel dat hun belangrijkste statistische eigenschappen niet veranderen over de tijd en die in het algemeen karakteristiek zijn voor een systeem in evenwichtstoestand. De meest stringente vorm van stationariteit is de eis dat de totale probabilistische structuur van een proces niet varieert onder een verschuiving van de tijdsgrondslag. Meestal wordt volstaan met de notie van zwakke stationariteit, waarbij niet de volledige probabilistische structuur, maar de belangrijkste eigenschappen van de kansverdeling invariant moeten zijn voor een verschuiving van de tijdsgrondslag.

Ook in deze studie wordt beperkt tot zogenaamde 2^o orde stationaire processen, die als eis een tijdsinvariantie voor de eerste twee momenten oplegt en ook wel covariantie stationair genoemd wordt. De definitie luidt als volgt:

Een proces $\{y_t\}$ is 2^o orde stationair als voor alle toelaatbare t en $(t-s)$ geldt:

$$\begin{aligned} i) \quad & E\{y_t\} = E\{y_{t-s}\} = \mu \\ ii) \quad & E\{y_t - \mu\}^2 = E\{y_{t-s} - \mu\}^2 = \sigma_x^2 \quad (\text{Var}\{y_t\} = \text{Var}\{y_{t-s}\} = \sigma^2) \\ iii) \quad & E\{y_t - \mu\}E\{y_{t-s} - \mu\} = E\{y_{t-j} - \mu\}E\{y_{t-j-s} - \mu\} = \gamma(s) \quad (\text{Cov}(y_t, y_{t-s}) = \text{Cov}(y_{t-j}, y_{t-j-s}) = \gamma(s)) \end{aligned}$$

De functie $\gamma(s)$ wordt de autocovariantie van y_t genoemd en als er geen sprake is van mogelijke verwarring wordt dit afgekort tot covariantie. Om de eigenschappen van tijdreeksen te vergelijken is er een gestandaardiseerde maatstaf gedefinieerd, die niet wordt beïnvloed door de maateenheid van de onderliggende tijdreeksvariabele. Deze maatstaf is de autocorrelatiefunctie, veelal afgekort tot correlatie, en is gedefinieerd door:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\gamma(h)}{\sigma^2} \quad (2.2.1)$$

De correlatie is de covariantie genormaliseerd tot 1 als $h=0$. De eis voor stationariteit dat het gemiddelde en de variantie constant en tijdsonafhankelijk zijn betekent dat een tijdreeks noch explosief is noch uitdempt, waardoor dit type tijdreeks geschikt is om de cyclische, seizoensmatige beweging en de ruiscomponent van een economische tijdreeksvariabele te beschrijven.

2.3: De trendcomponent

Veel economische tijdreeksen zijn van nature niet stationair van karakter, zoals bijvoorbeeld industriële productie, nationaal inkomen en het prijsindexcijfer. Deze reeksen worden gekenmerkt door een gemiddelde waarde die een sterke opwaartse beweging over de tijd vertoont. Het reële nationale inkomen kent bijvoorbeeld een opgaande tendens. Deze kan door een lineaire trend als functie van de tijd in een vergelijking beschreven worden. Deze methode is echter controversieel omdat het een deterministische groeivoet van de reële economie impliceert. Er bestaat een macro-economische stroming die stelt dat technologische innovaties een permanente invloed hebben op de trend van de reële economie. Bovendien stelt deze stroming dat technologische innovaties stochastisch zijn en dat de trend deze onderliggende stochastiek moet reflecteren. De stochastische visie kan tevens met betrekking tot de huizenprijzen en de andere in deze studie gebruikte tijdreeksen gepostuleerd worden. Voor de huizenprijzreeks kunnen institutionele veranderingen als bijvoorbeeld fiscale voorzieningen, beleid met betrekking tot huurprijzen en toewijzing van vinx-locaties beschouwd worden als stochastische gebeurtenissen, die op toevallige momenten plaatsvinden, maar wel een grondige invloed uitoefenen op de huizenprijs. Het is dan ook nuttig om zowel modellen met een deterministische trend als modellen met een stochastische trend te beschouwen.

2.3.1: Deterministische trends

De deterministische trend is de structurele beweging van een reeks, die wordt beschreven door een louter van de tijdsvariabele t afhankende functie. Door de afhankelijkheid van t wordt deze trend tijdtrend genoemd. Vergelijking (2.3.1) wordt kwadratische tijdtrend en in het geval dat $a_2=0$ lineaire tijdtrend genoemd.

$$y_t = a_0 + a_1 t + a_2 t^2 + e_t \quad (2.3.1)$$

Voor economische tijdreeksen wordt in de praktijk zeer zelden gebruik gemaakt van een t^n -term met $n > 2$. Het model kan worden uitgebreid met zowel vertraagde variabelen van de $\{y_t\}$ -reeks als van de $\{e_t\}$ -reeks.

De meest voor de hand liggende methode om de trend in (2.3.1) te elimineren is met een regressie van y_t op een gespecificeerde trendfunctie een schatting voor de trend te genereren en deze af te trekken van de werkelijke reeks. De overgebleven reeks is dan stationair en is voor (2.3.1) de $\{e_t\}$ -reeks. Merk op dat in de praktijk het datagenererende proces niet altijd volledig bekend is, waardoor er het risico van misspecificatie bestaat.

Een tweede methode om de met (2.3.1) gegenereerde tijdreeks stationair te maken is door het tweede orde verschil ervan te berekenen.

$$\begin{aligned} \Delta y_t &= a_0 + a_1 t + a_2 t^2 + e_t - (a_0 + a_1(t-1) + a_2(t-1)^2 + e_{t-1}) = a_1 - a_2 + 2a_2 t + e_t - e_{t-1}, \\ \Delta^2 y_t &= 2a_2 + e_t - 2e_{t-1} + e_{t-2}, \end{aligned} \quad (2.3.2)$$

Met (2.3.2) is eenvoudig in te zien dat de tweede differentiereeks⁷ van y_t stationair is. Bovendien is de autocovariantie over meer dan twee perioden van deze reeks gelijk aan nul en dat impliceert dat de autocovarianties absoluut sommeerbaar zijn. Stationariteit en absolute sommeerbaarheid zijn de twee voorwaarden voor toepassing van de spectraalanalyse⁸ en de tweede differentievergelijking (2.3.2) kwalificeert zich voor deze toepassing.

In figuur 2.1.1 is het tweede orde tijdtrend-model (2.3.1) weergegeven ($a_0=0$, $a_1=0.3$, $a_2=0.01$, $e_t \sim (0,1)$) samen met bijbehorende tweede orde differentiereeks (2.3.2). In de grafiek is duidelijk de opwaartse neiging van de tijdreeks (linkeras) in de vorm van een tweedegraads kromming zichtbaar. De bijbehorende differentiereeks (rechteras) toont duidelijk geen opwaartse of neerwaartse neiging en ook zijn beweeglijkheid is constant over de tijd, waarmee grafisch het stationaire karakter van de reeks tot uitdrukking komt.

2.3.2: Stochastische trends

Het essentiële kenmerk van een trend is dat het een permanent effect heeft op een tijdreeks. In termen van de in de inleiding gepresenteerde decompositie is de trend de niet-stationaire component en zijn de cyclische, de transitoire en de seizoenscomponent wel stationair. Deze niet-stationariteit komt bij een stochastische trend tot uitdrukking doordat de realisatie van een stochast in de loop van de tijd niet wegebt. Terwijl de deterministische trend zich op een a priori bekende wijze over de tijd ontwikkelt en daardoor beschreven kan worden als functie van t , wordt de stochastische trend ex ante weergegeven als een sommatie van stochasten, die pas ex post vorm krijgt als de sommatie van niet wegebbende realisaties van deze stochasten.

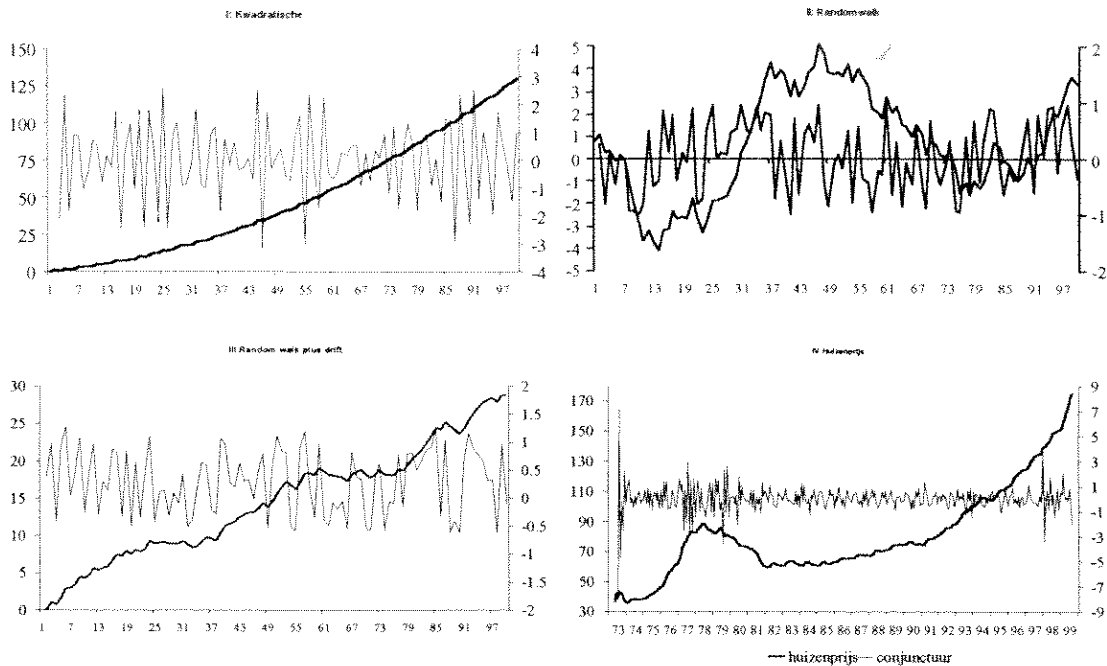
⁷ In het algemeen wordt het onderstaande n^e -orde polynomiale tijdtrend-model stationair gemaakt door de n^e -differentiereeks te bepalen, namelijk:

$$y_t = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n + e_t,$$

$$\Delta^n y_t = n a_n + \sum_{j=0}^n (-1)^j \binom{n}{j} e_{t-j}.$$

⁸ Merk op dat hoewel stationariteit voldoende is voor toepassing van de spectraalanalyse de tweede differentiereeks niet invertierbar is in de zin dat de reeks uitgedrukt kan worden in de vorm van een autoregressief proces. Dit komt door de (meerdere) eenheidswortel in de MA-component, zie Enders (1995) pagina 179. Het gevolg is dat de reeks niet invertierbar en de hele Box-Jenkins ARIMA-methode niet toepasbaar is.

Figuur 2.1: Tijdreeksen en stationaire reeksen



De meest eenvoudige verschijningsvorm van een stochastische trend is in het random walk-model, dat is gespecificeerd als een ARMA(1,0)-proces met $a_0=0$ en $a_1=1$:

$$y_t = y_{t-1} + e_t \quad (\text{ofwel } \Delta y_t = e_t) \quad (2.3.3)$$

De oplossing van deze eerste orde differentievergelijking wordt met herhaalde substitutie en met initiële beginwaarde y_0 verkregen en is gelijk aan:

$$y_t = y_0 + \sum_{i=1}^t e_i \quad (2.3.4)$$

Het a priori constante gemiddelde van de random walk blijkt uit de gelijkheid van de verwachte waarde op tijdstip t en tijdstip $t-s$, namelijk $E\{y_t\} = E\{y_{t-s}\} = y_0$. Het random walk-model is echter niet stationair blijkend uit de variantie, die met deze oplossing eenvoudig is te berekenen:

$$\text{var}\{y_t\} = \text{var}\{e_t + e_{t-1} + \dots + e_1\} = t\sigma^2,$$

en

$$\text{var}\{y_{t-s}\} = \text{var}\{e_{t-s} + e_{t-s-1} + \dots + e_1\} = (t-s)\sigma^2$$

Vanwege het niet constante karakter van de variantie over de tijd (d.w.z. $\text{var}(y_t) \neq \text{var}(y_{t-s})$) is het random walk-proces niet stationair en bovendien gaat de variantie voor $t \rightarrow \infty$ naar oneindig, waarmee de random walk voortkronkelt zonder een expliciet stijgende of dalende tendens te tonen. De stochastische trend in het random walk-model wordt geëlimineerd door de differentiereeks te bepalen. Deze differentiereeks $\Delta y_t = e_t$ is stationair en uitgezet (rechteras) in figuur 2.1.II samen met de random walk-reeks (linkeras), waarbij wederom geldt $y_0=0$ en $e_t \sim N(0,1)$.

Het random walk-model is geschikt voor het beschrijven van tijdreeksen die geen duidelijke neiging hebben in de tijd te stijgen of te dalen en tevens niet neigen naar een bepaald gegeven gemiddelde. In de economie laten vooral wisselkoersreeksen en reeksen met vertrouwensindicatoren een dergelijk patroon zien.

2.3.3: Stochastisch en deterministisch: random walk plus drift

Het volgende model genereert een trend, die is opgebouwd uit zowel een deterministisch als een stochastisch deel. Het random walk plus drift-model (RWPD) is een ARMA(1,0)-model met zowel $a_0=1$ als $a_1=1$ en breidt het random walk-model uit door toevoeging van een constante term a_0 , namelijk:

$$y_t = y_{t-1} + a_0 + e_t \quad (2.3.5)$$

Met de gegeven initiële beginvoorwaarde y_0 is de oplossing:

$$y_t = y_0 + a_0 t + \sum_{i=1}^t e_i \quad (2.3.6)$$

Het gedrag van y_t wordt nu voornamelijk geleid door twee niet-stationaire componenten: een lineaire deterministische trend $a_0 t$ en de stochastische trend $\sum e_i$. Voor de a priori verwachting van y geldt: $Ey_t = y_0 + a_0 t$ en $Ey_{t+s} = y_0 + a_0(t+s)$. De deterministische verandering voor elke realisatie van $\{y_i\}$ is a_0 , waardoor na t perioden de gecumuleerde deterministische verandering gelijk is aan $a_0 t$. Bovendien is er de stochastische trend $\sum e_i$, waarbij iedere gerealiseerde e_i net als in het pure random walk-model, een permanent effect heeft y_t .

Zowel de reeks van het random walk plus drift-model (linkeras) met $a_0=0.3$, $y_0=0$ en $e_t \sim (0,1)$, als de eerste differentiereeks van dit model (rechteras) zijn uitgezet in figuur 2.1.III. Merk op dat de eerste differentiereeks van het RWPD-model stationair is en dus zowel de stochastische als de deterministische trend uitfiltert.

Uit de grafiek blijkt dat de gesimuleerde reeks van het random walk plus drift-model gelijkenissen vertoont met het lineaire tijdtrend-model. Het RWPD-model beschrijft echter een grotere variatie aan beweging, omdat in dit model de sommatie van stochastische realisaties de variabiliteit bepaalt. In het lineaire tijdtrend-model wordt daarentegen de afwijking van de reeks ten opzichte van de lineaire trend bepaald door de realisatie van slechts een enkele stationaire component.

Tenslotte staat in figuur 2.1.IV de huizenprijsindexreeks weergegeven (linkeras) samen met de bijbehorende tweede orde differentiereeks (rechteras). Merk op dat het voor deze reeks in vergelijking met de gesimuleerde reeksen van de figuur moeilijk is om a priori te bepalen of er sprake is van een deterministische of stochastische trend.

2.4: Toetsen voor eenheidswortels

In het voorgaande is aangetoond dat in een tijdreeks potentieel een trendcomponent aanwezig is, die uit de reeks gefilterd wordt door de n^e -orde differentiereeks te bepalen. Deze differentiereeks is statistisch gezien stationair en wordt geïnterpreteerd als trendloos. Stationariteit is naast de absolute sommeerbaarheid van autocovarianties een voorwaarde voor toepassing van spectraalanalyse. Tot slot van dit hoofdstuk wordt een formele testprocedure beschreven, waarmee de stationariteit van een tijdreeks getoetst kan worden en welke orde differentie nodig is om een tijdreeks stationair te maken.

Om enig inzicht te krijgen in de werking van deze procedures beschouwen we eerst het meest eenvoudige model:

$$y_t = a_1 y_{t-1} + e_t \quad (2.4.1)$$

waarbij $e_t \sim (0, \sigma^2)$.

Dit model is stationair indien $|a_1| < 1$ en is een random walk-model indien $a_1 = 1$. In het laatste geval wordt gezegd dat er sprake is van een eenheidswortel en impliceert een niet-stationaire term die stochastische trend genoemd wordt.

De onderzoeksvraag is dan wanneer de nulhypothese $a_1 = 1$ verworpen kan worden. De klassieke statistische methoden voor het schatten en toetsen van a_1 gelden niet onder de nulhypothese vanwege het bestaan van een stochastische trend en de impliciete niet-stationariteit van het random walk-model.

Dickey en Fuller⁹ tonen aan dat onder de voorwaarde dat $\alpha_1=1$ de klassieke schatter voor α_1 een naar beneden afwijkende onzuiverheid kent. Bovendien convergeert deze schatter sneller naar zijn waarschijnlijkheidslimiet dan op basis van de klassieke methoden verwacht wordt. Als gevolg van deze twee verschijnselen kunnen de conventionele tests ten onrechte tot een verwerping van de nulhypothese $\alpha_1=1$ leiden. De verdelingen van de bijpassende toetsingsgrootheden zijn niet standaard en kunnen niet analytisch afgeleid worden. Dickey en Fuller creëerden een praktische oplossing door met Monte Carlo simulaties deze verdelingen en de bijbehorende tabellen met kritische waarden te genereren. De tabellen zijn variabel met betrekking tot het significantieniveau, het aantal waarnemingen en drie gebruikte regressievergelijkingen. Dickey-Fuller toetsprocedures zijn ontwikkeld voor het toetsen van de hypothese $\alpha_1=1$ voor een drietal modellen. Ten eerste het random walk-model met de nulhypothese van een stochastische trend, ook wel eenheidswortel genoemd, tegen de alternatieve hypothese van stationariteit met een verwachting van nul. Ten tweede het random walk plus drift-model met de nulhypothese van een eenheidswortel tegen het alternatief van stationariteit. Ten derde het random walk plus drift-model inclusief de incorporatie van een deterministische tijdtrendterm a_2t , waarbij de nulhypothese wederom de eenheidswortel bevat en de alternatieve hypothese bestaat uit de deterministische tijdtrend. De drie modellen bestaan tevens uit het passende aantal vertragingen¹⁰ van de afhankelijke variabele en zijn weergegeven als respectievelijk (2.4.2), (2.4.3) en (2.4.4).

$$\Delta y_t = \gamma y_{t-1} + \sum_{i=2}^p b_i \Delta y_{t-i+1} + e_t, \quad (2.4.2)$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=2}^p b_i \Delta y_{t-i+1} + e_t, \quad (2.4.3)$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \sum_{i=2}^p b_i \Delta y_{t-i+1} + e_t. \quad (2.4.4)$$

De passende vertragingenslengte (p) kan met standaard t -tests en F -tests bepaald worden. De Dickey-Fuller test bestaat uit het schatten van een of meerder van bovenstaande vergelijkingen met de standaard kleinste kwadraten regressietechniek om zo een schatting en een standaardfout voor γ te krijgen. Het vergelijken van de resulterende t -statistiek met de bijbehorende (gegeven significantieniveau, aantal waarnemingen en gebruikte model) kritische waarden uit de Dickey-Fuller tabellen maakt het mogelijk correct te bepalen of de nulhypothese $\gamma=0$ geaccepteerd of verworpen moet worden.

⁹ Zie voor een uitputtende beschrijving Fuller (1996) hoofdstuk 10

¹⁰ In de vergelijkingen geldt nu:

$$\gamma = - \left(1 - \sum_{i=1}^p a_i \right)$$

$$b_i = \sum_{j=1}^p a_j$$

Merk met Enders(1995) pagina 225 op dat als $\sum a_i=1$, $\gamma=0$ het systeem minimaal één eenheidswortel kent.

Hoofdstuk 3: Spectraalanalyse

In de econometrie worden net als in het vorige hoofdstuk tijdreeksen veelal geanalyseerd in een tijdsdomein. Er wordt gekeken naar de realisatie van een variabele in de loop van een tijdperiode. Een variabele wordt dan gerepresenteerd als een gewogen som van zijn realisaties uit het verleden en van storingen uit het verleden. Uiteindelijk is een variabele hiermee indirect een functie van de tijd en deze manier van representeren wordt het tijdsdomein genoemd. Een tijdreeksvariabele kan daarentegen ook beschouwd worden als een gewogen som van periodieke golven van verschillende frequenties. In dit geval is de variabele dan een functie van de frequentie en deze manier van representeren wordt het frequentiedomein genoemd. Spectraalanalyse behelst het totaal van functies en bewerkingen van functies in het frequentiedomein en de transformatie van functies van het frequentiedomein naar het tijdsdomein en vice versa.

Gegeven de doelstelling van deze studie om de huizenprijs te voorspellen met een leidende indicatorenmodel biedt spectraalanalyse zowel de mogelijkheid om leidende indicatoren te identificeren als om een lineair leidende indicatorenmodel multivariaat te schatten. Een tijdreeksvariabele wordt geïdentificeerd als leidende indicator voor de huizenprijs als deze variabele voldoende samenhang met de huizenprijs toont en een voldoende voorlooptijd op de huizenprijs kent. Met spectraalanalyse worden maatstaven gedefinieerd, die deze samenhang en voorlooptijd kwantificeren. Het leidende indicatorenmodel is een voorspelmodel waarbij de huizenprijs een lineaire combinatie is van de met hun individuele voorlooptijd verschoven leidende indicatoren. De maximale voorspelperiode van de huizenprijs is dan gelijk aan het minimum van de individuele voorlooptijden van de leidende indicatoren. Spectraalanalyse biedt een integraal raamwerk om dit leidende indicatorenmodel multivariaat te schatten.

Dit hoofdstuk concentreert zich op de definities, opbouw en afleidingen van de spectraalanalyse. De toepassing van spectraalanalyse op de constructie van het voorspelmodel voor de huizenprijs komt in het volgende hoofdstuk aan bod. In paragraaf 3.1. wordt met de Fourier transformatie het fundament gelegd voor de spectraalanalyse. Hier zal onder bepaalde voorwaarden de spectraalfunctie worden afgeleid evenals de eenduidige relatie tussen de spectraalfunctie in het frequentiedomein en de covariantiefunctie in het tijdsdomein. In paragraaf 3.2. wordt de univariate spectraalanalyse uitgewerkt op basis van stationaire tijdreeksen. Stationariteit is een voorwaarde voor toepassing van de spectraalanalyse. Bovendien zal blijken dat in het tijdsdomein gedefinieerde van de tijd onafhankelijke lineaire transformaties, filters genaamd, ook een equivalent hebben in het frequentiedomein in de vorm van een zogenaamde transferfunctie. Bovendien is aan de hand van deze transferfunctie de verandering van het spectrum door toepassing van het filter relatief eenvoudig te beschrijven in het frequentiedomein. Paragraaf 3.3. is de belangrijkste paragraaf van dit hoofdstuk. Hier worden de relevante concepten en maatstaven over de relatie tussen variabelen gedefinieerd en toegepast in voorbeelden. Deze maatstaven kunnen worden afgeleid uit een transferfunctie, die analoog aan de transferfunctie van paragraaf 3.2 alle informatie over de relatie tussen twee tijdreeksvariabelen incorporeert. Tenslotte zal in paragraaf 3.4. worden ingegaan op de meest voor de hand liggende schatter van het spectrum, die wel asymptotisch zuiver is, maar niet consistent. Door het toepassen van een kernel op deze schatter kan consistentie worden afgedwongen, wat voor kleine waarnemingsaantallen echter ten koste gaat van de zuiverheid. De bewijzen van de gepresenteerde theorema's zijn uitgewerkt in appendix A.

3.1: Fourier Analyse: een introductie

Deze paragraaf heeft als doel de lezer intuïtie te verschaffen over de wiskundige bouwstenen waarop de spectraalanalyse gestoeld is. Er wordt toegewerkt naar de zogenaamde Fourier transformatiestelling, die een eenduidige koppeling bewerkstelligt tussen de representatie van een stationaire tijdreeks in het tijdsdomein en in het frequentiedomein. In het tijdsdomein is de representatie van een stochastische tijdreeksvariabele een functie van zijn vertragingen, tijd en een gewogen som van onafhankelijke en

identieke verdeelde innovaties en eventueel enkele initiële beginwaarden. In het frequentiedomein is dezelfde stochastische tijdreeksvariabele een functie van de frequentie en opgebouwd uit een oneindige gewogen som van periodieke sinus- en cosinusfuncties. De Fourier analyse analyseert mede de voorwaarden waaronder en de wijze waarop functies beschreven kunnen worden als een som of integraal van sinus- en cosinusfuncties.

Allereerst wordt aangetoond dat een stelsel van sinus- en cosinusfuncties, het goniometrische stelsel, een orthogonale ruimte is, waarbij de Fourier-reeks een kleinste kwadraten projectie is van een functie op deze ruimte. Vervolgens wordt gekeken aan welke condities een functie moet voldoen om benaderd te kunnen worden door een Fourier-reeks.

Centraal in de Fourier analyse zijn de Fourier transformatie en de inverse Fourier transformatie, die de omzetting van een functie in zijn Fourier equivalent en vice versa beschrijven. In deze studie wordt beperkt tot de Fourier transformatie van een discrete functie, waarvan de realisaties absoluut sommeerbaar zijn. De Fourier transformatie van een specifieke discrete functie, namelijk de correlatiefunctie, wordt spectraalfunctie of spectrum genoemd. Dit hoofdstuk behandelt slechts enkele theorema's van de Fourier analyse en beoogt zeker geen uitputtend overzicht van dit vakgebied te geven.

3.1.1: Systemen van orthogonale functies

Het is in de algebra gebruikelijk een ruimte te creëren aan de hand van een verzameling vectoren, die basis genoemd wordt en waarbij alle andere vectoren worden uitgedrukt als een lineaire combinatie van elementen uit deze basis. De basisvectoren zijn vaak zodanig geconstrueerd dat ze orthogonaal zijn, wat wil zeggen dat de som van de produkten van de elementen van elk tweetal verschillende vectoren gelijk is aan nul. De Fourier-analyse beschouwt de volgende goniometrische verzameling:

$$\left\{ \cos\left(\frac{2\pi m}{n} t\right), \sin\left(\frac{2\pi m}{n} t\right) : t = 0, 1, \dots, n-1 \text{ en } m = 0, 1, \dots, L(n) \right\} \quad (3.1.1)$$

waarbij $L(n)$ het grootste grootste gehele getal $\leq n/2$.

In onderstaand theorema 3.1.1. zal bewezen worden dat bovenstaande goniometrische functies, gedefinieerd op de natuurlijke getallen $0, 1, \dots, (n-1)$, orthogonaal zijn aan de hand van een uitdrukking voor de som van de produkten van een tweetal verschillende functies. Het bewijs van de stelling staat in de appendix A.3.1.1.

Theorema 3.1.1: Een orthogonaal goniometrisch stelsel

Onder de aanname dat m en r element zijn van de verzameling $\{0, 1, \dots, L(n)\}$ geldt :

$$\sum_{t=0}^{n-1} \cos \frac{2\pi m}{n} t \cos \frac{2\pi r}{n} t = \begin{cases} n, & m = r = 0 \text{ of } n/2 \\ n/2, & m = r \neq 0 \text{ of } n/2 \\ 0, & m \neq r \end{cases}$$

$$\sum_{t=0}^{n-1} \sin \frac{2\pi m}{n} t \cos \frac{2\pi r}{n} t = 0 \quad \forall_{m,r}$$

$$\sum_{t=0}^{n-1} \sin \frac{2\pi m}{n} t \sin \frac{2\pi r}{n} t = \begin{cases} n/2, & m = r \neq 0 \text{ of } n/2 \\ 0, & m \neq r \end{cases}$$

□

Uit Theorema 3.1.1. is af te leiden dat de som van de produkten van een tweetal verschillende vectoren ($m \neq r$) gelijk is aan nul en daarmee is de orthogonaliteit van de n functies gedemonstreerd. Met het orthogonale stelsel kan elke functie $f(t)$ gedefinieerd op n gehele getallen, beschreven worden door:

$$f(t) = \sum_{m=0}^{L(n)} (a_m \cos(\omega_m t) + b_m \sin(\omega_m t)), \quad t = 0, 1, \dots, (n-1) \quad (3.1.2)$$

$$\text{waarbij: } \omega_m = \frac{2\pi m}{n} \quad m = 0, 1, \dots, L(n)$$

$$a_m = \begin{cases} \frac{2 \sum_{t=0}^{N-1} f(t) \cos(\omega_m t)}{n}, & m = 1, 2, \dots, L(n-1) \\ \frac{\sum_{t=0}^{N-1} f(t) \cos(\omega_m t)}{n}, & m = 0 \text{ en } m = N/2 \text{ (if } n \text{ even)} \end{cases} \quad (3.1.3)$$

$$b_m = \frac{2 \sum_{t=0}^{N-1} f(t) \sin(\omega_m t)}{n}, \quad m = 1, 2, \dots, L(n-1) \quad (3.1.4)$$

a_m en b_m worden Fourier coëfficiënten genoemd. Hoewel voor de bewijsvoering niet per se noodzakelijk is het voor de interpretatie duidelijker de Fouriercoëfficiënten te berekenen door een minimalisatie van het kwadratische verschil, namelijk door:

$$\min_{a_m, b_m} \sum_{t=0}^{n-1} \left\{ f(t) - \sum_{m=0}^{L(n)} a_m \cos(\omega_m t) + b_m \sin(\omega_m t) \right\}^2.$$

Eerste orde voorwaarden :

$$\sum_{t=0}^{n-1} \left\{ f(t) - \sum_{m=0}^{L(n)} a_m \cos(\omega_m t) + b_m \sin(\omega_m t) \right\} \cos(\omega_j t) = 0, \quad j = 0, 1, 2, \dots, L(n).$$

$$\sum_{t=0}^{n-1} \left\{ f(t) - \sum_{m=0}^{L(n)} a_m \cos(\omega_m t) + b_m \sin(\omega_m t) \right\} \sin(\omega_j t) = 0, \quad j = 0, 1, 2, \dots, L(n-1).$$

Met theorema 3.1.1 is de produktsom van sinus en cosinus gelijk aan nul en dus :

$$\sum_{t=0}^{n-1} f(t) \cos(\omega_m t) = a_m \sum_{t=0}^{n-1} \cos^2(\omega_m t),$$

$$\sum_{t=0}^{n-1} f(t) \sin(\omega_m t) = b_m \sum_{t=0}^{n-1} \sin^2(\omega_m t).$$

Wederom met theorema 3.1.1. volgen hieruit de coëfficiënten (3.1.3) en (3.1.4). Uit deze benadering blijkt dat deze coëfficiënten de regressiecoëfficiënten zijn, die verkregen worden door $f(t)$ ¹¹ te regresseren op $\cos(\omega_m t)$ en $\sin(\omega_m t)$. De in theorema 3.1.1. gestelde orthogonaliteit van de regressoren impliceert dat de meervoudige regressiecoëfficiënten gelijk zijn aan de enkelvoudige. Een enkelvoudig regressiecoëfficiënt is de coëfficiënt die wordt verkregen uit een regressiemodel met slechts een enkele regressor. Een meervoudige regressiecoëfficiënt wordt verkregen uit een regressiemodel met meerdere regressoren, waarbij de regressiecoëfficiënt de relatie weergeeft tussen de regressand en de regressor, gegeven alle overige regressoren. Een meervoudig regressiecoëfficiënt corrigeert voor beïnvloeding door de overige regressoren van het regressiemodel. Echter in het geval dat de regressoren orthogonaal zijn, zoals in het goniometrische stelsel, is er van onderlinge beïnvloeding van de regressoren geen sprake en is de enkelvoudige regressiecoëfficiënt gelijk aan de meervoudige.

¹¹ In deze studie wordt gewerkt met (stationaire) differentiereksen, waardoor niet de golfbewegingen van de oorspronkelijke tijdreeksen $f(t)$ geanalyseerd worden, maar de golfbewegingen van de eerste-orde verschilreeksen. Het continue equivalent van de verschilreeks is de afgeleide $f_i(t)$ en aangetoond wordt dat de golfbewegingen van de afgeleide hetzelfde zijn in termen van frequentie en golflengte als van de oorspronkelijke tijdreeks $f(t)$. De afgeleide luidt als volgt:

$$f_i(t) = \omega_m \sum_{m=0}^{L(n)} (b_m \cos(\omega_m t) - a_m \sin(\omega_m t))$$

Hieruit volgt dat de amplitude van de golven van een differentiereeks anders zijn dan van de oorspronkelijke reeks, maar niet de frequentie en golflengte. Het minteken voor de sinusfunctie veroorzaakt wel een verschuiving van de golven met een halve periode. Zie hiervoor mede voorbeeld 4.3.5. over het differentiefilter.

Nu is aangetoond dat een op een eindig aantal punten gedefinieerde functie goniometrisch gerepresenteerd kan worden is de volgende stap het onderzoeken van een dergelijke representatie voor functies die gedefinieerd zijn op een eindig interval van de reële lijn. Hierbij wordt het in de goniometrie meest gangbare interval, $[-\pi, \pi]$, gebruikt. De beperking tot dit interval is zonder verlies aan algemeenheid, omdat het gedrag van een functie op een groter interval geheel is af te leiden uit het gedrag op dit meest gangbare interval. Geheel in lijn met de bovenstaande discrete analyse wordt de definitie van orthogonaliteit van een oneindig systeem van continue functies gegeven en bewezen dat het goniometrische stelsel aan deze definitie voldoet.

Definitie 3.1.2: Orthogonaliteit van continue functies

Een oneindig systeem van kwadratisch integreerbare functies $\{\varphi_j\}_{j=-\infty}^{\infty}$ is orthogonaal op $[a, b]$ als

$$\int_a^b \varphi_j(x) \varphi_m(x) dx = 0, \quad j \neq m, \quad j, m = 0, 1, \dots$$

en $\int_a^b \varphi_j^2(x) dx \neq 0, \quad j = 0, 1, \dots$

□

De orthogonaliteit van het goniometrische stelsel wordt ook in het continue geval bewezen door bovenstaande definitie toe te passen. Het bewijs staat wederom in de appendix A.3.1.3.

Theorema 3.1.3: Orthogonaliteit van het oneindige goniometrische stelsel

Gegeven dat m en j niet - negatieve gehele getallen zijn, dan :

$$\int_{-\pi}^{\pi} \sin mx \cos jx dx = 0 \quad \forall m, j;$$

$$\int_{-\pi}^{\pi} \sin mx \sin jx dx = \begin{cases} 0, & m \neq j, \\ \pi, & m = j \neq 0, \\ 0, & m = j = 0; \end{cases}$$

$$\int_{-\pi}^{\pi} \cos mx \cos jx dx = \begin{cases} 0, & m \neq j, \\ \pi, & m = j \neq 0, \\ 2, & m = j = 0. \end{cases}$$

□

Er is, analoog aan het discrete geval, een integreerbare functie $f(x)$ gedefinieerd op $[-\pi, \pi]$, waarvoor een goniometrisch representatie voor wordt gezocht. Nu $f(x)$ niet meer gedefinieerd is op N discrete punten, maar op het interval $[-\pi, \pi]$ geldt tevens voor de term ω_m in (3.1.2): $\omega_k = 2\pi k / 2\pi = k$. De afleiding van de goniometrische benadering van een continue functie $f(x)$ voor een eindige N gaat als volgt:

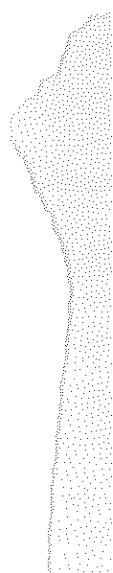
$$\min_{a_k, b_k} \int_{-\pi}^{\pi} \left[f(x) - \sum_{k=0}^n (a_k \cos kx + b_k \sin kx) \right]^2 dx.$$

Eerste orde condities voor a_k :

$$2 \int_{-\pi}^{\pi} \left[\left(f(x) - \sum_{k=0}^n (a_k \cos kx + b_k \sin kx) \right) \cos kx \right] dx = 0 \Leftrightarrow$$

$$\int_{-\pi}^{\pi} f(x) \cos kx dx = \int_{-\pi}^{\pi} a_k \cos^2 kx dx = a_k \pi \Leftrightarrow a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx. \tag{3.1.5}$$

$$\text{Evenzo volgt } b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx. \tag{3.1.6}$$



Deze paragraaf heeft het bestaan van zowel een eindig als een oneindig orthogonaal goniometrisch stelsel aangetoond. Tevens zijn voor functies, zowel gedefinieerd op een eindig aantal punten als op een interval van de reële lijn, de coëfficiënten bepaald van een kleinste kwadraten projectie op het goniometrische stelsel. In hoeverre de som van de produkten van de berekende coëfficiënten en de bijbehorende goniometrische functies een Fourier reeks is wordt in de volgende paragraaf bepaald.

3.1.2: Fourier-reeksen

De lezer is wellicht bekend met het Weierstrass approximatie theorema dat stelt dat elke continue functie op een compact interval benaderd kan worden door een polynoom. Soortgelijke theorema's bestaan tevens met betrekking tot het goniometrische stelsel en worden uitgebreid behandeld in Tolstov (1962). In deze paragraaf worden summier enkele resultaten hiervan weergegeven met als uiteindelijke doel de formulering van de Fourier transformatiestelling toegepast op de correlatiefunctie. In de vorige subparagraaf is aangetoond dat een continue functie benaderd kan worden door een eindig goniometrisch polynoom en in deze subparagraaf wordt de convergentie van een oneindig goniometrisch polynoom bekeken. Ter aanvang enige definities:

$$\text{De som } S_n(x) = \sum_{k=0}^n (a_k \cos kx + b_k \sin kx) \quad (3.1.7)$$

wordt een goniometrisch polynoom van de orde n genoemd. Als n toeneemt zonder grens ontstaat het oneindige goniometrische polynoom:

$$S(x) = \sum_{k=0}^{\infty} (a_k \cos kx + b_k \sin kx) \quad (3.1.8)$$

Definitie 3.1.4: Fourierreeks

Een oneindig goniometrisch polynoom is een Fourierreeks als de reeks $S_n(x)$ uniform convergeert naar de bijbehorende functie f . De bijbehorende Fourier-coëfficiënten a_k en b_k zijn gedefinieerd in (3.1.5) en (3.1.6). □

Met definitie A.3.1.9 van uniforme convergentie in appendix A is eenvoudig een voorwaarde af te leiden waarbij een goniometrische reeks een Fourier-reeks is:

Theorema 3.1.5: convergentievoorwaarde Fourier-reeks

Beschouw de goniometrische reeks (3.1.8):

$$\sum_{k=0}^{\infty} (a_k \cos kx + b_k \sin kx),$$

die niet a priori verondersteld is de Fourier-reeks te zijn van een bepaalde functie. Als voor de bijbehorende reeks coëfficiënten geldt:

$$\sum_{k=0}^{\infty} |a_k| + |b_k| < \infty$$

dan convergeert (3.1.7) voor $k \rightarrow \infty$ absoluut en continu en heeft daarom een continue som waarvan het de Fourier-reeks is.

Bewijs:

$$|a_k \cos kx + b_k \sin kx| \leq |a_k \cos kx| + |b_k \sin kx| \leq |a_k| + |b_k| < \infty.$$

De termen van de goniometrische reeks (3.1.8) zijn kleiner dan de termen van de convergente reeks, waarmee het resultaat volgt met Weierstrass M-test (lemma A.3.1.10). □

In de vorige paragraaf zijn de kleinste kwadraten-coëfficiënten (3.1.3), (3.1.4) voor discrete functies en (3.1.5), (3.1.6) voor continue functies afgeleid van een projectie van een functie op de goniometrische ruimte. Met theorema 3.1.5. is bekend dat als de oneindige absolute som van deze coëfficiënten convergeert, d.w.z. strikt kleiner dan oneindig is, de goniometrische reeks (3.1.8) een Fourier-reeks is.

Een volgende stap is dan ook het bepalen van de klasse van functies, waarvan de bijbehorende kleinste kwadraten-coëfficiënten voldoen aan de convergentie-eis van theorema 3.1.5.

Een startpunt om convergentie van een reeks te beschouwen is een bekend theorema genaamd Bessel's ongelijkheid. Merk op dat voor (3.1.7) geldt:

$$S_n^2(x) = \left\{ \sum_{k=0}^n (a_k \cos kx + b_k \sin kx) \right\}^2 = \pi \sum_{k=0}^n (a_k^2 + b_k^2)$$

Deze som is compleet analoog aan de kwadratensom uit een regressie op n orthogonale variabelen, zoals bekend uit de eindige regressie-theorie. Het verklarend vermogen van een regressiemodel, R^2 , wordt berekend als een verhouding van de som van de gekwadrateerde verwachte modelwaarden voor de exogene variabele en de som van de gekwadrateerde werkelijke waarden. Bessel's ongelijkheid stelt een soortgelijke kwaliteitsmaat voor de kwaliteit van de projectie van een functie op het goniometrische stelsel.

Theorema 3.1.6: Bessel's ongelijkheid

Laat a_k , b_k en $S_n(x)$ gedefinieerd zijn door (3.1.5), (3.1.6) en (3.1.7) respectievelijk. Als $f(x)$ gedefinieerd op $[-\pi, \pi]$ kwadratisch integreerbaar is, geldt:

$$\frac{1}{\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx \geq \sum_{k=0}^{\infty} (a_k^2 + b_k^2) \tag{3.1.9}$$

Bewijs :

$$0 \leq \int_{-\pi}^{\pi} (f(x) - S_n(x))^2 dx = \int_{-\pi}^{\pi} f^2(x) dx - 2 \int_{-\pi}^{\pi} f(x) S_n(x) dx + \int_{-\pi}^{\pi} S_n^2(x) dx =$$

, door definities van a_k , b_k en orthogonaliteit van sinus en cosinus functies,

$$= \int_{-\pi}^{\pi} f^2(x) dx - 2\pi \left\{ \sum_{k=0}^n (a_k^2 + b_k^2) \right\} + \pi \left\{ \sum_{k=0}^n (a_k^2 + b_k^2) \right\} \Leftrightarrow \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(x) dx \geq \sum_{k=0}^n (a_k^2 + b_k^2)$$

waarbij n eindig is.

□

Bessel's ongelijkheid genereert een bovengrens aan de kwadratische som van (3.1.7) en stelt daarmee feitelijk dat de kwadratensom die verkregen is uit de projectie nooit groter kan zijn dan de integraal van de kwadratische functie. De mate van ongelijkheid is daarbij negatief evenredig met het verklarend vermogen van deze projectie. Zou (3.1.9) als ratio gedefinieerd worden met in de teller de kwadratensom en in de noemer de integraal, dan is deze ratio feitelijk het verklarend vermogen van de projectie en compleet analoog aan de bekende R^2 -maatstaf uit de regressie-analyse.

In de Fourier theorie bestaat er een stelling, Parseval's theorema, die stelt dat Bessel's ongelijkheid (3.1.9) een gelijkheid is als er sprake is van een kleinste kwadraten projectie op het oneindige goniometrische stelsel. De bewijsvoering stoelt op de compleetheid van het onderliggende oneindige goniometrische stelsel en daarmee samenhangend de uniciteit van de Fourier-coëfficiënten. Een systeem van functies, zoals het goniometrische systeem, is compleet als er geen functie buiten het systeem bestaat zodanig dat de integraal van deze functie ongelijk nul is en de integraal van het product van deze functie met een functie uit het systeem wel gelijk nul is. Deze overweging van compleetheid is de reden waarom een Fourier-reeks een oneindig goniometrisch polynoom is. Voor gedetailleerde notities over Parseval's theorema wordt verwezen naar Tolstov (1962).

Met behulp van Bessel's ongelijkheid is het mogelijk een klasse van functies te identificeren, waarvan een Fourier-reeks bestaat. Met een technische handigheid wordt de convergentie van de coëfficiënten afgedwongen door de convergentie van de kwadratensom met Bessel's ongelijkheid. Vanwege het technische karakter is het bewijs niet opgenomen in de hoofdtekst en wordt verwezen naar de appendix A.3.1.7.

Theorema 3.1.7: Klasse van functies waarvan Fourier-reeks bestaat

Stel $f(x)$ een continue periodieke functie van periode 2π met een afgeleide $f'(x)$, die kwadratisch integreerbaar is. Dan convergeert de goniometrische reeks (3.1.8) absoluut en uniform naar $f(x)$ en is daarmee de Fourier-reeks.

□

Indien de Fourier-reeks bestaat is er feitelijk sprake van een unieke equivalentie tussen een functie van $x \in \mathbb{R}$ en een functie van $\omega \in [-\pi, \pi]$. Deze equivalentie wordt beschreven door een transformatie en bestaat voor allerlei klassen van functies. Voor de volledigheid wordt vermeld dat Fourier-reeksen bepaald kunnen worden voor een veel grotere klasse van functies dan de verzameling die theorema 3.1.7 dekt en niet per se continu of kwadratisch integreerbaar hoeven te zijn. Zo bestaan Fourier-reeksen voor type functies als discrete, continue, continue stuksgewijs gladde, continue periodieke stuksgewijs gladde, etc. Een stuksgewijs gladde functie is een functie, waarbij zowel de functie als de afgeleide van de functie continu zijn en/of slechts een aftelbaar aantal discontiuiteiten kennen.

In deze studie wordt met betrekking tot de spectraalanalyse beperkt tot de Fourier transformatie van een bepaalde discrete functie, waarvan de waarden absoluut sommeerbaar zijn, namelijk de correlatiefunctie. Theorema 3.1.8. stelt de geldigheid van deze discrete transformatie en een instructieve afleiding van de getransformeerde functie volgt aan het eind van de subparagraaf. Merk op dat $f(x)$ een even functie is als geldt: $f(x) = f(-x) \quad \forall x$.

Theorema 3.1.8: Fourier transformatie theorema

Neem aan dat de autocorrelatie functie van een stationair tijdreeks absoluut sommeerbaar is,

dat wil zeggen dat $\{\rho(j)\}_{j=-\infty}^{\infty}$ een absoluut sommeerbare reeks van autocorrelaties is.

Dan bestaat er een continue functie $f(\omega)$ zodanig dat :

$$i) \quad \rho(h) = \int_{-\pi}^{\pi} f(\omega) \cos(\omega h) d\omega = \int_{-\pi}^{\pi} f(\omega) e^{i\omega h} d\omega,$$

$$ii) \quad f(\omega) \geq 0,$$

$$iii) \quad \int_{-\pi}^{\pi} f(\omega) d\omega = 1,$$

$$iv) \quad f(\omega) \text{ is een even functie.}$$

□

In het Fourier transformatietheorema wordt de voorwaarde opgelegd dat de correlaties van een stationaire tijdreeks absoluut sommeerbaar zijn. In het volgende hoofdstuk worden stationaire tijdreeksen uitgebreid gekarakteriseerd aan de hand van correlaties. Opgemerkt wordt dat aan de eis dat een tijdreeks stationair en de autocorrelaties absoluut sommeerbaar moeten zijn om toepassing van spectraalanalyse mogelijk te maken in de praktijk echter niet a priori voldaan is. De tijdreeksen moeten hiervoor veelal eerst van hun trendcomponent ontdaan worden.

$f(\omega)$ wordt de Fourier transformatie genoemd. Ander benamingen zijn spectraaldichtheid, spectrum, spectrogram of spectraalfunctie. Bewering (i) in de transformatiestelling wordt de inverse transformatie genoemd of ook wel de karakteristieke functie. Uit de Fourier-transformatiestelling blijkt dat elk covariantie-stationaire tijdreeks met absoluut sommeerbare correlaties zowel een representatie in het tijdsdomein heeft, $f(\omega)$, als in het frequentiedomein, $\rho(h)$. Elke eigenschap van de data komt in beide beschrijvingen tot uitdrukking. Een van de belangrijkste doelstellingen van deze studie is het bepalen in hoeverre reeksen ten opzichte van elkaar voor- of achterlopen en later al blijken dat het frequentiedomein hiervoor een geijkte methode is.

Ter afsluiting van deze paragraaf wordt er als samenvatting een alternatieve en eenvoudige afleiding gegeven van de spectraalfunctie aan de hand van de geïntroduceerde begrippen.

$$\begin{aligned}
f(x) &= \sum_{k=0}^{\infty} [a_k \cos(kx) + b_k \sin(kx)] \stackrel{(3.1.5)}{=} \sum_{k=0}^{\infty} \left[\frac{1}{\pi} \int_{-\pi}^{\pi} f(\omega) \cos(k\omega) \cos(kx) d\omega + \frac{1}{\pi} \int_{-\pi}^{\pi} f(\omega) \sin(k\omega) \sin(kx) d\omega \right] \\
&= \sum_{k=0}^{\infty} \left[\frac{1}{\pi} \int_{-\pi}^{\pi} f(\omega) \cos(k(\omega - x)) d\omega \right] \stackrel{(3.1.8.i)}{=} \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{-ikx} \int_{-\pi}^{\pi} f(\omega) e^{ik\omega} d\omega \stackrel{(3.1.8.i)}{=} \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{-ikx} \rho(k)
\end{aligned}$$

3.2: Univariate spectraalanalyse

Er bestaat een eenduidige relatie tussen de representatie van een reeks in het tijdsdomein en de representatie in het frequentiedomein. In het tijdsdomein wordt een variabele gerepresenteerd als een gewogen som van zijn realisaties uit het verleden en van storingen uit het verleden. Een variabele is in het frequentiedomein een som van periodieke golven van verschillende frequenties. De uitsplitsing van een variabele naar de frequentie bepaalt de relatieve importantie van de afzonderlijke golven, die het gedrag van de variabele beschrijven en wordt het spectrum genoemd. Zo is het mogelijk om te bepalen of lange termijn golven in een tijdreeks relatief dominant zijn ten opzichte van korte termijn golven. Als bijvoorbeeld het spectrum in een bepaald frequentiegebied piekt betekent dit dat de bij deze frequenties behorende golven dominant zijn in de beweeglijkheid van de tijdreeks.

Nadat de belangrijke Wold-decompositie voor stationaire tijdreeksen is behandeld wordt er een nadere interpretatie van het spectrum gegeven, waarbij de spectraalfunctie feitelijk een toepassing is van een bepaalde statistische samenvattende functie, de zogenaamde autocovariantiegenererende functie, met een complex argument. Met deze kennis wordt de analytische spectraalfunctie van een ARMA-model afgeleid en grafisch gepresenteerd.

Vervolgens wordt aangetoond dat gepredetermineerde operaties in het tijdsdomein een uniek gepredetermineerd equivalent in het frequentiedomein hebben. Ter illustratie wordt de operatie van een lineair filter in het tijdsdomein middels een Fourier-transformatie omgezet naar een unieke bijbehorende operatie in het frequentiedomein. Deze omzetting van het filter levert een zogenaamde transferfunctie op, die alle veranderingen in het spectrum veroorzaakt door toepassing van het filter beschrijft. De transferfunctie van een bekend filter beschrijft de door het filter bepaalde deterministische relatie in het frequentiedomein tussen de ongefilterde en de gefilterde reeksen en daarmee wordt een eerste stap gezet in de multivariate spectraalanalyse van de volgende paragraaf.

3.2.1: Stationaire processen

De klasse van alle stochastische processen is veel te omvangrijk voor het ontwikkelen van algemeen toepasbare analysemethoden die geschikt zijn voor alle type processen. Het gevolg is dat de ontwikkeling van het vakgebied voornamelijk verlopen is langs studielijnen van speciale type processen, zoals Markov-processen, geboorte- en sterfteprocessen, vernieuwingsprocessen, boomprocessen, etc. De spectraalanalyse beschouwt de klasse van stationaire processen, die eerder in hoofdstuk 2 gedefinieerd zijn door het verschijnsel dat hun belangrijkste statistische eigenschappen niet veranderen over de tijd. Uit verdere analyse van stationaire tijdreeksen volgen een aantal eigenschappen, die verderop in de studie waardevol zullen zijn. Ten eerste geldt de volgende eigenschap voor zowel de covariantie als voor de correlatie:

Theorema 3.2.1:

$$\gamma(s) = \gamma(-s)$$

Bewijs:

$$\begin{aligned}
\gamma(s) &= E(x_t - \mu)(x_{t-s} - \mu) = E(x_{t-j} - \mu)(x_{t-j-s} - \mu) = \\
&= E(x_{t-(-s)} - \mu)(x_t - \mu) = E(x_t - \mu)(x_{t-(-s)} - \mu) = \gamma(-s)
\end{aligned}$$

□

Aan het eind van dit hoofdstuk worden er spectraalschatters gedefinieerd en worden er afleidingen gepresenteerd die aantonen dat deze schatters consistent zijn. Deze afleidingen zijn gebaseerd op de aanname dat een tijdreeksvariabele beschreven kan worden als een oneindige gewogen som van normaal verdeelde stochasten, ook wel innovaties genaamd. Dat met deze aanname een grote klasse van tijdreeksen gedekt wordt is het resultaat van een fundamenteel theorema uit de tijdreeksliteratuur, namelijk de Wold-decompositie. Met Priestley (1981) pagina 756 luidt deze als volgt:

Theorema 3.2.2: Wold-decompositie (zonder bewijs):

Neem aan dat Y_t een covariantie stationaire tijdreeks is gedefinieerd op $T=\{0, \pm 1, \pm 2, \dots\}$.

Dan kan Y_t uitgedrukt worden in de vorm:

$Y_t = U_t + V_t$, waarbij

- i) U_t en V_t ongecorreleerde processen zijn,
- ii) U_t een regulier proces is met als representatie:

$$U_t = \sum_{u=0}^{\infty} g_u e_{t-u},$$

met $g_0 = 1, \sum_{u=0}^{\infty} g_u^2 < \infty, e_t \sim (0, \sigma^2)$ o.i.v., waarbij:

$$E\{e_t V_s\} = 0 \quad \forall s, t.$$

De reeks $\{g_u\}$ en het proces $\{e_t\}$ zijn uniek bepaald.

- iii) V_t is singulier en kan dus voorspeld worden vanuit zijn verleden met een voorspelvariantie van nul.

□

Elke stationaire tijdreeks kan met de Wold-decompositie herschreven worden als een (voorspelbare) constante en een oneindige gewogen som innovaties. Gegeven deze heldere structuur waarin elke stationaire tijdreeks gegoten kan worden is het mogelijk de covariantiereeks behorende bij dit stationaire proces eenvoudig samen te vatten in een bepaalde statistische functie, de autocovariantie-genererende functie, die gedefinieerd is als volgt:

Definitie 3.2.3: Autocovariantie-genererende functie:

Laat $\{y_t\}$ een stationaire tijdreeks zijn met $\sum_{j=0}^{\infty} |\gamma(j)| < \infty$, dan is

$$g_y(z) = \sum_{j=-\infty}^{\infty} \gamma(j) z^j \tag{3.2.1}$$

de autocovariantie-genererende functie.

□

De autocovariantie-genererende functie faciliteert een eenvoudige structuur waarin de j^e covariantie de multiplicatiefactor is behorende bij de j^e machtsvariabele van z . Een en ander wordt geïllustreerd aan de hand van een MA-reeks.

Een MA(q) proces kan met vertragingbewerker L als volgt worden gerepresenteerd:

$$y_t = \mu + \sum_{j=1}^q \varphi_j e_{t-j} = \mu + (1 + \varphi_1 L + \dots + \varphi_q L^q) e_t \tag{3.2.2}$$

waarbij $e_t \sim N(0, \sigma^2)$ o.i.v..

Voor $j = 1, 2, \dots, q$ geldt:

$$\begin{aligned} \gamma_j &= E(e_t + \varphi_1 e_{t-1} + \dots + \varphi_q e_{t-q})(e_{t-j} + \varphi_1 e_{t-j-1} + \dots + \varphi_q e_{t-j-q}) = \\ &= E(\varphi_j e_{t-j}^2 + \varphi_{j+1} \varphi_1 e_{t-j-1}^2 + \varphi_{j+2} \varphi_2 e_{t-j-2}^2 + \dots + \varphi_q \varphi_{q-j} e_{t-q}^2) = \sigma^2 (\varphi_j + \varphi_{j+1} \varphi_1 + \dots + \varphi_q \varphi_{q-j}) \end{aligned}$$

Nu geldt voor de autocovariantie-genererende functie:

$$g_y(z) = \sigma^2 (1 + \varphi_1 z + \dots + \varphi_q z^q) (1 + \varphi_1 z^{-1} + \dots + \varphi_q z^{-q}) \tag{3.2.3}$$

Namelijk:

$$g_y(z) = \sigma^2 \left\{ \varphi_q z^q + (\varphi_{q-1} + \varphi_q \varphi_1) z^{q-1} + (\varphi_{q-2} + \varphi_{q-1} \varphi_1 + \varphi_q \varphi_2) z^{q-2} + \dots + (\varphi_1 + \varphi_2 \varphi_1 + \dots + \varphi_q \varphi_{q-1}) z \right. \\ \left. + (1 + \varphi_1^2 + \dots + \varphi_q^2) z^0 + (\varphi_1 + \varphi_2 \varphi_1 + \dots + \varphi_q \varphi_{q-1}) z^{-1} + \dots + \varphi_q z^{-q} \right\} = \sigma^2 \sum_{j=-q}^q \gamma(j) z^j,$$

waarbij voor $j=0,1,\dots,q$ de bijbehorende $\gamma(j)$ gelijk is aan de vermenigvuldigingsfactor van de term z^j . Deze resultaten worden gegeneraliseerd naar een $MA(\infty)$ -proces, waarbij een tijdreeksvariabele gelijk is aan een constante plus de oneindige gewogen som van normaal verdeelde innovaties. Laat y_t zo'n $MA(\infty)$ -proces zijn, namelijk:

$$y_t = \mu + \psi(L)e_t$$

met: $e_t \sim N(0, \sigma^2)$ o.i.v.,

en $\psi(L) = \psi_0 + \psi_1 L + \psi_2 L^2 + \dots$, waarbij $\sum_j |\psi_j| < \infty$.

Dan geldt met (3.2.3) en $q \rightarrow \infty$:

$$g_y(z) = \sigma^2 \psi(z) \psi(z^{-1}) \quad (3.2.4)$$

Merk op dat nu variabelen oneindig ver uit het verleden meegenomen worden de zwaardere restrictie van absoluut sommeerbare coëfficiënten wordt opgelegd. Deze restrictie impliceert echter dat de covarianties absoluut sommeerbaar zijn en dit wordt aangetoond in appendix A.3.2.7.

De autocovariantie-genererende functie is een samenvattende statistische functie, waarmee eenvoudig alle covarianties berekend kunnen. Er bestaat echter een theoretisch addertje dat niet voor elke stationaire reeks er een autocovariantie-genererende functie bestaat. De enige restrictie voor deze functie is de absolute sommeerbaarheid van de covarianties en deze eis wordt afgedwongen met de voorwaarde dat de coëfficiënten absoluut sommeerbaar zijn. De Wold-decompositie stelt dat er voor elke stationaire tijdreeks een gewogen oneindige reeks van normaal verdeelde stochasten bestaat, waarbij de gewichten kwadratisch sommeerbaar zijn. Echter, absolute sommeerbaarheid is een sterkere voorwaarde dan kwadratische sommeerbaarheid¹², waarmee niet voor alle reeksen waarbij de Wold-decompositie geldt automatisch een autocovariantie-genererende functie hoeft te bestaan. Dit verschijnsel is echter meer een theoretisch curiosum dan een praktische handicap, omdat een variabele in de praktijk zeer onwaarschijnlijk geautocorreleerd is over een oneindig lange periode.

3.2.2: Het spectrum: interpretatie en toepassing

In de vorige paragraaf over de Fourier-analyse is de spectraalfunctie afgeleid en bovendien is zojuist aangetoond dat voor bijna alle stationaire reeksen er een autocovariantie-genererende functie bestaat. Onderstaande vergelijking toont een verband tussen deze twee functies onder de voorwaarde van absolute sommeerbaarheid van de covarianties. Namelijk:

$$f_y(\omega) \stackrel{\text{Th 3.1.8}}{=} \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma(j) e^{-i\omega j} \stackrel{\text{Def 3.1.4}}{=} \frac{1}{2\pi} g_y(e^{-i\omega}) \quad (3.2.5)$$

Bovendien:

$$f_y(\omega) = \frac{1}{2\pi} g_y(e^{-i\omega}) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma(j) e^{-i\omega j} = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma(j) (\cos(\omega j) - i \sin(\omega j)) = \\ \frac{1}{2\pi} \left\{ \sum_{j=-\infty}^{\infty} \gamma(j) (\cos(\omega j) + \cos(\omega j) - i \sin(\omega j) - i \sin(-\omega j)) \right\} = \frac{1}{2\pi} \left\{ \gamma(0) + 2 \sum_{j=1}^{\infty} \gamma(j) \cos(\omega j) \right\}. \quad (3.2.6)$$

¹² bijv. de reeks $\{1/x\}$ gedefinieerd op de gehele getallen is wel kwadratisch sommeerbaar, maar is niet absoluut sommeerbaar.

De spectraalfunctie is feitelijk gelijk aan de autocovariantie-genererende functie met een complex argument gedeeld door 2π , maar met (3.2.6) blijkt deze functie toch reëelwaardig te zijn. Dit betekent voor het bereik van de functie: $f(\omega) \in \mathbb{R}$. Bovendien is de spectraalfunctie als som van continue, periodieke cosinusfuncties zelf ook continu en periodiek. Periodiciteit betekent dat $f_y(\omega) = f_y(\omega + 2\pi k)$ voor elk geheel getal k , waarmee de waarde van f bekend is voor elke ω als de waarde van f bekend is voor elke $\omega \in [-\pi, \pi]$. Sinds f bovendien een even functie is, $f(\omega) = f(-\omega)$, is zijn gedrag zelfs volledig bepaald door zijn gedrag op het interval $[0, \pi]$.

In voorbeeld 3.1. is een AR(2)-model gesimuleerd, met 350 realisaties, waarna de spectraalfunctie van deze tijdreeks is geschat met het MATLAB-programma in appendix C. Bovendien is van het AR(2)-model met behulp van (3.2.4) een analytische spectraalfunctie berekend. Beide functies zijn weergegeven in figuur 3.2.

Voorbeeld 3.1: Het spectrum: frequentie en golflengte

Beschouw het volgende AR(2)-model:

$$y_t = -0.3y_{t-1} - 0.6y_{t-2} + e_t, \quad \text{waarbij } e_t \sim N(0,1)$$

\Leftrightarrow

$$(1 + 0.3L + 0.6L^2)y_t = e_t.$$

Dan geldt¹³:

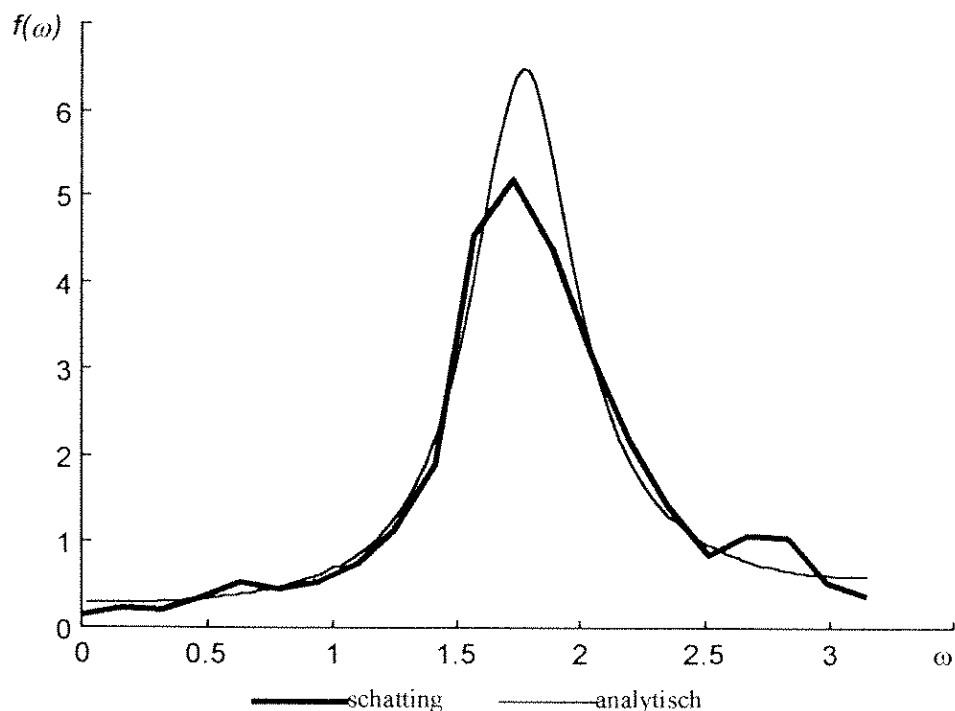
$$y_t = \psi(L)e_t, \quad \text{met}$$

$$\psi(L) = \frac{1}{1 + 0.3L + 0.6L^2}.$$

$$f_y(\omega) = \frac{1}{2\pi} g_y(e^{-i\omega}) = \frac{\sigma^2}{2\pi} \psi(e^{-i\omega})\psi(e^{i\omega}) = \frac{1}{2\pi(1 + 0.3e^{-i\omega} + 0.6e^{-2i\omega})(1 + 0.3e^{i\omega} + 0.6e^{2i\omega})}.$$

□

Figuur 3.2: De spectraalfunctie, schatting en analytisch



¹³ de karakteristieke wortels van de vertragingopolynoom liggen buiten de eenheidskring, waarmee voldaan is aan de stabiliteitsconditie.

Met de Fourier transformatiestelling, theorema 3.1.8, is bekend dat de totale oppervlakte onder de (analytische) grafiek gelijk is aan de variantie van de stationaire tijdreeks. Bovendien is uit de Fourier-analyse bekend dat een tijdreeks is opgebouwd uit golven met verschillende frequenties. Uit figuur 3.2 blijkt dat grofweg de helft van de variantie wordt veroorzaakt door golven met frequenties in het gebied [1.5, 2]. De berekening van de bij deze frequenties behorende golflengtes wordt inzichtelijk gemaakt door onderstaande redenering. De golven behorende bij de termen $\cos(\omega t)$ en $\sin(\omega t)$ kennen elk een periode van $(2\pi/\omega)$ en daarmee doorloopt elke term in één tijdseenheid $(\omega/2\pi)$ maal een complete golfcyclus.

De berekening van de golflengtes in voorbeeld 3.1. geschiedt door toepassing van bovenstaande concepten op cosinusfunctie en de bijbehorende waarde voor ω met (3.1.2) van $\omega_m = 2\pi * m/n$. De cosinusfunctie doorloopt dan in een enkele tijdseenheid $\omega_m/2\pi = m/n$ volledige cycli en met t gemeten in maanden heeft de term $\cos(\omega_m t)$ een omloopsnelheid van (m/n) cycli per maand. De golflengte l is dan de hoeveelheid maanden waarin een enkele cyclus doorlopen wordt en is gedefinieerd als de reciproque van de omloopsnelheid gelijk aan (n/m) maanden per cyclus. De golflengtes behorende bij $\omega \in [1.5, 2]$ zijn dus te berekenen met $(l = n/m = 2\pi/\omega_m)$ en dit genereert een resultaat van $l \in [3.14, 4.2]$ in maanden.

De interpretatie van het spectrum als een uitsplitsing van de beweeglijkheid van een tijdreeks naar de opbouwende karakteristieke golven wordt geïllustreerd in voorbeeld 3.3. Om de decompositie van de variantie van een tijdreeks in partities behorende bij de afzonderlijke frequenties te illustreren wordt er in het voorbeeld een discreet model beschouwd, waarvan de coëfficiënten behorende bij de goniometrische functies onafhankelijk en identiek verdeelde (o.i.v.) stochasten zijn.

Voorbeeld 3.3: Variantie decompositie

Beschouw de volgende tijdreeks :

$$y_t = \sum_{j=1}^n \{a_j \cos(\omega_j t) + b_j \sin(\omega_j t)\}$$

waarbij de reeksen $\{a_j\}_{j=1}^n$ en $\{b_j\}_{j=1}^n$ stochastische reeksen zijn met eigenschappen :

$$\begin{aligned} E(a_j^2) &= E(b_j^2) = \sigma_j^2, \quad j = 1, \dots, N \\ E(a_j a_k) &= E(b_j b_k) = 0, \quad j \neq k \\ E(a_j b_k) &= 0 \quad \forall j, k. \end{aligned}$$

De variantie van y_t is dan :

$$E(y_t^2) = \sum_{j=1}^n [E(a_j^2) \cos^2(\omega_j t) + E(b_j^2) \sin^2(\omega_j t)] = \sum_{j=1}^n \sigma_j^2 [\cos^2(\omega_j t) + \sin^2(\omega_j t)] = \sum_{j=1}^n \sigma_j^2,$$

en verder :

$$\begin{aligned} \gamma_y(k) &= E(y_t y_{t-k}) = \sum_{j=1}^n \{E(a_j^2) \cos(\omega_j t) \cos[\omega_j (t-k)] + E(b_j^2) \sin(\omega_j t) \sin[\omega_j (t-k)]\} = \\ &= \sum_{j=1}^n \sigma_j^2 \{ \cos(\omega_j t) \cos[\omega_j (t-k)] + \sin(\omega_j t) \sin[\omega_j (t-k)] \} = \sum_{j=1}^n \sigma_j^2 \cos[\omega_j t - \omega_j (t-k)] = \sum_{j=1}^n \sigma_j^2 \cos(\omega_j k). \end{aligned}$$

□

In bovenstaand voorbeeld bestaat de tijdreeksvariabele y_t uit een eindige gewogen som goniometrische functies en zijn (co)variantie kan ontrafeld worden in een som partiële (co)varianties. De uitzetting van deze partiële varianties tegen de frequentie is dan feitelijk het discrete analoog van de spectraalfunctie en de sommatie van deze partiële varianties is het discrete equivalent van theorema 3.1.8, dat stelt dat de integraal van de spectraalfunctie gelijk is aan de totale variantie.



3.2.3: Filters

Een filter is een gepredetermineerde bewerking op een tijdreeks, die zowel in het tijdsdomein als in het frequentiedomein gerepresenteerd kan worden. Vanuit een andere invalshoek is een filter de beschrijving van een gepredetermineerde relatie tussen twee reeksen, namelijk de ongefilterde en de gefilterde. In de volgende paragraaf is deze beschrijving van de relatie tussen twee reeksen niet meer gepredetermineerd en spreken we over multivariate analyse. In de loop van het hoofdstuk zal verder blijken dat de concepten, methoden en maatstaven voor de beschrijving van relaties in beide gevallen identiek zijn, maar dat de parameters gepredetermineerd zijn voor filters en geschat moeten worden voor multivariate analyse.

De invloed van een filter op een tijdreeks in het frequentiedomein wordt ontleed in een effect dat de uitslag van het spectrum veranderd en in een effect dat het spectrum verschuift. Nadat een filter is gedefinieerd worden beide effecten middels een poolcoördinaten-representatie geïsoleerd.

Vervolgens wordt de representatie van een algemeen lineair filter zowel in het tijdsdomein als in het frequentiedomein uiteengezet en worden er enkele bekende, veel gebruikte toepassingen van dit filter weergegeven.

In deze studie worden filters gedefinieerd als tijdsinvariante lineaire transformaties. Hierbij transformeert een filter F een input-tijdreeks $\{x(t)\}$, de input, in een output-tijdreeks $\{y(t)\}$, dat wil zeggen: $\{y(t)\}=F\{x(t)\}$ en heeft volgende eigenschappen:

$$\text{lineariteit: } F\left[\sum_{j=1}^n a_j \{x_j\}\right] = \sum_{j=1}^n a_j F[\{x_j\}] \quad (3.2.7)$$

$$\text{tijds - invariantie: als } F\{x_t\} = \{y_t\}, \text{ dan } F\{x_{t+h}\} = \{y_{t+h}\} \quad \forall h \quad (3.2.8)$$

Met deze definities wordt er een relatie tussen het input spectrum en het output spectrum afgeleid zoals bijvoorbeeld volgens Koopmans (1974) en loopt als volgt:

Beschouw de situatie waarbij de input bestaat uit slechts een enkele complexe exponentiële term:

$$x_t = e^{i\omega t},$$

dan geldt voor de bijbehorende output y het volgende:

$$y_{t+h} = F\{x_{t+h}\} = F\{e^{i\omega t} e^{i\omega h}\} = e^{i\omega h} F\{e^{i\omega t}\} = e^{i\omega h} y_t.$$

Omdat deze relatie geldt voor elke t en h kan deze vergelijking beschouwd worden als functie van h .

In het bijzonder, stel $t=0$, dan geldt:

$$y_h = y_0 e^{i\omega h} \quad \forall h.$$

Omdat h een waarde is van de tijdparameter, kan worden gesteld $h=t$, waarbij:

$$y_t = F\{e^{i\omega t}\} = T(\omega) e^{i\omega t} = |T(\omega)| e^{i\varphi(\omega)} e^{i\omega t} \quad (3.2.9)$$

en $T(\omega) = y_0$. Merk echter op dat y_0 de reactie is van het filter op tijdstip 0 op een input van $e^{i\omega t}$ en daarmee afhangt van ω .

Het laatste '='-teken in (3.2.9) volgt uit toepassing van de poolcoördinaten. De invloed van het filter kan met poolcoördinaten¹⁴ ontleed worden in een vermenigvuldigingsfactor $|T(\omega)|$, wat de uitslag van de reeks veranderd en daarom amplitude-effect genoemd wordt. De tweede veranderingsfactor is de term $e^{i\varphi(\omega)}$, waarbij $\varphi(\omega) = \arg\{T(\omega)\}$, wat ervoor kan zorgen dat golven een faseverschuiving ondergaan. Een meer precieze interpretatie komt nog aan bod.

Met (3.2.9) blijkt dat de transferfunctie de invloed van het filter compleet beschrijft indien de input bestaat uit slechts een enkele complexe exponentiële term. In onderstaand theorema 3.2.5. wordt aangetoond dat ook voor een heel algemeen gedefinieerde lineaire filter een dergelijke transferfunctie bestaat. De stelling is te vinden in Fuller (1996) en het bewijs in theorema A.3.2.5.:

¹⁴ Van het complexe getal $z=x+iy$ de amplitude $|z|=(x^2+y^2)^{1/2}$ en de fase $\varphi=\arg(z)=\tan^{-1}(y/x)$. Zie voor uitgebreide behandeling appendix A.3.5.4.

Theorema 3.2.5: Spectraaltransformatie van een algemeen lineair filter

Veronderstel x_t een stationaire tijdreeks met een absoluut sommeerbare covariantie - functie en veronderstel dat de deterministische rij $\{a_j\}_{j=-\infty}^{\infty}$ absoluut sommeerbaar is.

$$\text{Beschouw het algemene lineaire filter } F(x_t) = \sum_{j=-\infty}^{\infty} a_j L^j(x_t) \tag{3.2.10}$$

Dan is de spectraaldichtheid van $y_t = F(x_t) = \sum_{j=-\infty}^{\infty} a_j x_{t-j}$ gelijk aan :

$$f_y(\omega) = (2\pi)^2 f_x(\omega) f_a(\omega) f_a^*(\omega) = |T(\omega)|^2 f_x(\omega) \tag{3.2.11}$$

waarbij $f_x(\omega)$ het spectrum is van x_t en

$$f_a(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} a_j e^{-i\omega j} \quad \text{de Fourier - transformatie van } a_j \tag{3.2.12}$$

$$f_a^*(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} a_j e^{i\omega j} \quad \text{is de complex geconjugeerde Fourier transformatie of } a_j,$$

$$T(\omega) = \sum_{j=-\infty}^{\infty} a_j L^j(e^{i\omega t}) \text{ op } t=0 \text{ de bij filter } F \text{ behorende transfer - functie.} \tag{3.2.13}$$

□

Uit theorema 3.2.5 blijkt dat alle veranderingen in het frequentiedomein die door het algemene lineaire filter worden veroorzaakt in de transferfunctie geïncorporeerd zijn en met deze complexe functie ontleedt kunnen worden in een amplitude-effect, $|T(\omega)|$, wat de uitslag van de reeks verandert en een faseverschuiving, $e^{i\varphi(\omega)}$, waarbij $\varphi = \arg\{T(\omega)\}$, wat ervoor kan zorgen dat golven een faseverschuiving ondergaan. Om beide veranderingsfactoren in het frequentiedomein van het algemene lineaire filter te illustreren volgen er een aantal voorbeelden van bekende filters. In voorbeeld 3.4 wordt aangetoond hoe een eenvoudige verschuiving van een tijdreeks tot uitdrukking komt in een fasemaatstaf.

Voorbeeld 3.4: Fase en shift

Beschouw het shiftmodel als speciaal geval van het algemene lineaire model :

$$a_j = \begin{cases} 1, & j = d \\ 0, & j \neq d \end{cases}$$

waarbij een tijdreeks d - perioden vooruit wordt geschoven. Dan geldt met :

$$\varphi(\omega) = \arg\{T(\omega)\} = \arg\{e^{-i\omega d}\} = -\omega d$$

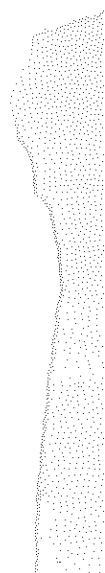
$$\text{definieer shift als : } s(\omega) := -\left\{ \frac{\varphi(\omega)}{\omega} \right\} \tag{3.2.14}$$

$$\text{met als toepassing : } s(\omega) = -\left\{ \frac{-\omega d}{\omega} \right\} = d.$$

□

Er is sprake van een shift, ofwel faseverandering als de maatstaf fase, φ , ongelijk is aan nul. Het shiftmodel van voorbeeld 3.4 kent een faseverschuiving en dit komt tot uitdrukking in een rechte lijn met het aantal perioden verschuiving, d , als negatieve hellingshoek.

Om daadwerkelijk het aantal perioden verschuiving te meten is de shift $s(\omega)$ met (3.2.14) gedefinieerd als fase gedeeld door de frequentie. De shift meet het feitelijke aantal perioden dat een tijdreeks is verschoven ten opzichte van een referentiereeks, uitgesplitst naar de frequenties. De referentiereeks is in deze paragraaf de ongefilterde reeks en in dit voorbeeld kennen alle frequenties dezelfde shift, namelijk d perioden.



De Fourier-transformatie stelt dat een tijdreeks opgebouwd kan worden uit golffuncties en de shiftmaatstaf bepaalt hoeveel tijdseenheden een bepaalde golf, waarbij de bepaling geschiedt met de frequentie, moet verschuiven ten opzichte van dezelfde bepaalde golf van de referentietijdreeks zodanig dat beide golven in dezelfde fase zitten, d.w.z. tegelijkertijd hun toppunt, dalpunt, nulpunt, etc bereiken.

Sommige filters kennen slechts een amplitude-effect en veroorzaken geen faseverandering. Dit geldt bijvoorbeeld voor het bekende filter voortschrijdend gemiddelde, zoals blijkt in voorbeeld 3.5.

Voorbeeld 3.5: Voortschrijdend gemiddelde

Beschouw (3.2.10) met de volgende karakteristieken:

$$a_j = \begin{cases} \frac{1}{m}, & |j| \leq m \\ 0, & |j| > m \end{cases}$$

Dan geldt voor de transferfunctie van dit filter:

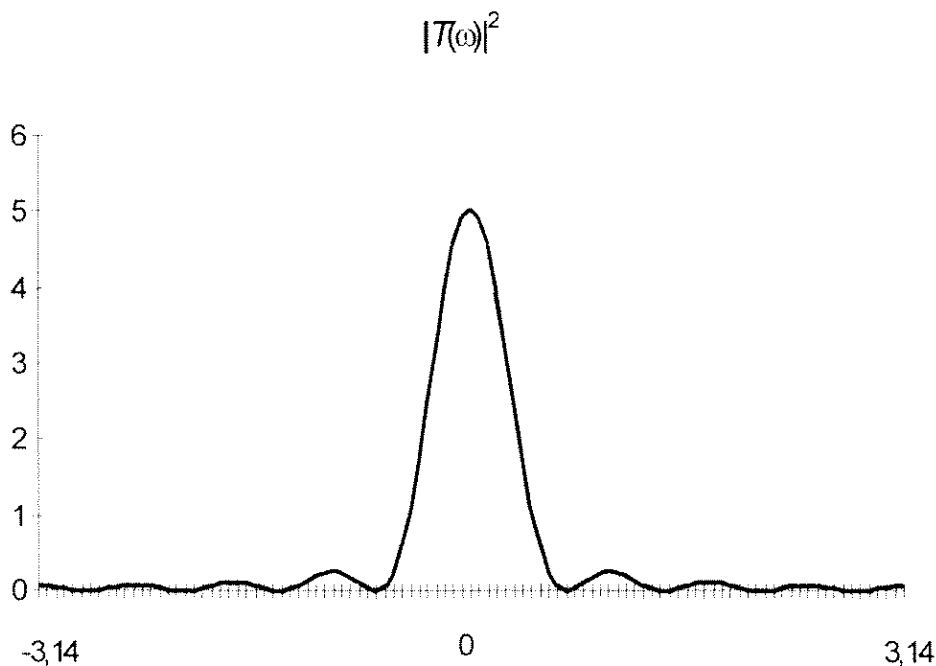
$$T(\omega) = \sum_{j=-\infty}^{\infty} a_j e^{-i\omega j} = \sum_{j=-m}^m \frac{1}{m} e^{-i\omega j} = \frac{1}{m} \sum_{j=-m}^m \cos(\omega j) \stackrel{\text{Th A.3.5.1}}{=} \frac{1}{m} \left\{ \frac{\sin\left[\left(m + \frac{1}{2}\right)\omega\right]}{\sin\left(\frac{\omega}{2}\right)} \right\},$$

waarbij $|T(\omega)|^2 = \frac{\sin^2\left[\left(m + \frac{1}{2}\right)\omega\right]}{m^2 \sin^2\left(\frac{\omega}{2}\right)}$.

□

Merk behulp van de definitie van fase in appendix A.3.5.4 op dat er geen faseverschuiving, ($\varphi(\omega)=0$), bestaat door het ontbreken van een complex getal in de transferfunctie. Omdat het filter symmetrisch is sommeren de $i \cdot \sin$ -termen weg uit de transferfunctie. In het algemeen geldt dat symmetrische filters, waarbij $a_j = a_{-j}$, geen faseverandering veroorzaken. In figuur 3.6 is de gekwadrateerde absolute transferfunctie van dit filter weergegeven, waarmee volgens (3.2.13) grafisch zichtbaar wordt op welke manier dit filter het spectrum van de ongefilterde reeks wijzigt.

Figuur 3.6: Transferfunctie behorende bij het filter van voortschrijdend gemiddelde



Uit de figuur blijkt dat deze filter een zogenaamde low-pass filter is, dat wil zeggen een filter die, in absolute termen gezien, hoge frequentiewaarden elimineert en lage frequentiewaarden ongewijzigd laat. Omdat hoge frequentiewaarden corresponderen met korte golven elimineert dit filter deze korte golven, die ook wel ruis genoemd worden. In het volgende voorbeeld wordt een differentiefilter getoond, die gekarakteriseerd kan worden als een zogenaamde high-pass filter, het tegenovergestelde van een low-pass filter. Dit filter heeft als doel om lage frequenties en daarmee corresponderende lange golven te elimineren. Dit filter is bekend, doordat het filter in het tijdsdomein wordt gebruikt om een stochastische trend uit een tijdreeks te halen.

Voorbeeld 3.7: Differentiefilter

Een differentiefilter transformeert een tijdreeks zodanig dat de output de differentiereeks, dat wil zeggen het eerste orde verschil, is van de inputreeks. Een differentiefilter wordt als toepassing van het algemene lineaire filter als volgt gedefinieerd:

$$y_t = F(x_t) = (1 - L)x_t = x_t - x_{t-1}.$$

Dit is feitelijk de algemene lineaire filter met $a_0 = 1$, $a_1 = -1$ en $a_j = 0 \quad \forall j, j \neq 0, 1$.

De transferfunctie is met appendix A.3.5.1:

$$T(\omega) = 1 - e^{-i\omega} = e^{-i\omega/2} (e^{i\omega/2} - e^{-i\omega/2}) = 2ie^{-i\omega/2} \sin(\omega/2) = 2e^{i(\pi-\omega)/2} \sin(\omega/2) = \begin{cases} 2|\sin(\omega/2)|e^{i(\pi-\omega)/2}, & \omega > 0 \\ 2|\sin(\omega/2)|e^{-i(\pi-\omega)/2}, & \omega < 0 \end{cases}$$

en daarmee geldt voor de amplitude en fase:

$$|T(\omega)| = 2|\sin(\omega/2)|,$$

$$\mathcal{A}(\omega) = \begin{cases} e^{i(\pi-\omega)/2}, & \omega > 0, \\ e^{-i(\pi-\omega)/2}, & \omega < 0. \end{cases}$$

□

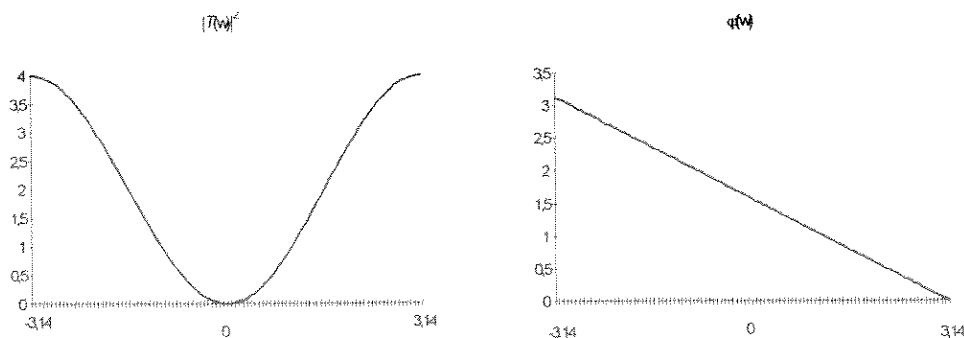
Beide veranderingsfactoren van het filter zijn grafisch uitgezet in figuur 3.8, waarbij het amplitude-effect hier weer absoluut genomen en gekwadrateerd is. Uit figuur 3.8.B blijkt er voor dit filter een fase-effect te bestaan. Eenvoudig is in te zien dat dit filter een shift van een halve periode veroorzaakt, waarmee er slechts sprake is van een marginale en verwaarloosbare verschuiving.

De in de voorbeelden beschouwde filters worden veel gebruikt en uit hun transformaties in het frequentiedomein blijkt dat het dempen van lage frequenties overeenkomt met eliminatie van trend en het dempen van hoge frequenties met eliminatie van ruis. De relatie tussen twee tijdreeksen, een gefilterde en een ongefiterde is door het a priori bekende filter deterministisch. Dit komt er concreet op neer dat de coëfficiënten van het algemene lineaire model bekend zijn.

Figuur 3.8: Amplitude en fase-effect behorende bij het differentiefilter

A) amplitude-effect

B) fase-effect



In het vervolg over multivariate spectraalanalyse zijn er twee tijdreeksen a priori gegeven en worden de best passende coëfficiënten van het algemene lineaire model berekend door een kleinste kwadraten minimalisatie. Met de berekende coëfficiënten is het weer mogelijk een transferfunctie te bepalen en daarmee amplitude- en fasemaatstaven te expliciteren als functie van de frequentie. Hiermee is het mogelijk een verdere analyse te maken door de uitsplitsing van de spectraalmaatstaven in frequenties behorende bij trendachtige golven, middellange conjunctuurgolven en korte ruisgolven.

3.3: Multivariate spectraalanalyse

De multivariate spectraalanalyse is een generalisatie van de univariate spectraalanalyse en de bijbehorende uitbreidingen van de definities en de concepten worden in subparagraaf 3.3.1 gegeven. In deze paragraaf wordt bovendien het kruislingse spectrum geïntroduceerd, die alle informatie over de relatie tussen twee tijdreeksen beschrijft in het frequentiedomein. Dit kruislingse spectrum is equivalent aan de transferfunctie uit de vorige paragraaf met het verschil dat de transferfunctie nu een schatter is. Nadat het kruislingse spectrum nader is geïnterpreteerd wordt deze in subparagraaf 3.3.2 geëxpliciteerd naar maatstaven die relatiekarakteristieken tussen twee variabelen kwantificeren. De belangrijkste maatstaven zijn de coherentie K^2 , die samenhang meet en verschuiving s , die het faseverschil tussen twee reeksen in het frequentiedomein omzet naar een, mogelijk negatieve, voorlooptijd in het tijdsdomein. Criteria voor beide maatstaven zijn uiteindelijk in het volgende hoofdstuk doorslaggevend voor de identificatie van de leidende indicatoren voor de huizenprijs.

Een uitgebreid voorbeeld in subparagraaf 3.3.3 poogt duidelijkheid te scheppen over de geïntroduceerde begrippen en concepten door twee tijdreeksen te simuleren en voor elke spectraalfunctie en –maatstaf zowel een geschatte als een analytische kwantificering te berekenen. De constructie van de spectraalschatter komt overigens pas in paragraaf 3.4 aan bod.

Tot slot worden de bivariate spectraalfuncties en –maatstaven multivariaat gegeneraliseerd, waarbij de multivariate spectraalfuncties en –maatstaven handelen over de relatie tussen twee variabelen, waarbij gecorrigeerd wordt voor beïnvloeding door de overige variabelen. Gedurende deze paragraaf zal tevens blijken dat de spectraalschatter van het multivariate lineaire model optimaal is.

3.3.1: Multivariate generalisatie van de spectraaldefinities

De n -dimensionale tijdreeks $\{y_t\}$ is gedefinieerd door $y_t = [y_{1t}, y_{2t}, \dots, y_{nt}]$, waarbij $\{y_{it} : t = 0, \pm 1, \pm 2, \dots\}$, $i = 1, 2, \dots, n$ univariate tijdreeksen zijn.

Onder de veronderstelling dat voor het gemiddelde geldt $E\{y_t\} = \underline{\mu} = \underline{0}$,

is de h° autocovariantie - matrix gedefinieerd als :

$$\Gamma(h) = E\{y_t y_{t+h}'\} = \begin{bmatrix} E\{y_{1t} y_{1,t+h}\} & E\{y_{1t} y_{2,t+h}\} & \dots & E\{y_{1t} y_{n,t+h}\} \\ E\{y_{2t} y_{1,t+h}\} & E\{y_{2t} y_{2,t+h}\} & \dots & E\{y_{2t} y_{n,t+h}\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{y_{nt} y_{1,t+h}\} & E\{y_{nt} y_{2,t+h}\} & \dots & E\{y_{nt} y_{n,t+h}\} \end{bmatrix} = \begin{bmatrix} \gamma_{y_1 y_1}(h) & \gamma_{y_1 y_2}(h) & \dots & \gamma_{y_1 y_n}(h) \\ \gamma_{y_2 y_1}(h) & \gamma_{y_2 y_2}(h) & \dots & \gamma_{y_2 y_n}(h) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{y_n y_1}(h) & \gamma_{y_n y_2}(h) & \dots & \gamma_{y_n y_n}(h) \end{bmatrix}$$

waarbij in het algemeen de kruislingse covariantie gedefinieerd is als :

$$\gamma_{y_i y_j}(h) = E\{(y_{it} - \mu_{y_i})(y_{j,t+h} - \mu_{y_j})\} \quad \forall i, j \text{ toelaatbaar} \quad (3.3.1)$$

$\{y_t\}$ wordt covariantie stationair genoemd als: $\forall t, t+h, j, j+h$ toelaatbaar geldt :

$$1) E\{y_t\} = E\{y_{t-s}\} = \underline{\mu}$$

$$2) \Gamma(h) = E\{y_t y_{t+h}'\} = E\{y_j y_{j+h}'\}$$

Bovendien geldt de volgende belangrijke gelijkheid :

$$\Gamma(h) = E\{y_t y_{t+h}'\} = \left(E\{y_{t+h} y_t'\}\right)^{(j=t+h)} = \left(E\{y_j y_{j+(-h)}'\}\right) = \Gamma(-h) \quad (3.3.2)$$

Onder de voorwaarde van absoluut sommeerbare $\{\Gamma(j)\}$ is de spectraalfunctie als volgt gedefinieerd :

$$f_{y_u y_v}(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \Gamma(j) e^{-i\omega j} = \{f_{y_u y_v}(\omega) : u, v = 1, \dots, n\} \quad (3.3.3)$$

De multivariate spectraalfunctie is volgens appendix A.3.3.1 een positief semi-definiete matrix met op de diagonaalelementen de univariate spectraalfuncties zijn van de afzonderlijke tijdreeksen. De niet-diagonaalelementen zijn de kruislingse spectraalfuncties en beschrijven de spectraalrelaties tussen twee tijdreeksen. De individuele spectraalfunctie is gedefinieerd als:

$$f_{y_u y_v}(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_{y_u y_v}(j) e^{-i\omega j} \quad (3.3.4)$$

(3.3.2) impliceert dat $\gamma_{y_u y_v}(h) = \gamma_{y_u y_v}(-h)$, maar niet noodzakelijk $\gamma_{y_u y_v}(h) = \gamma_{y_v y_u}(-h)$

Daarom is $f_{y_u y_v}$ in het algemeen een complexe functie :

$$f_{y_u y_v}(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_{y_u y_v}(j) \{\cos(\omega j) - i \sin(\omega j)\} = c_{y_u y_v}(\omega) - i q_{y_u y_v}(\omega),$$

waarbij c de samenvallende spectraaldichtheid genoemd wordt, kortweg cospectrum en q de kwadratuur spectraaldichtheid. Merk op:

$$c_{y_u y_v}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_{y_u y_v}(k) \{\cos(\omega k)\}, \text{ waarbij } c_{y_u y_v}(\omega) = c_{y_v y_u}(\omega), \text{ omdat } \cos(-\omega j) = \cos(\omega j), \quad (3.3.5)$$

$$q_{y_u y_v}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_{y_u y_v}(k) \{\sin(\omega k)\}, \text{ waarbij } q_{y_u y_v}(\omega) = -q_{y_v y_u}(\omega), \text{ omdat } \sin(-\omega j) = -\sin(\omega j) \quad (3.3.6)$$

Deze resultaten impliceren dat de niet-diagonale elementen van (3.3.3) complex geconjugeerden van elkaar zijn, dat wil zeggen dat het rij j , kolom m element van (3.3.3) de complex geconjugeerde is van het rij m kolom j element van (3.3.3). Bovendien geldt:

$$\int_{-\pi}^{\pi} q_{y_u y_v}(\omega) d\omega = \int_{-\pi}^{\pi} \{q_{y_u y_v}(\omega) + q_{y_u y_v}(-\omega)\} d\omega = \int_0^{\pi} \{q_{y_u y_v}(\omega) - q_{y_u y_v}(\omega)\} d\omega = 0$$

Met behulp van de inverse Fouriertransformatie geldt:

$$\int_{-\pi}^{\pi} c_{y_u y_v}(\omega) d\omega = \int_{-\pi}^{\pi} \{c_{y_u y_v}(\omega) + i q_{y_u y_v}(\omega)\} d\omega = \int_{-\pi}^{\pi} f_{y_u y_v}(\omega) d\omega = E(y_{u,t} - \mu)(y_{v,t} - \mu) \quad (3.3.7)$$

Uit het bovenstaande volgt al meteen de belangrijke interpretatie dat de integraal van het cospectrum over het interval $[-\pi, \pi]$ gelijk is aan de simultane covariantie van de beide reeksen. Tevens kan het cospectrum tussen y_u en y_v op frequentie ω worden geïnterpreteerd als het gedeelte van de covariantie tussen y_u en y_v , wat kan worden toegeschreven aan golven met frequentie ω . Ook voor het cospectrum geldt dat een piek in de uitslag betekent dat een groot gedeelte van de gelijktijdige beweeglijkheid tussen de reeksen wordt veroorzaakt door golven met de bij dit gebied horende frequenties.

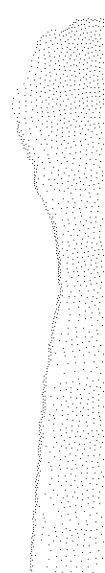
Om de interpretatie van het cospectrum te versterken en die van het kwadratuur spectrum te ontwikkelen volgt voorbeeld 3.9 uit Hamilton (1990), over twee tijdreeksen met een eindig aantal realisaties in een discrete setting.

Voorbeeld 3.9: Discrete multivariate interpretatie van co- en kwadratuurspectrum

Beschouw de volgende twee deterministische processen:

$$y_t = \mu_y + \sum_{j=1}^m a_j \cos[\omega_j(t-1)] + d_j \sin[\omega_j(t-1)]$$

$$x_t = \mu_x + \sum_{j=1}^m \alpha_j \cos[\omega_j(t-1)] + \delta_j \sin[\omega_j(t-1)]$$



waarbij N het aantal waarnemingen en $m=L(N)$ het grootste gehele getal kleiner dan of gelijk aan $N/2$. Met de Fourier-analyse wordt de volgende Fourier coëfficiënt berekend:

$$a_j = \frac{2}{N} \sum_{t=1}^N y_t \cos(\omega_j(t-1)) = \frac{2}{N} \sum_{t=1}^N (y_t - \mu_y) \cos(\omega_j(t-1))$$

omdat geldt met appendix A.3.3.2 $\sum_{t=1}^N \cos(\omega_j(t-1)) = 0$.

Nu geldt voor de kruislingse spectraalfunctie met appendix A.3.3.3 het volgende:

$$f_{yx}(\omega_j) = \frac{N}{8\pi} ((\alpha_j + i\delta_j)(a_j - id_j)) = \frac{N}{8\pi} (a_j\alpha_j + d_j\delta_j) + i\frac{N}{8\pi} (d_j\alpha_j - a_j\delta_j) = c_{yx}(\omega_j) + iq_{yx}(\omega_j) \quad (3.3.8)$$

Voor het cospectrum geldt een soortgelijke relatie als (3.3.7):

$$\begin{aligned} E(y_t - \mu_y)(x_t - \mu_x) &= \frac{1}{N} \sum_{t=1}^T (y_t - \mu_y)(x_t - \mu_x) = \\ &= \frac{1}{N} \sum_{t=1}^N \left(\sum_{j=1}^M \alpha_j \cos[\omega_j(t-1)] + \delta_j \sin[\omega_j(t-1)] \right) \times \left(\sum_{j=1}^M a_j \cos[\omega_j(t-1)] + d_j \sin[\omega_j(t-1)] \right) \\ &= \frac{1}{N} \sum_{t=1}^N \left(\sum_{j=1}^M a_j \alpha_j \cos^2[\omega_j(t-1)] + d_j \delta_j \sin^2[\omega_j(t-1)] \right) \stackrel{\text{Th}(2.1.1)}{=} \frac{1}{2} \sum_{j=1}^M (a_j \alpha_j + d_j \delta_j) \stackrel{(4.3.8)}{=} \frac{4\pi}{N} \sum_{j=1}^M c_{yx}(\omega_j) \end{aligned}$$

Het gedeelte van de covariantie tussen x en y dat wordt veroorzaakt door hun gezamenlijke afhankelijkheid van golven met frequentie ω_j wordt gegeven door $(a_j\alpha_j + d_j\delta_j)/2$.

Merk op dat als de twee processen identiek zijn geldt: $\sigma^2 = \frac{1}{2} \sum_{j=1}^M (a_j^2 + d_j^2)$.

Beschouw voor het kwadratuur spectrum een nieuwe reeks x^* , die ontstaat uit x door een faseshift van een kwart cyclus voor elke golf:

$$x_t^* = \mu_x + \sum_{j=1}^M (a_j \{\cos[\omega_j(t-1) + 1/2\pi]\} + d_j \{\sin[\omega_j(t-1) + 1/2\pi]\}) \Leftrightarrow$$

$$x_t^* = \mu_x + \sum_{j=1}^M (d_j \cos[\omega_j(t-1)] - a_j \sin[\omega_j(t-1)])$$

Uit soortgelijke calculaties volgt:

$$E(y_t - \mu_y)(x_t^* - \mu_x) = \frac{1}{T} \sum_{t=1}^T (y_t - \mu_y)(x_t^* - \mu_x) = \frac{1}{2} \sum_{j=1}^M (\alpha_j d_j - \delta_j a_j) = \frac{4\pi}{N} \sum_{j=1}^M q_{yx}(\omega_j)$$

Het kwadratuur spectrum tussen x en y op frequenties ω_j is gelijk aan de proportie van de gelijktijdige covariantie tussen x^* en y , dat wordt veroorzaakt door golven van frequentie ω_j .

Merk op dat als de twee processen identiek zijn geldt: $f_{yx} = \frac{N}{8\pi} \sum_{j=1}^M (a_j^2 + d_j^2)$ en $q_{yx}(\omega_j) = 0 \quad \forall \omega_j$.

□

Dit voorbeeld toont aan dat de berekende gelijktijdige covariantie tussen twee reeksen lager zou kunnen zijn dan op basis van de werkelijke relatie te verwachten is, wat veroorzaakt wordt doordat op een bepaald punt in de tijd de twee reeksen zich in een verschillende fase van de cyclus bevinden. In voorbeeld 3.9 is met de introductie van x^* elke opbouwende golf een kwart cyclus verschoven. Additionele complexiteit wordt geïntroduceerd als de in het frequentiedomein constituerende golven van twee tijdreeksvariabelen variërende faseshifts ten opzichte van elkaar kennen. In de economie kunnen bijvoorbeeld alleen de middellange conjunctuurgolven van een bepaalde tijdreeks eerder reageren op een recessie dan de conjunctuurgolven van een andere tijdreeks. In spectraalterminologie zijn dan de lange trendachtige golven en de korte ruisgolven in dezelfde fase, terwijl er een faseshift bestaat voor de middellange conjunctuurgolven.

Het kwadratuur spectrum beschrijft dergelijke faserelaties en in de volgende subparagraaf worden maatstaven gedefinieerd om faseshifts in het kwadratuur te expliciteren.

3.3.2: Multivariate extensie: samenhang en verschuiving

In de vorige subparagraaf is de toolkit van de spectraalanalyse uitgebreid naar een multivariate context, die noodzakelijk is voor het bepalen van relaties tussen afzonderlijke tijdreeksen. Deze relaties zijn samengevat in het cospectrum, dat de gelijktijdige kruislingse covariantie ontleedt naar de afzonderlijke frequenties en het kwadratuur spectrum, dat faseshifts genereert als functie van de frequentie. Het doel van deze paragraaf is om uit beide spectra praktische maatstaven te definiëren voor samenhang en verschuiving, die zullen worden aangeduid als spectraalmaatstaven.

Beschouw wederom (3.3.4), de kruislingse spectraalfunctie :

$$f_{xy}(\omega) = c_{xy}(\omega) - iq_{xy}(\omega).$$

Deze functie kan gerepresenteerd worden in de vorm van poolcoördinaten :

$$f_{xy}(\omega) = A_{xy}(\omega)e^{i\varphi_{xy}(\omega)}, \quad (3.3.9)$$

waarbij $A_{xy}(\omega)$ wordt gedefinieerd als :

$$A_{xy}(\omega) = [c_{xy}^2(\omega) + q_{xy}^2(\omega)]^{1/2} = [f_{xy}(\omega) * \overline{f_{xy}(\omega)}] \text{ is het kruislingse amplitude - spectrum} \quad (3.3.10)$$

en $\overline{f_{xy}(\omega)}$ de complex geconjugeerde is van $f_{xy}(\omega)$.

Het kruislingse amplitudespectrum is de multivariate pendant van het amplitude-effect zoals geïntroduceerd in de paragraaf over de filters. Deze maatstaf is een verhouding van de absolute uitslag van de spectraalfuncties van de twee reeksen ten opzichte van elkaar. Van subparagraaf 3.2.3 over filters is bekend dat de absolute waarde van de transferfunctie de multiplicatiefactor is waarmee de uitslag van het spectrum van x vermenigvuldigd moet worden om de uitslag van het spectrum van y te krijgen, vergelijk figuur 3.6 en figuur 3.8. Het multivariate equivalent is het winstspectrum en is een standaardisatie van het kruislingse amplitudespectrum. De winst zal gelijk blijken te zijn aan de absolute transferfunctie tussen twee reeksen. De definitie luidt als volgt:

$$\Psi_{xy}(\omega) = \frac{A_{xy}(\omega)}{f_{yy}(\omega)} \text{ is de winst van } y \text{ over } x. \quad (3.3.11)$$

Een gerelateerde belangrijke maatstaf is de kwadratische coherentie en dit is een dimensieloze maat voor de samenhang tussen twee tijdreeksen als functie van de frequentie. Door de normalisatie geldt voor deze maatstaf een dimensieloosheid en kan deze kwadratische coherentie niet groter zijn dan één, zie hiervoor appendix A.3.3.4. De kwadratische coherentie is gedefinieerd als de voor beide spectra gestandaardiseerde kruislingse amplitude en luidt als volgt:

$$K_{xy}^2(\omega) = \frac{|f_{xy}(\omega)|^2}{f_{xx}(\omega)f_{yy}(\omega)} = \frac{c_{xy}^2(\omega) + q_{xy}^2(\omega)}{f_{xx}(\omega)f_{yy}(\omega)} \text{ is de kwadratische coherentie.} \quad (3.3.12)$$

Feitelijk kan deze maatstaf geïnterpreteerd worden als de mate van lineaire samenhang, uitgesplitst naar de frequentie, tussen twee variabelen analoog aan de R^2 -maatstaf uit de lineaire regressie. Vooruitlopend zal in (3.3.24) blijken dat de kwadratische coherentie evenredig is aan de mate van lineaire samenhang tussen twee tijdreeksen.

De kwadratische coherentie is een gestandaardiseerde maat en als er filters toegepast worden op de tijdreeksen is er als gevolg van deze standaardisatie geen sprake van verandering in deze maatstaf. Deze bewering uit Priestley(1981) wordt als volgt gesteld:

Theorema 3.3.1: Invariantie van kwadratische coherentie voor lineaire filters

Veronderstel een bivariaat stationair proces, waarbij op elke tijdreeks het algemene lineaire filter wordt toegepast, namelijk:

$$x_t = \sum_{u=-\infty}^{\infty} a_u L^u(x_t) = \alpha(B)x_t \text{ en } y_t = \sum_{u=-\infty}^{\infty} b_u L^u(y_t) = \beta(B)y_t.$$

Onder de voorwaarde dat $\alpha(e^{-i\omega}) = \beta(e^{-i\omega}) \neq 0 \quad \forall \omega$ geldt : $K_{xy}^2 = K_{x'y'}^2$.



Bewijs :

Met de transfer - functie van (3.2.13) geldt :

$$f_{x'x'}(\omega) = \alpha(e^{-i\omega})^2 f_{xx}(\omega)$$

$$f_{y'y'}(\omega) = \beta(e^{-i\omega})^2 f_{yy}(\omega)$$

$$f_{x'y'}(\omega) = \alpha(e^{-i\omega})\beta(e^{-i\omega})f_{xy}(\omega)$$

$$K_{x'y'}^2 = \frac{|f_{x'y'}(\omega)|^2}{\{f_{x'x'}(\omega)f_{y'y'}(\omega)\}} = \frac{|\alpha(e^{-i\omega})\beta(e^{-i\omega})f_{xy}(\omega)|^2}{\{\alpha(e^{-i\omega})^2 f_{xx}(\omega)\beta(e^{-i\omega})^2 f_{yy}(\omega)\}} = \frac{|f_{xy}(\omega)|^2}{\{f_{xx}(\omega)f_{yy}(\omega)\}} = K_{xy}^2.$$

□
Merk echter op dat zowel x als y stationaire processen zijn, waarvan de spectraalfunctie en de kruislingse spectraalfunctie bestaan voor iedere frequentie. Als een van beide tijdreeksen een 'trend' bevat (die door de filters eruit gehaald wordt) en dus niet stationair is, zal de bovenstaande stelling in het algemeen niet opgaan.

De gedefinieerde spectraalmaatstaven (3.2.2), (3.2.3) en (3.2.4) zijn amplitude-gerelateerde maatstaven en meten alle een vorm van samenhang tussen reeksen. Er is tevens een groep maatstaven die verschuiving van reeksen meet.

Voor filters is een fase-effect afgeleid dat de verschuiving van een tijdreeks veroorzaakt door toepassing van een filter beschrijft. Het multivariate equivalent van deze maatstaf beschrijft de fase van twee gegeven tijdreeksen ten opzichte van elkaar en de bijbehorende concepten zijn als volgt gedefinieerd:

$$\sin \varphi_{xy}(\omega) = \frac{-q_{xy}(\omega)}{A_{xy}(\omega)}, \quad \cos \varphi_{xy}(\omega) = \frac{c_{xy}(\omega)}{A_{xy}(\omega)},$$

$$\varphi_{xy}(\omega) = \tan^{-1} \left(\frac{-q_{xy}(\omega)}{c_{xy}(\omega)} \right) \quad \text{is het fasespectrum,} \quad (3.3.13)$$

$$s(\omega) := - \left\{ \frac{\varphi(\omega)}{\omega} \right\} \quad \text{is het shiftspectrum.} \quad (3.3.14)$$

Hierbij meet $s(\omega)$ ¹⁵ wederom, geheel analoog aan de shift van de filters, de verschuiving in tijdsperioden van twee tijdreeksen ten opzichte van elkaar uitgesplitst naar frequentie ω . Deze maatstaf bepaalt voor de Fourier-representatie van een tijdreeks als de oneindige som van golffuncties per golf de verschuiving van twee tijdreeksen ten opzichte van elkaar.

Om de interpretatie van de gedefinieerde maatstaven te versterken worden ter illustratie een aantal voorbeelden uitgewerkt. Het eenvoudigste voorbeeld is een model met twee ongecorreleerde processen, namelijk:

Voorbeeld 3.10: Ongecorreleerde processen

Beschouw het eenvoudige voorbeeld, waarbij x_t en y_s ongecorreleerde processen zijn, namelijk :

$$\text{cov}\{x_t, y_{t+s}\} = 0 \quad \forall s, t. \quad \text{Dan geldt vanzelfsprekend :}$$

$$\gamma_{xy}(s) = 0 \quad \forall s, \quad \rho_{xy}(s) = 0 \quad \forall s,$$

$$\text{en met definitie (3.1.4) } f_{xy}(\omega) = 0 \quad \forall \omega.$$

De gedefinieerde maatstaven genereren dan :

$$c_{xy}(\omega) = 0 \quad q_{xy}(\omega) = 0 \quad A_{xy}(\omega) = 0 \quad K_{xy}^2(\omega) = 0 \quad \forall \omega$$

□

¹⁵ de definitie van shift is overgenomen van Koopmans(1974) bladzijde 95.

Indien de onderliggende tijdreeksen ongecorrleerd zijn detecteren de spectraalmaatstaven geen enkele samenhang, wat tot uitdrukking komt in de waarde van nul voor alle spectraalmaatstaven. Het cospectrogram, c_{xy} , incorporeert de covarianties en genereert de waarde nul voor ongecorrleerde processen. Dit geldt in gestandaardiseerde vorm ook van de winst A_{xy} . Het kwadratuur spectrum, q_{xy} , beschrijft faseverschillen van samenhangende golven en genereert de waarde nul als de golven ongecorrleerd zijn. De kwadratische coherentie is een equivalent van de correlatie uitgesplitst naar frequentie en genereert bij afwezigheid van correlatie de waarde nul.

In voorbeeld 3.11¹⁶ wordt er een vertraagde relatie met ruis geïntroduceerd en geanalyseerd op welke manier deze vertraging met ruis gedetecteerd wordt door de spectraalmaatstaven, waarbij de interesse vooral uitgaat naar de maatstaven coherentie en shift.

Voorbeeld 3.11: Lineaire regressie met vertraging

In dit voorbeeld geldt de volgende relatie:

$$x_t = ay_{t-d} + e_t \quad e_t \sim (0, \sigma^2) \text{ o.i.v.}$$

Dan geldt :

$$\gamma_{xy}(s) = a\gamma_{yy}(s-d)$$

$$f_{xy}(\omega) = \frac{a}{2\pi} \sum_{s=-\infty}^{\infty} \gamma_{yy}(s-d)e^{-i\omega s} = ae^{-i\omega d} f_{yy}(\omega) = af_{yy}(\omega) \{ \cos(\omega d) - i \sin(\omega d) \}$$

$$= af_{yy}(\omega) \{ c_{xy}(\omega) - iq_{xy}(\omega) \}$$

$$f_{xx}(\omega) = \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} \{ a^2 \gamma_{yy}(s-d) + \gamma_{ee}(s) \} e^{-i\omega s} = a^2 f_{yy}(\omega) + f_e(\omega), \text{ waarbij } f_e(\omega) = \frac{\sigma^2}{2\pi}.$$

Hiermee wordt verkregen :

$$c_{xy}(\omega) = a \cos(\omega d) f_{yy}(\omega), \quad q_{xy}(\omega) = a \sin(\omega d) f_{yy}(\omega),$$

$$A_{xy}(\omega) = [c_{xy}^2(\omega) + q_{xy}^2(\omega)]^{1/2} = af_{yy}(\omega),$$

$$\Psi_{xy}(\omega) = \frac{A_{xy}(\omega)}{f_{yy}(\omega)} = a.$$

$$\varphi_{xy}(\omega) = \tan^{-1} \left(\frac{-q_{xy}(\omega)}{c_{xy}(\omega)} \right) = \tan^{-1} (-\tan(\omega d)) = -\omega d,$$

$$s_{xy}(\omega) = - \left\{ \frac{\varphi_{xy}(\omega)}{\omega} \right\} = d,$$

$$K_{xy}^2(\omega) = \frac{|f_{xy}(\omega)|^2}{f_{xx}(\omega)f_{yy}(\omega)} = \frac{a^2 f_{yy}^2(\omega)}{a^2 f_{yy}^2(\omega) + f_e(\omega)} = \frac{1}{1 + \left(\frac{f_e(\omega)}{a^2 f_{yy}(\omega)} \right)}$$

Merk op dat als $e_t \equiv 0$, dan geldt $f_e(\omega) = 0$ en $K_{xy}^2(\omega) = 1 \quad (\forall \omega)$.

□

In het voorbeeld genereert het winstspectrum de multiplicatiewaarde tussen de tijdreeksen, a , en genereert de shift, s , de verschuiving van d tijdseenheden. De kwadratische coherentie genereert een waarde van een als er sprake is van een exact lineair verband bij afwezigheid van de storingsterm. De verhouding van de spectraalfunctie van de storingsterm ten opzichte van de spectraalfunctie van de onafhankelijke variabele y in bovenstaand voorbeeld bepaalt in hoeverre de kwadratische coherentie verwijderd is van de perfecte samenhang.

¹⁶ bron: Priestly (1981)



3.3.3: Bivariate spectraalanalyse toegepast op het algemene lineaire model

Het is al eerder gesteld dat de uiteindelijke doelstelling van deze studie het kwantificeren is van relaties tussen tijdreeksen in termen van samenhang en verschuiving, zoals gedemonstreerd in voorbeeld 3.10 en voorbeeld 3.11. Deze illustraties zijn twee toepassingen van het algemene lineaire model en de spectraalmaatstaven worden in deze paragraaf voor dit algemene model gegeneraliseerd. Het algemene lineaire model stelt dat een tijdreeks x is opgebouwd als een lineaire combinatie van een andere tijdreeks y en een met y ongecorreleerde storingsterm. Deze veronderstelling van ongecorreleerdheid is een uitvloeisel van het feit dat x in het algemene lineaire model een kleinste kwadraten lineaire benadering is van y . Dit zal worden aangetoond door een minimaliseringsoplossing in het frequentiedomein. Beschouw het volgende algemene lineaire model met ruis:

$$x_t = \sum_{u=-\infty}^{\infty} g(u)y_{t-u} + N_t \quad (3.3.15)$$

en veronderstel dat het ruis N_t niet gecorreleerd is met y_t .

Er geldt:

$$\gamma_{xy}(s) = E[x_t y_{t+s}] = \sum_{u=-\infty}^{\infty} g(u) E[y_t y_{t+s-u}] = \sum_{u=-\infty}^{\infty} g(u) \gamma_{yy}(s-u)$$

Merk op dat bij het laatste '='-teken de kruislingse produkttermen wegvallen omdat N_t niet gecorreleerd is met y_t en een verwachting van nul heeft.

Nu geldt dat:

$$f_{xy}(\omega) = \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} e^{-i\omega s} \left\{ \sum_{u=-\infty}^{\infty} g(u) \gamma_{yy}(s-u) \right\} = \sum_{u=-\infty}^{\infty} g(u) e^{-i\omega u} \left\{ \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} e^{-i\omega(s-u)} \gamma_{yy}(s-u) \right\} = T(\omega) f_{yy}(\omega)$$

$$\Leftrightarrow T(\omega) = \frac{f_{xy}(\omega)}{f_{yy}(\omega)} \quad (3.3.16)$$

$$\gamma_{xx}(s) = E[x_t x_{t+s}] = E \left\{ \left(\sum_{u=-\infty}^{\infty} g(u) y_{t-u} + N_t \right) \left(\sum_{v=-\infty}^{\infty} g(v) y_{t+s-v} + N_{t+s} \right) \right\} =$$

$$\sum_{v=-\infty}^{\infty} g(v) \sum_{u=-\infty}^{\infty} g(u) \gamma_{yy}(s+u-v) + \gamma_{NN}(s)$$

Met theorema 3.2.5, geldt nu de volgende relatie:

$$f_{xx}(\omega) = |T(\omega)|^2 f_{yy}(\omega) + f_{NN}(\omega) \quad (3.3.17)$$

Omdat $f_{yy}(\omega)$ een reële functie is geldt voor het fasespectrum met appendix A.4.5.4:

$$\varphi_{xy}(\omega) = \arg \{ f_{xy}(\omega) \} = \arg \{ T(\omega) f_{yy}(\omega) \} = \arg \{ T(\omega) \} \quad (3.3.18)$$

Het fasespectrum hangt dus louter af van de fase - shifts van de transfer - functie $T(\omega)$.

Bovendien geldt voor het winst - spectrum:

$$\Psi_{xy}(\omega) = \frac{\alpha_{xy}(\omega)}{f_{yy}(\omega)} = \frac{|f_{xy}(\omega)|}{f_{yy}(\omega)} = \frac{|T(\omega) f_{yy}(\omega)|}{f_{yy}(\omega)} = |T(\omega)| \quad (3.3.19)$$

De shiftrelatie tussen twee reeksen hangt, onder de veronderstelling van ongecorreleerdheid van de storing met de reeks y , slechts af van de transferfunctie (3.3.16). Deze functie is exact analoog aan de transferfunctie die is gebruikt voor de filters, waarbij de functie in dit model bepaald wordt door de spectraalfuncties van de tijdreeksen. Alle relatiekarakteristieken tussen beide tijdreeksen worden beschreven door deze transferfunctie. Zo blijkt uit (3.3.18) dat de fasemaatstaf van de twee reeksen louter een functie is van de transferfunctie, waarbij de fase besloten ligt in het complexe gedeelte van de transferfunctie. Uit (3.3.19) blijkt het winstspectrum ook louter een functie te zijn van de transfer van de transferfunctie, waarbij de multiplicatiefactor van de relatieve uitslag van beide spectra gelijk is aan de absolute transferfunctie.

Om de importantie van de transferfunctie en de interpretatie van de coherentie als correlatiemaatstaf nader te ontwikkelen wordt een kleinste kwadraten benadering in het lineaire model voor de reeksen afgeleid. De argumentatie gaat met Priestley (1981) als volgt:

Beschouw de volgende kleinste kwadraten benadering van het algemene lineaire model van (3.3.15):

$$\min_{g(u)} E \left\{ x_t - \sum_{u=-\infty}^{\infty} g(u) y_{t-u} \right\}^2, \quad (3.3.20)$$

$$\text{waarbij } N_t = x_t - \sum_{u=-\infty}^{\infty} g(u) y_{t-u}.$$

$$\text{De variantie van } N_t \text{ is dan } \gamma_{NN}(0) = E \{ N_t^2 \} = \int_{-\pi}^{\pi} f_{NN}(\omega) d\omega \stackrel{\text{Th 3.2.5+(3.3.16)}}{=} =$$

$$\int_{-\pi}^{\pi} \left\{ f_{xx}(\omega) - 2\pi \left[f_g(\omega) f_{yx}(\omega) + f_g^*(\omega) f_{xy}(\omega) \right] + (2\pi)^2 f_g(\omega) f_g^*(\omega) f_{yy}(\omega) \right\} d\omega. \quad (3.3.21)$$

Het minimaliserings - probleem (3.3.20) is nu omgezet in de minimalisatie van (3.3.21) met $f_g(\omega)$.

De oplossing wordt gegeven door: $f_g(\omega) = \frac{1}{2\pi} \frac{f_{xy}(\omega)}{f_{yy}(\omega)}$ ($f_{yy}(\omega) > 0$) en verder geldt:

$$\gamma_{NN}(0) = \int_{-\pi}^{\pi} \left[f_{xx}(\omega) - \frac{f_{xy}(\omega) f_{xy}^*(\omega)}{f_{yy}(\omega)} \right] d\omega = \int_{-\pi}^{\pi} \left[f_{xx}(\omega) (1 - K_{xy}^2(\omega)) \right] d\omega. \quad (3.3.22)$$

De gewichten g_j worden gegeven door de inverse Fourier transformatie van $f_g(\omega)$,

$$g(u) = \int_{-\pi}^{\pi} f_g(\omega) e^{i\omega u} d\omega \stackrel{(3.3.16)}{=} \frac{1}{2\pi} \int_{-\pi}^{\pi} T(\omega) e^{i\omega u} d\omega \quad u = 0, \pm 1, \pm 2, \dots \quad (3.3.23)$$

In appendix A.3.3.5 wordt aangetoond dat $f_g(\omega)$ inderdaad de minimumfunctie is. Vergelijking (3.3.22) is een grenswaarde, omdat dit de waarde van de resterende variantie geeft nadat de best mogelijke, in kleinste kwadraten opzicht, lineaire relatie tussen x en y is bepaald met behulp van een oneindig aantal parameters. Als er een eindig aantal parameters lineaire benadering bepaald wordt zal de residuele variantie niet lager zijn dan deze grens en in het algemeen zelfs hoger.

Uit (3.3.22) blijkt tevens het correlatiekarakter van de coherentie. Een exact lineair verband zou namelijk impliceren dat de variantie gelijk is aan nul en dit correspondeert met een kwadratische coherentie van één.

Voorbeeld 3.12: Een uitgebreid voorbeeld

$$y_t = b_1 y_{t-1} + e_{y,t}$$

$$x_t = a_1 y_{t-4} + a_2 y_{t-5} + e_{x,t},$$

waarbij $|b_1| < 1$ en $\{e_{y,t}, e_{x,t}\}$ een vector - reeks ongecorreleerde toevalsvariabelen met:

$$E(e_{y,t}^2) = \sigma_{yy} = 1, \quad E(e_{x,t}^2) = \sigma_{xx} = 1, \quad E(e_{y,t} e_{x,t+h}) = 0 \quad \forall t, h.$$

Dan geldt:

$$\begin{aligned} f_{yy}(\omega) &= g_{yy}(\omega) (1 - b_1 e^{-i\omega})^{-1} (1 - b_1 e^{i\omega})^{-1}, \\ f_{xx}(\omega) &= (a_1 e^{-4i\omega} + a_2 e^{-5i\omega}) (a_1 e^{4i\omega} + a_2 e^{5i\omega}) f_{yy}(\omega) + g_{xx}(\omega), \\ f_{yx}(\omega) &= (a_1 e^{-4i\omega} + a_2 e^{-5i\omega}) f_{yy}(\omega) \quad [NB: \gamma_{yx}(h) = E\{y_t x_{t+h}\} = a_1 \gamma_y(h) + a_2 \gamma_y(h-1)] \\ \text{waarbij } g_{yy}(\omega) &= 1, \quad g_{xx}(\omega) = 1. \end{aligned}$$

Het cospectrum en de kwadratuur spectraaldichtheid zijn:



$$c_{yx}(\omega) = \operatorname{re} \ddot{\operatorname{el}}(f_{yx}(\omega)) = (a_1 \cos(4\omega) + a_2 \cos(5\omega)) f_{yy}(\omega),$$

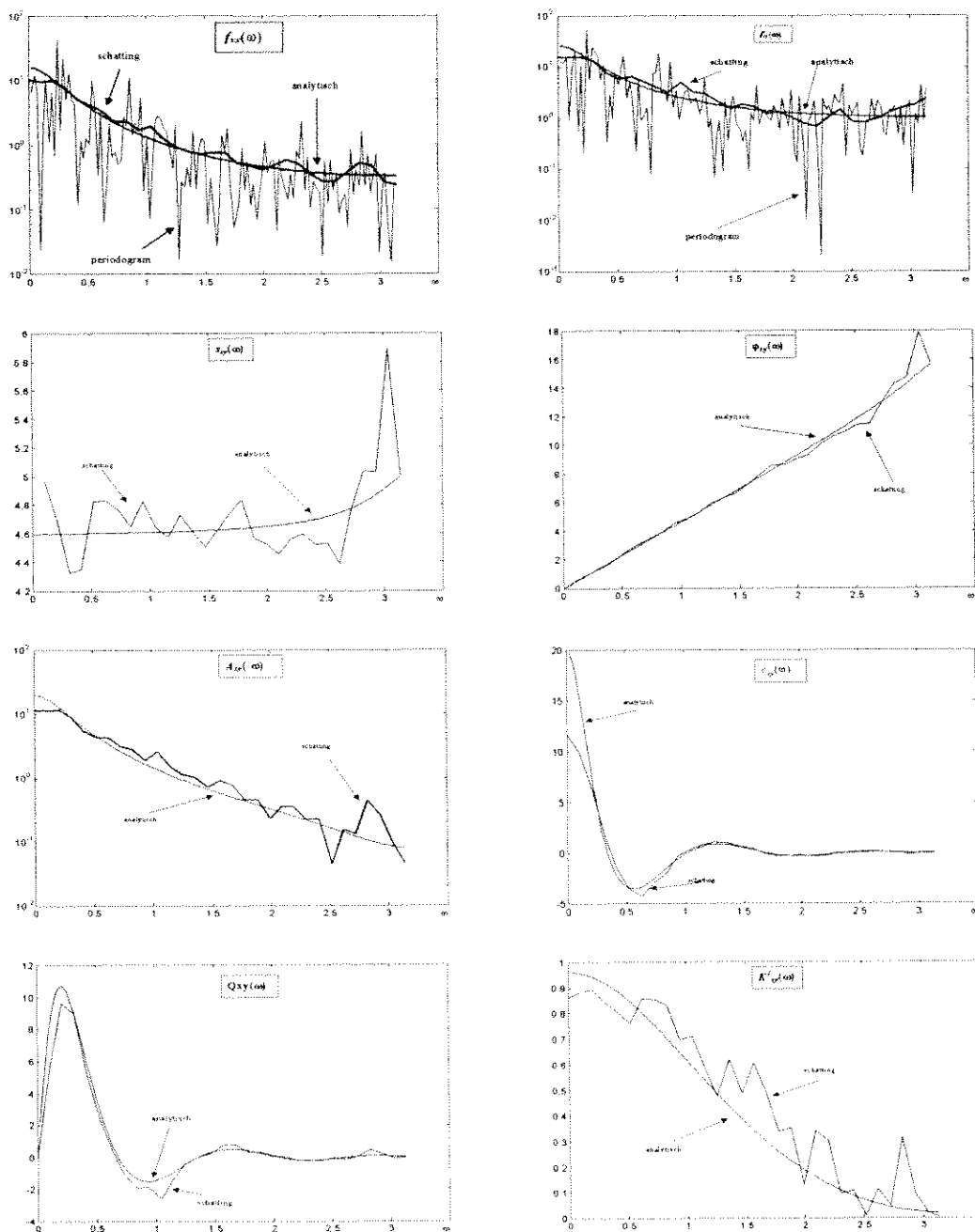
$$q_{yx}(\omega) = \operatorname{imaginair}(f_{yx}(\omega)) = (a_1 \sin(4\omega) + a_2 \sin(5\omega)) f_{yy}(\omega),$$

$$\varphi_{yx}(\omega) = \tan^{-1} \left[\frac{-q_{yx}(\omega)}{c_{yx}(\omega)} \right] = \tan^{-1} \left[\frac{-(a_1 \cos(4\omega) + a_2 \cos(5\omega))}{(a_1 \sin(4\omega) + a_2 \sin(5\omega))} \right]$$

$$A_{yx}(\omega) = [c_{yx}^2(\omega) + q_{yx}^2(\omega)]^{1/2} = f_{yy}(\omega) [a_1^2 + 2a_1 a_2 (\cos(4\omega)\cos(5\omega) + \sin(4\omega)\sin(5\omega)) + a_2^2]^{1/2},$$

$$K_{yx}^2(\omega) = \frac{|f_{yx}(\omega)|^2}{f_{xx}(\omega)f_{yy}(\omega)} = \frac{f_{yy}^2(\omega) [a_1^2 + 2a_1 a_2 (\cos(4\omega)\cos(5\omega) + \sin(4\omega)\sin(5\omega)) + a_2^2]}{f_{yy}(\omega) \{a_1^2 + 2a_1 a_2 \cos(\omega) + a_2^2\} f_{yy}(\omega) + g_{xx}(\omega)}$$

Figuur 3.13: spectraalmaatstaven behorende bij voorbeeld 3.12



Voorbeeld 3.12 gaat over een AR(1) proces y en een daarvan afgeleid proces x , waarbij alleen de storingsterm ervoor zorgt dat er geen exacte lineaire benadering tussen beide reeksen bestaat. Het voorbeeld is gekwantificeerd in de grafieken met als waarden $a_1=0.5$, $a_2=0.75$, $b_1=0.75$.

In figuur 3.13 zijn de spectra van beide reeksen, f_{xx} en f_{yy} uitgezet tegen de frequentie, waarbij de y -as een logaritmische schaal kent. In beide grafieken zijn drie reeksen uitgezet, te weten het periodogram, de spectraalschatter en de analytische reeks. Het periodogram wordt uitgebreid besproken in paragraaf 5 over spectraalschatters en voor dit moment is het voldoende om deze te interpreteren als ruwe schatting. Omdat de reeksen bekend zijn, is in het voorbeeld de spectraalschatter ook analytisch bepaald en uitgezet in de grafiek. De gelijkheid van de analytische en de geschatte reeks is dus een maat voor de kwaliteit van de schatter. Er wordt overigens gewerkt met 1000 waarnemingen. Uit beide grafieken blijkt dat het spectrogram een grote waarde aanneemt bij lage frequenties, wat erop duidt dat veel van de totale beweeglijkheid zit in de lange golven.

In het voorbeeld is tevens de kruislingse spectraalfunctie berekend en ontleed in een reële component, het cospectrum en een imaginaire component, het kwadratuur spectrum. Het cospectrum meet de gelijktijdige beweeglijkheid van beide reeksen en uit de figuur blijkt dat veel van deze gezamenlijke beweeglijkheid wordt verklaard door de hele lange golven, waarbij er sprake is dat de y reeks en de x -reeks positief gecorreleerd zijn. Voor de middellange golven geldt een negatieve correlatie, waarmee is gesteld dat op deze frequenties beide golven tegengesteld bewegen.

Van het kwadratuur is bekend dat het de verschuiving van reeksen ten opzichte van elkaar meet. Uit de figuur blijkt dat de bij korte frequenties behorende lange golven het meeste faseverschil kennen en de korte golven juist helemaal geen. De middellange golven kennen een negatief kwadratuur spectrum, wat wil zeggen dat voor deze golven de reeks x voorloopt op de reeks y in plaats van omgekeerd. Deze vier maatstaven, dat wil zeggen beide spectra en het kruislingse spectrum ontleedt in cospectrum en kwadratuur spectrum, worden geschat aan de hand van de data. Alle overige maatstaven zijn wiskundige combinaties van deze vier schatters.

Het winstpectrum, niet weergegeven, geeft de multiplicator waarmee y vermenigvuldigd moet worden om x te krijgen. Het amplitudespectrum is een maat voor deze multiplicatie en geeft dus het spectrum van x op een indirecte manier weer. De belangrijkste maatstaven zijn coherentie en fase/shift.

De coherentie meet de correlatie tussen de beide reeksen. Uit het voorbeeld is bekend dat de reeksen op de storingsterm na gecorreleerd zijn. Dat betekent dat voor de lange golven een hoge waarde voor de coherentie wordt verwacht en voor de korte ruisgolven een coherentie van nul. Dit fenomeen doet zich inderdaad voor en is zichtbaar in de figuur.

Uit het voorbeeld is tevens bekend dat x een verschuiving is van y met tussen de vier en vijf perioden. Het fasespectrum genereert ook, zoals grafisch weergegeven, een redelijke rechte lijn, waarvan de hellingshoek is weergegeven in het shiftspectrum. Deze hellingshoek is tevens het aantal perioden verschuiving, wat zoals verwacht tussen de vier en de vijf ligt. Merk overigens op dat als blijkt dat er, zeg 4.7 perioden verschoven is, er sprake van is dat de waarneming van vijf perioden geleden meeweegt. Het maximale aantal perioden dat teruggekeken wordt is dus de naar boven afgeronde absolute waarde van de faseschatting.

In het voorbeeld zijn twee bekende reeksen beschouwd en is gekeken in hoeverre de karakteristieken van deze reeksen tot uitdrukking komen in de verschillende maatstaven. Indien het uitgangspunt de analyse van de conjuncturele verhouding tussen twee reeksen is, wordt eerst gedefinieerd welke frequenties bij conjuncturele golven horen en vervolgens gekeken naar het gedrag van de afzonderlijke maatstaven op dit frequentie-interval. Het is mogelijk om referentiewaarden voor de gemiddelden van de maatstaven op dit interval te definiëren, waarmee een selectiemechanisme verkregen wordt om uit een verzameling reeksen de met een referentiereeks samenhangende reeksen uit te filteren.



3.3.4: Het multivariate algemene lineaire model en partiële maatstaven

Tot slot van deze paragraaf wordt het algemene lineaire model multivariaat gegeneraliseerd. Relaties in het algemene lineaire model worden doorgaans geïnterpreteerd in de context van lineaire systemen, waarbij y en x respectievelijk de input en output van een model voorstellen. In het multivariate model bestaan er q input processen, $y_{1,t}, y_{2,t}, \dots, y_{q,t}$ en p output processen $x_{1,t}, x_{2,t}, \dots, x_{p,t}$, alle met een verwachting van nul. De meest algemene vorm is een representatie van elk output als de gewogen som van alle input processen. Dit wordt als volgt genoteerd¹⁷:

$$x_{i,t} = \sum_{u=-\infty}^{\infty} g_{i1}(u)y_{1,t-u} + \sum_{u=-\infty}^{\infty} g_{i2}(u)y_{2,t-u} + \dots + \sum_{u=-\infty}^{\infty} g_{iq}(u)y_{q,t-u} + N_{i,t}, \quad i=1,2,\dots,p, \quad (3.3.24)$$

waarbij $N_{i,t} \sim N(0, \sigma^2)$ o.i.d en ongecorrleerd met iedere $y_{j,t}$.

In vectornotatie:

$$\underline{x}_t = \sum_{u=-\infty}^{\infty} \underline{g}(u)\underline{y}_{t-u} + \underline{N}_t. \quad (3.3.25)$$

Geheel in lijn met de univariate case:

$$\text{de kruislings spectraalmatrix } \underline{f}_{xy}(\omega) = T(\omega)\underline{f}_{yy}(\omega), \quad (3.3.26)$$

$$T(\omega) = \underline{f}_{xy}(\omega)\underline{f}_{yy}^{-1}(\omega), \text{ een } p \times q \text{ - transfer - matrix als geldt dat } \left(\underline{f}_{yy}(\omega)\right) \text{ niet - singulier is,} \quad (3.3.27)$$

$$\underline{f}_g(\omega) = \frac{1}{2\pi} T(\omega) \text{ het spectraalequivalent van de optimale coëfficiëntenmatrix.} \quad (3.3.28)$$

$$\underline{f}_{xx}(\omega) = T(\omega)\underline{f}_{yy}(\omega)\underline{T}^*(\omega) + \underline{f}_{NN}(\omega) = \underline{f}_{xy}(\omega)\underline{f}_{yy}^{-1}(\omega)\underline{f}_{xy}^*(\omega) + \underline{f}_{NN}(\omega),$$

$$\text{met de gekwadrateerde coherentie matrix: } K_{x_1 \dots x_p, y_1 y_2 \dots y_q}^2(\omega) = \frac{\underline{f}_{xy}(\omega)\underline{f}_{yy}^{-1}(\omega)\underline{f}_{xy}^*(\omega)}{\underline{f}_{xx}(\omega)}. \quad (3.3.29)$$

$$\text{De fasematrix wordt als volgt berekend: } \varphi_{xy}(\omega) = \arg\{\underline{f}_g(\omega)\} \quad (3.3.30)$$

Geheel in analogie met (3.3.22) en (3.3.23) is (3.3.30) het spectraalequivalent van de optimale coëfficiëntenmatrix. In de praktijk zal er in tegenstelling tot het algemene multivariate lineaire model (3.3.24) een restrictie gelden voor de parameter u : $-a < u < b$. Het is immers zeer onwaarschijnlijk dat er oneindig veel data uit het verleden bekend zijn en het is per definitie onmogelijk dat er oneindig veel data van de toekomst beschikbaar zijn. Volgens Priestley (1991)¹⁸ blijft de spectraalmatrix (3.3.28) echter de optimale matrix, waaruit de $g_{ij}(u)$'s berekend kunnen worden, ongeacht de restricties a en b . Hannan (1970) in paragraaf 8.7. bevestigt dit door te stellen dat de gevolgde procedure een groot voordeel heeft, omdat elke coëfficiënt $g_{ij}(u)$ individueel berekend wordt. De restricties a en b kunnen worden verhoogd zonder dat de eerder berekende $g_{ij}(u)$'s veranderen en hoeven daarmee niet herberekend te worden.

In het multivariate model kennen de gegeneraliseerde maatstaven een veelvoudig karakter. De elementen van de gekwadrateerde coherentiematrix (3.3.29) zijn bijvoorbeeld maatstaven voor de individuele coherenties tussen twee reeksen in een multivariate context. Een hoge coherentie kan veroorzaakt worden door een intrinsieke associatie tussen deze reeksen, maar ook door het feit dat beide reeksen hoog gecorreleerd zijn met een derde reeks. Om de intrinsieke waarde van de coherentie te isoleren in een multivariate context wordt de partiële gekwadrateerde coherentiematrix geïntroduceerd. Om dezelfde reden wordt er ook een partiële fasematrix geïntroduceerd. Deze partiële maatstaven meten hun karakteristieke relatie tussen twee reeksen, nadat de gezamenlijke invloed van de overige reeksen is verwijderd.

¹⁷ bron Priestly(1981)

¹⁸ vergelijk in Priestly paragraaf 9.2.1, waarbij u geen restricties kent en paragraaf 10.4.3, waarbij u strikt positief moet zijn.

Beschouw de volgende vereenvoudigde versie van het multivariate model, waarbij er een enkele outputvariabele, x_t , is en er twee input variabelen, y_{1t} en y_{2t} , zijn. De afleiding van de partiële coherentie begint met het elimineren van de invloed van y_{2t} op x_t en y_{1t} . De voor de invloed van y_{2t} geschoonde nieuwe variabelen zijn als volgt gedefinieerd:

$$\eta_{1,t} = x_t - \sum_{u=-\infty}^{\infty} b_1(u)y_{2,t-u}, \quad \eta_{2,t} = y_{1,t} - \sum_{u=-\infty}^{\infty} b_2(u)y_{2,t-u},$$

waarbij $\{b_1(u)\}$ en $\{b_2(u)\}$ worden gevormd door minimalisatie van $E\{\eta_{1,t}^2\}$ and $E\{\eta_{2,t}^2\}$

$$\text{De transfer - functies zijn : } B_1(\omega) = \frac{f_{xy_2}(\omega)}{f_{y_2y_2}(\omega)} \quad \text{en} \quad B_2(\omega) = \frac{f_{y_1y_2}(\omega)}{f_{y_2y_2}(\omega)}.$$

Analoog aan eerdere calculaties is de kruislingse spectraaldichtheid :

$$f_{\eta_1\eta_2}(\omega) = f_{xy_2}(\omega) - B_1(\omega)f_{y_1y_2}(\omega) - B_2^*(\omega)f_{xy_2}(\omega) + B_1(\omega)B_2^*(\omega)f_{y_2y_2}(\omega).$$

$f_{\eta_1\eta_2}(\omega)$ is de partiële kruislingse spectraalfunctie van x_t and y_{1t} (in aanwezigheid van y_{2t}) en wordt ook wel weergegeven als $f_{xy_1|y_2}(\omega)$. De partiële coherentie is dan gedefinieerd als de coherentie tussen $\eta_{1,t}$

en $\eta_{2,t}$ en wordt weergegeven door :

$$K_{xy_1|y_2}(\omega) = \frac{f_{\eta_1\eta_2}(\omega)}{\left\{f_{\eta_2\eta_2}(\omega)f_{\eta_1\eta_1}(\omega)\right\}^{1/2}},$$

$$f_{\eta_2\eta_2}(\omega) = f_{y_2y_2}(\omega) \left\{1 - |K_{xy_2}(\omega)|^2\right\}$$

$$f_{\eta_1\eta_1}(\omega) = f_{y_1y_1}(\omega) \left\{1 - |K_{y_2y_2}(\omega)|^2\right\} \quad \text{en dus :}$$

$$K_{xy_1|y_2}(\omega) = \frac{K_{xy_1}(\omega) - K_{xy_2}(\omega)K_{y_2y_1}(\omega)}{\left\{\left(1 - |K_{xy_2}(\omega)|^2\right)\left(1 - |K_{y_2y_2}(\omega)|^2\right)\right\}^{1/2}}. \quad (3.3.31)$$

De functie $K_{xy_1|y_2}(\omega)$ meet de coherentie tussen x_t en y_{1t} nadat voor de invloed van y_{2t} is gecorrigeerd.

Het partiële fase - spectrum is op een gelijke wijze gedefinieerd als : $\varphi_{xy_1|y_2}(\omega) = \arg\left\{f_{\eta_1\eta_2}(\omega)\right\}$ (3.3.32)

en meet de fase tussen x_t en y_{1t} nadat voor de invloed van y_{2t} is gecorrigeerd.

3.4: De spectraalschatter

Deze paragraaf handelt louter over schatters voor de spectraalfuncties en de bijbehorende statistische eigenschappen als convergentie, zuiverheid en consistentie. De schatters voor alle in het vorige hoofdstuk gedefinieerde maatstaven zijn wiskundige afleidingen van de vier multivariate spectraalschatters. Het startpunt is de meest voor de hand liggende schatter voor discrete tijdreeksen en wordt het periodogram genoemd. Deze schatter blijkt echter statistisch slechte eigenschappen te bezitten, die verholpen worden door een wegingsfactor in de schatter te integreren, waarmee consistentie afgedwongen wordt. Een dergelijke schatter staat bekend als een zogenaamde kernschatter voor de covariantie over een fractie van het aantal waarnemingen.

Omdat de interesse in deze studie uitgaat naar puntschattingen voor de spectraalfuncties worden er geen kansverdelingen en betrouwbaarheidsintervallen voor de schatters berekend.

Voor de volledigheid wordt vermeld dat de kansverdeling van de multivariate spectraalschatter kan worden benaderd door een zogenaamde 'complexe Wishart verdeling'¹⁹.

¹⁹ Priestley (1981) bladzijde 696.

3.4.1: Het periodogram en de impliciete inconsistentie

Van een stationaire tijdreeks $f(t)$ met een eindig aantal observaties is in paragraaf 3.1 geconstateerd dat de coëfficiënten van (3.1.3) en (3.1.4), a_k en b_k , feitelijk de regressiecoëfficiënten zijn, die verkregen worden door $f(t)$ te regresseren op $\sin(\omega_m t)$ en $\cos(\omega_m t)$. Deze regressiecoëfficiënten vormen de basis voor de met (3.4.1) gedefinieerde initiële schatter, het periodogram²⁰. Deze initiële schatter is met geschatte covarianties volgens (3.4.3) het eindige steekproefequivalent van het met (3.2.6) gedefinieerde spectrum. In essentie is het periodogram de in paragraaf 3.2 afgeleide definitie van het spectrum, waarbij de theoretische covarianties zijn vervangen door de geschatte covarianties en is daarmee een lineaire (eventueel complexe) combinatie van deze steekproefcovarianties. Als het steekproefequivalent van de spectraalfunctie is het periodogram de meest voor de hand liggende schatter ervan. De schatter poogt echter een continue functie te schatten in plaats van een eindige verzameling parameters en het gevolg is dat de schatter eigenaardige statistische eigenschappen blijkt te bezitten. In de volgende paragraaf wordt het periodogram aangepast om voor de eigenaardige statistische eigenschappen te corrigeren. Onderstaand volgen eerst de definitie en interpretatie van het periodogram.

Definitie 3.4.1: Het periodogram

$$I_n(\omega_k) = \frac{n}{2} (a_k^2 + b_k^2) \text{ is het periodogram, waarbij} \quad (3.4.1)$$

$k = 1, 2, \dots, m$ en m het kleinste gehele getal $\leq (n-1)/2$

□

Invullen van de oorspronkelijke observaties in de definitie van het periodogram geeft:

$$\begin{aligned} I_n(\omega_k) &= \frac{2}{n} \left[\left(\sum_{t=1}^n x_t \cos(\omega_k t) \right)^2 + \left(\sum_{t=1}^n x_t \sin(\omega_k t) \right)^2 \right] \stackrel{\text{Th A.3.2.9}}{=} \\ &= \frac{2}{n} \left[\left(\sum_{t=1}^n (x_t - \mu) \cos(\omega_k t) \right)^2 + \left(\sum_{t=1}^n (x_t - \mu) \sin(\omega_k t) \right)^2 \right] = \\ &= \frac{2}{n} \left[\sum_{t=1}^n \sum_{j=1}^n (x_t - \mu)(x_j - \mu) \cos(\omega_k t) \cos(\omega_k j) + \sum_{t=1}^n \sum_{j=1}^n (x_t - \mu)(x_j - \mu) \sin(\omega_k t) \sin(\omega_k j) \right] = \\ &= \frac{2}{n} \left[\sum_{t=1}^n \sum_{j=1}^n (x_t - \mu)(x_j - \mu) \cos(\omega_k (t-j)) \right], \end{aligned}$$

voor $k=1, 2, \dots, m$. In deze dubbelsom zitten n combinaties met $(t-j)=0$, $(n-1)$ met $(t-j)=1$, etcetera. Door te stellen $p=t-j$ geldt:

$$I_n(\omega_k) = \begin{cases} 2n\bar{x}_n^2, & k=0 \\ 4n\tilde{\gamma}(\omega_k), & k=1, 2, \dots, m \end{cases} \quad (3.4.2)$$

$$\text{waarbij } \tilde{\gamma}(\omega) = \frac{1}{2\pi} \sum_{p=-(n-1)}^{(n-1)} \frac{n-|p|}{n} \hat{\gamma}(p) \cos(\omega p), \quad (-\pi \leq \omega \leq \pi) \quad (3.4.3)$$

$$\hat{\gamma}(p) = \hat{\gamma}(-p) = \begin{cases} \frac{1}{n-p} \sum_{j=1}^{n-p} (x_j - \bar{x}_n)(x_{j+p} - \bar{x}_n), & 0 \leq p \leq n-1 \\ 0, & p > n-1 \end{cases}$$

$$\bar{x}_n = \frac{1}{n} \sum_{j=1}^n x_j.$$

²⁰ de coëfficiënten a_t en b_t zijn functies van t en het periodogram is als functie van a_t en b_t impliciet ook een functie van t . Een alternatieve interpretatie is $t=\omega_n/2\pi$, waarbij t een periode voorstelt en hierdoor komt deze functie aan zijn naam.

Voor de kruislingse covarianties geldt:

$$\hat{\gamma}_{ij}(p) = \begin{cases} \frac{1}{n-p} \sum_{j=1}^{n-p} (x_{j,t} - \bar{x}_n)(y_{j,t+p} - \bar{y}_n) & 0 \leq p \leq n-1 \\ 0, & p > n-1 \end{cases}$$

$$\hat{\gamma}_{ij}(p) = \begin{cases} \frac{1}{n+p} \sum_{j=(t-p)}^n (x_{j,t} - \bar{x}_n)(y_{j,t+p} - \bar{y}_n) & 1-n \leq p \leq 0 \\ 0. & p < 1-n \end{cases}$$

Het geschatte spectrum (3.4.3) is een continue functie van ω en gedefinieerd voor alle ω . Het periodogram is slechts gedefinieerd op de volgende discrete verzameling punten:

$$\omega_k = 2\pi k / n, \quad k = 0, 1, 2, \dots, m$$

Voor het onderzoek naar de limiet eigenschappen van het periodogram is het geschikter een functie te definiëren voor iedere $\omega \in [-\pi, \pi]$. Daartoe wordt gedefinieerd:

$$K(n, \omega) = k \quad \text{voor} \quad \frac{\pi(2k-1)}{n} < \omega < \frac{\pi(2k+1)}{n} \quad k = 0, \pm 1, \pm 2, \dots$$

$$\text{en neem } I_n(\omega) = I_n(\omega_{K(n, \omega)})$$

Voor $\omega \in [0, \pi]$ is $I_n(\omega)$ een trapfunctie met de waarde $I_n(\omega_k)$ op het interval $\left(\frac{\pi(2k-1)}{n}, \frac{\pi(2k+1)}{n}\right)$.

Zoals in de inleiding op dit hoofdstuk en door de titel van deze paragraaf al is gesuggereerd is het periodogram een niet consistente schatter. Deze statistische eigenschap wordt afgeleid nadat in theorema 3.4.2 is gesteld dat de schatter echter wel zuiver is.

Theorema 3.4.2: De zuiverheid van het periodogram

Veronderstel x_t is een stationaire tijdreeks, waarbij $E(x_t) = \mu$ en die een absoluut sommeerbare covariantiereeks kent. Dan geldt :

$$\lim_{n \rightarrow \infty} E\{I_n(\omega)\} = 4\pi f(\omega), \quad \omega \neq 0$$

$$\lim_{n \rightarrow \infty} E\{I_n(0) - 2n\mu^2\} = 4\pi f(0), \quad \omega = 0.$$

Bewijs:

Omdat $E\{\hat{\gamma}(p)\} = \gamma(p)$ volgt :

$$E\{I_n(0)\} = 2 \sum_{h=-(n-1)}^{(n-1)} \frac{n-|h|}{n} \gamma(h) + 2n\mu^2,$$

$$E\{I_n(\omega_k)\} = 2 \sum_{h=-(n-1)}^{(n-1)} \frac{n-|h|}{n} \gamma(h) \cos(\omega_k(h)), \quad k=1, 2, \dots, m$$

Nu geldt : $\left| \sum_{h=-n}^n \frac{|h|}{n} \gamma(h) \cos(\omega_k(h)) \right| \leq \sum_{h=-n}^n \frac{|h|}{n} |\gamma(h)| \xrightarrow{n \rightarrow \infty} 0$, (Kronecker's lemma, appendix A.3.1.12)

$$g_n(\omega) = 2 \sum_{h=-(n-1)}^{(n-1)} \gamma(h) \cos(\omega h) \text{ convergeert uniform naar } 4\pi f(\omega) \text{ (theorema 3.1.8)}$$

en $\omega_{K(n, \omega)}$ convergeert naar ω door de constructie.

□

Uit de stelling volgt dat het periodogram een asymptotisch zuivere schatter is. Het periodogram is statistisch gezien echter een slechte schatter voor de spectraalfunctie. De reden hiervoor is tweeledig:

- $I_n(\omega)$ is geen consistente schatter van $f(\omega)$ in de zin dat $\text{Var}\{I_n(\omega)\}$ niet naar nul gaat als $n \rightarrow \infty$,
- Zijnde een functie van ω heeft $I_n(\omega)$ een grillige en erg fluctuerende vorm, vergelijk figuur 3.5. Dit verschijnsel wordt veroorzaakt door het feit dat voor elke twee dichtbij elkaar liggende frequenties ω_i en ω_r geldt: $\text{Cov}\{I_n(\omega_i), I_n(\omega_r)\} = 0$ als $n \rightarrow \infty$.

Het bewijs van bovenstaande twee beweringen wordt geleverd voor een ruime klasse van tijdreeksen, die gekenmerkt worden door het feit dat de tijdreeksvariabele gerepresenteerd kan worden als een oneindige gewogen som van normaal verdeelde stochasten. Deze klasse van tijdreeksen dekt als gevolg van de Wold-decompositie alle stationaire tijdreeksen. Voor de bewijsvoering over inconsistentie van het periodogram, consistentie en zuiverheid van de kernelschatter worden additieve restricties met betrekking tot de sommeerbaarheid van de covarianties opgelegd aan tijdreeksen. In theorema 3.4.2., dat de inconsistentie van het periodogram aantoont, wordt een restrictie opgelegd aan de absolute sommeerbaarheid van de weegcoëfficiënten van de innovaties. Deze restrictie impliceert een restrictie ten aanzien van de sommeerbaarheid van de covarianties en deze implicatie wordt aangetoond in appendix A.3.4.6. De bewijsvoering omtrent de statistische eigenschappen van kernelschatter in de volgende paragraaf veronderstelt sterkere restricties ten aanzien van de sommeerbaarheid van de covarianties.

Echter, voor een grote klasse van tijdreeksen geldt het volgende theorema, dat komt uit Fuller (1996) en waarvan het bewijs staat gegeven in appendix A.3.4.3:

Theorema 3.4.3: Niet convergerende variantie van het periodogram

Veronderstel een tijdreeks x_t , waarbij: $x_t = \sum_{j=0}^{\infty} a_j e_{t-j}$ met $e_t \sim N(0, \sigma^2)$. o.i.v

en vierde moment $\eta\sigma^4$. Bovendien $\sum_{j=1}^{\infty} j^{1/2} |a_j| < \infty$.

Dan geldt :

$$Cov\{I_n(\omega_j), I_n(\omega_k)\} = \begin{cases} 2(4\pi)^2 f^2(0) + O(n^{-1}) & \omega_j = \omega_k = 0 \\ (4\pi)^2 f^2(\omega_k) + O(n^{-1}) & \omega_j = \omega_k \neq 0, \pi \\ O(n^{-1}) & \omega_j \neq \omega_k \end{cases}$$

□

Theorema 3.4.3. stelt hiermee beide beweringen dat ten eerste de variantie niet naar nul convergeert (geval als $\omega_j = \omega_k$) en de covariantie wel ($\omega_j \neq \omega_k$), waarmee het FDDDDfluctuerende gedrag van het periodogram verklaard wordt.

3.4.2: De kernelschatter: een consistente gewogen integraal van het periodogram

In deze subparagraaf wordt in eerste instantie gepoogd een interpretatie te geven van het inconsistentiefenomeen van theorema 3.4.3 en vervolgens een heuristische uitleg over de oplossing, die vervolgens mathematisch aangetoond wordt.

Voor grote n is het periodogram gedeeld door 4π dezelfde functie van de geschatte covarianties als dat het spectrum de functie is van de theoretische covarianties en het is bekend dat de geschatte covarianties consistent zijn, dat wil zeggen zuiver met een naar nul convergerende variantie, zie appendix A.3.4.4. De paradox is dan dat verwacht zou worden dat een lineaire combinatie van de covariantieschatters een consistente schatter is voor dezelfde lineaire combinatie van theoretische covarianties.

De reden dat $var\{I_n(\omega)\}$ niet naar nul convergeert is gelegen in het feit dat (3.4.1) de som is van n steekproefcovarianties en hoewel elk een variantie kent van $O(1/n)$ geeft het cumulatieve effect van n termen een variantie die $O(1)$ is.

Een reductie van de variantie kan bewerkstelligd worden door simpelweg een gedeelte van de termen in (3.4.3) weg te laten en slechts de eerste m steekproefcovarianties op te nemen. De variantie van deze nieuw geconstrueerde schatter, de afknotschatter, gaat naar nul als $O(m/n)$. Hiermee wordt echter wel een onzuiverheid in de schatter geconstrueerd. In deze studie wordt gewerkt met stationaire processen, die een absoluut sommeerbare covariantiereeks kent. Dit impliceert dat $\gamma(s) \rightarrow 0$ als $|s| \rightarrow \infty$ en indien alleen de termen in de staart van de steekproefcovariantie reeks worden weggelaten wordt de zuiverheid niet te veel aangetast.

Een iets algemenere schatter is de kernelschatter, waarbij de toegelaten steekproefcovarianties gewogen gesommeerd worden. De afknotschatter is dan een toepassing van deze kernelschatter met een weegfunctie, die waarde één heeft voor iedere toegelaten steekproefcovariantie. In deze studie worden er aan de weegfunctie een aantal eisen opgelegd, zoals dat deze functie een even functie is en dat deze de waarde één genereert op tijdstip nul. Een en ander wordt als volgt mathematisch geformuleerd:

$$\text{De kernelschatter is gedefinieerd als: } \hat{f}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \lambda(k) \hat{\gamma}(k) e^{-i\omega k}, \quad (3.4.4)$$

$$\text{waarbij de weegfunctie gelijk is aan: } \lambda(s) = \begin{cases} k\left(\frac{s}{m}\right) & |s| \leq m \\ 0, & \text{elders} \end{cases}, \quad (3.4.5)$$

die voldoet aan: $k\left(\frac{s}{m}\right) = k\left(-\frac{s}{m}\right)$ en $k(0) = 1$.

De centrale bewering in deze paragraaf is dat de kernelschatter een consistente schatter is. De te bewijzen bewering wordt als volgt gesteld:

Onder de voorwaarde dat $m \rightarrow \infty$ als $n \rightarrow \infty$, maar langzaam genoeg, zodanig dat $\left(\frac{m}{n}\right) \rightarrow 0$ als $n \rightarrow \infty$,

dat wil zeggen als $m = n^\alpha$ $0 < \alpha < 1$, dan geldt dat zowel de onzuiverheid van $\hat{f}(\omega)$ verdwijnt als de variantie van $\hat{f}(\omega)$ naar nul gaat voor $n \rightarrow \infty$. Dat betekent dat $\hat{f}(\omega) \forall \omega$ een consistente schatter is van $f(\omega)$, ofwel $\hat{f}(\omega) \xrightarrow{n \rightarrow \infty} f(\omega)$.

In theorema 3.4.1. is gesteld dat het periodogram een zuivere schatter is voor de spectraalfunctie. De kernelschatter van deze paragraaf beschouwt slechts een aantal van m steekproefcovarianties en reduceert daarmee de variantie van deze schatter ten laste van de introductie van onzuiverheid. Deze onzuiverheid²¹ wordt met op de volgende manier gekwantificeerd:

Met $E\{\hat{\gamma}_k\} = \gamma_k$ geldt:

$$b(\omega, m) = E[\hat{f}(\omega)] - f(\omega) \stackrel{(3.4.4)+(3.4.5)}{=} \frac{1}{2\pi} \sum_{s=-(n-1)}^{(n-1)} \left\{ k\left(\frac{s}{m}\right) - 1 \right\} \gamma(s) e^{-ias} - \frac{1}{2\pi} \sum_{|s| \geq n} \gamma(s) e^{-ias}. \quad (3.4.6)$$

Met $k(x)$ van (3.4.5), neem als $r > 0$ het grootste gehele getal waarvoor $\exists k > 0$ zodanig dat

$$\lim_{x \rightarrow 0} \frac{1 - k(x)}{|x|^q} = k^{(r)}$$
 bestaat, eindig en ongelijk nul is.

Veronderstel bovendien dat $\sum_{s=-\infty}^{\infty} |s|^q \gamma(s) < \infty$ voor $q \leq r$ en dat $m \rightarrow \infty$ en $\left(\frac{m^r}{n}\right) \rightarrow 0$ als $n \rightarrow \infty$. Dan gaat

de tweede term van (3.4.6) naar nul, terwijl de eerste term asymptotisch gelijk is aan:

$$-k^{(r)} m^{-r} f^{(r)}(\omega), \text{ waarbij } f^{(r)}(\omega) = \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} |s|^r \gamma(s) e^{-ias}, \text{ ofwel: } \lim_{n \rightarrow \infty} \{m^r b(\omega)\} = -k^{(r)} f^{(r)}(\omega).$$

Een asymptotische expressie voor de onzuiverheid²² is dan: $b(\omega, m) \sim -m^{-r} k^{(r)} f^{(r)}(\omega)$

Het bewijs voor de convergentie van de variantie van de kernelschatter is technisch veelomvattend en verwezen wordt naar Anderson (1971). Op basis van de opgebouwde kennis over (on)zuiverheid van de raamschatters kan wel op een relatief eenvoudige wijze convergentie in waarschijnlijkheid aangetoond worden. De restrictie van absolute sommeerbaarheid van de covariantiereeks wordt wel weer iets aangescherpt. De redenering gaat als volgt:

²¹ de Engelse term voor onzuiverheid is bias, waarmee de naamkeuze van de functie b is.

²² Met (3.4.8) wordt de Tuckey-Hanning weegfunctie geïntroduceerd, waarbij $k(x) = 1 - 2\alpha + 2\alpha \cos(\pi x)$. Nu geldt:

$$\lim_{x \rightarrow 0} \frac{1 - k(x)}{|x|^r} = \lim_{x \rightarrow 0} \frac{2\alpha - 2\alpha \cos(\pi x)}{|x|^r} \stackrel{\text{Hôpital}}{=} \lim_{x \rightarrow 0} \frac{2\alpha \pi \sin(\pi x)}{r|x|^{r-1}} \stackrel{\text{Hôpital en } r=2}{=} \lim_{x \rightarrow 0} \frac{2\alpha \pi^2 \cos(\pi x)}{r(r-1)|x|^{r-2}} = 2\alpha \pi^2 \text{ en } r = 2.$$

$\hat{f}(\omega)$ is een kernel schatter, waarbij voor de weegfunctie $k(u)$ geldt :

$$\int_{-\infty}^{\infty} k(u) du < \infty, \quad m = O\left(n^{1/2}\right) \rightarrow \infty, \quad \sum_{s=-\infty}^{\infty} |s|^2 \gamma(s) < \infty.$$

Beschouw het proces $x_t = \sum_{j=-\infty}^{\infty} a_j e_{t-j}$ met $e_t \sim N(0, \sigma^2)$ stochastische variabelen en de reeks $\{a_j\}$ absoluut

sommeerbaar. Consistentie betekent : $\forall \varepsilon > 0$ geldt $P\left\{|\hat{f}(\omega) - f(\omega)| < \varepsilon\right\} \rightarrow 1$ als $n \rightarrow \infty$ (3.4.7)

$$\text{Er geldt } \hat{f}(\omega) - f(\omega) = [\hat{f}(\omega) - E\{\hat{f}(\omega)\}] + [E\{\hat{f}(\omega)\} - f(\omega)]$$

De tweede term gaat naar nul wegens eerder aangetoonde asymptotische zuiverheid.

$$\text{Voor de eerste term geldt : } \max_{\omega} |\hat{f}(\omega) - E\{\hat{f}(\omega)\}| \leq \sum_{s=-(n-1)}^{(n-1)} \left|k\left(\frac{s}{m}\right)\right| \gamma(s) - E\{\gamma(s)\}.$$

De bovenstaande voorwaarde op $\gamma(s)$ betekent dat $\gamma(s) = O(|s|^{-3})$ en dus kwadratisch sommeerbaar.

Met behulp van appendix A.3.4.4, die stelt dat $\text{var}\{\gamma(s)\} \ll K/n$ waarbij K onafhankelijk van s . Dan geldt

$$\max_{\omega} |\hat{f}(\omega) - E\{\hat{f}(\omega)\}| \leq \sqrt{\frac{K}{n}} \sum_{s=-(n-1)}^{(n-1)} \left|k\left(\frac{s}{m}\right)\right|. \text{ Als geldt } m \rightarrow \infty \text{ dan :}$$

$$\left\{\frac{1}{m} \sum_s k\left(\frac{s}{m}\right)\right\} \rightarrow \int_{-\infty}^{\infty} k(u) du = O(1) \text{ en dus, gegeven dat } \left(\frac{m}{n}\right) \rightarrow 0, E\left\{\max_{\omega} |\hat{f}(\omega) - E\{\hat{f}(\omega)\}|\right\} \rightarrow 0.$$

Nu is aangetoond dat de kernelschatter (3.4.4) een statistisch goede schatter is rest voor een nadere interpretatie nog het afleiden van een alternatieve representatie van deze schatter als een gewogen integraal van het periodogram: Met (3.4.3) en Kronecker's lemma, appendix A.3.1.12, kan het periodogram beschreven worden als:

$$\tilde{f}(\omega) = \frac{1}{2\pi} \sum_{s=-(n-1)}^{(n-1)} \gamma(s) e^{-i\omega s} \text{ en is in Fourier - geïnvverteerde vorm, voor } |s| \leq n-1: \gamma(s) = \int_{-\pi}^{\pi} \tilde{f}(\omega) e^{is\vartheta} d\vartheta.$$

Met substitutie in de kernelschatter (3.4.4) volgt :

$$\hat{f}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \left(\int_{-\pi}^{\pi} \tilde{f}(\omega) e^{ik\vartheta} d\vartheta \right) \lambda(k) e^{-i\omega k} = \int_{-\pi}^{\pi} \tilde{f}(\omega) \left\{ \frac{1}{2\pi} \sum_{s=-(n-1)}^{(n-1)} \lambda(s) e^{-is(\omega-\vartheta)} \right\} d\vartheta =$$

$$\hat{f}(\omega) = \int_{-\pi}^{\pi} \tilde{f}(\omega) W(\omega - \vartheta) d\vartheta, \quad (3.4.8)$$

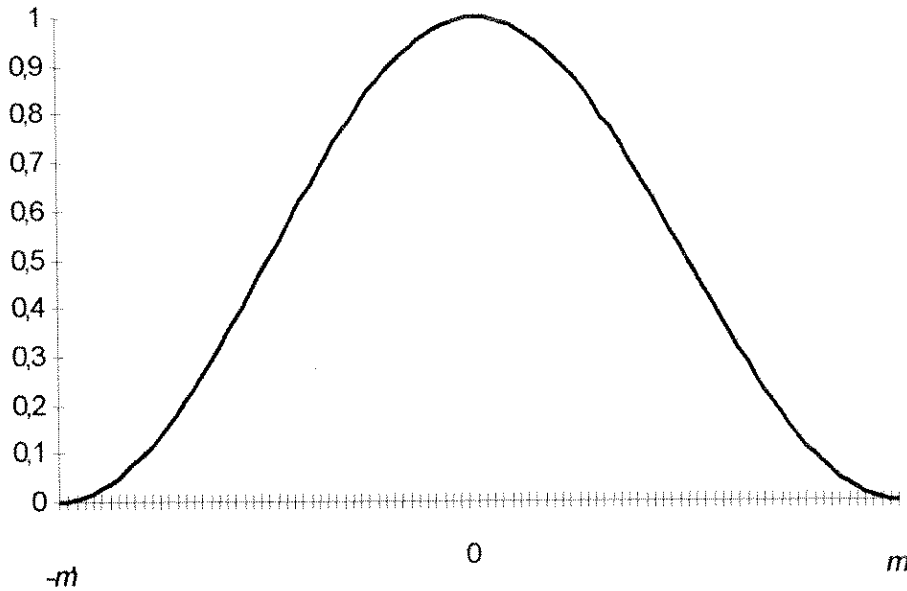
$$\text{waarbij het Fourier - transformatiepaar is } W(\vartheta) = \frac{1}{2\pi} \sum_{s=-(n-1)}^{(n-1)} \lambda(s) e^{-is\vartheta}$$

$$\text{en } \lambda(s) = \int_{-\pi}^{\pi} e^{is(\omega-\vartheta)} W(\omega - \vartheta) d\vartheta = \int_{-\pi}^{\pi} e^{is\varphi} W(\varphi) d\varphi, \quad s = \{-(n-1), \dots, (n-1)\}.$$

De notie om de geschatte autocovariantie gewogen te middelen om daarmee de bijdrage van de staart te reduceren geeft exact hetzelfde resultaat als het gladstrijken van het periodogram met een gewogen integraal van de vorm (3.4.8).

Tot slot wordt er een toepassing gepresenteerd van de algemene kernelschatter. In de literatuur zijn verschillende kernels geopperd, die voldoen aan de bij (3.4.4) gestelde eisen. Deze voorwaarden voor de weegfunctie zijn een waarde van één op tijdstip nul en symmetrie rond dit tijdstip. De meest gebruikte weegfunctie hebben in de literatuur de naam gekregen van hun uitvinders, zoals: Bartlett, Daniell, General Tuckey, Parzen, Priestly, Tuckey-Hanning en Tuckey Hamming. In deze studie zal gebruikt gemaakt worden van Tuckey-Hanning, een getransformeerde Dirchilett kernel, die in figuur 3.14 grafisch is weergegeven.

Figuur 3.14: De Tuckey-Hanning kernel



De Tuckey-Hanning weegfunctie is een specificatie van (3.4.5) met de volgende karakteristieken:

$$\lambda(s) = \begin{cases} 1 - 2a + 2a \cos\left(\frac{\pi s}{M}\right), & |s| \leq m \\ 0, & |s| > m \end{cases} \quad (3.4.9)$$

$0 < a \leq 0.25$, such that $\lambda(s) \geq 0 \quad \forall s$

als $a = 0.25$ wordt dit de Tuckey - Hanning kernel genoemd.

De functie D_m wordt Dirchilet kernel genoemd. Met appendix A.3.5.2 over cosinusommatie:

$$D_m := \frac{1}{2\pi} \sum_{s=-m}^m e^{-is\vartheta} = \frac{1}{2\pi} \sum_{s=-m}^m \cos(s\vartheta) = \frac{1}{2\pi} \left\{ \frac{\sin\left(m + \frac{1}{2}\right)\vartheta}{\sin\left(\frac{\vartheta}{2}\right)} \right\}$$

Met appendix A.3.5.3 en (5.2.5):

$$W(\vartheta) = \frac{1}{2\pi} \sum_{s=-m}^m \lambda(s) e^{-is\vartheta} = \frac{1}{2\pi} \sum_{s=-m}^m \left\{ 1 - 2a + a(e^{i\pi s/m} + e^{-i\pi s/m}) \right\} e^{-is\vartheta} =$$

$$aD_m\left(\vartheta - \frac{\pi}{m}\right) + (1 - 2a)D_m(\vartheta) + aD_m\left(\vartheta + \frac{\pi}{m}\right) \quad (3.4.10)$$

Als (3.4.10) in (3.4.8) gesubstitueerd wordt blijkt dat Tuckey's schatting van de spectraalfunctie geschreven kan worden als een lineaire combinatie van de afgekapte periodogram schattingen op bepaalde frequenties, namelijk:

$$\hat{f}(\omega) = a\tilde{f}_0\left(\omega - \frac{\pi}{m}\right) + (1 - 2a)\tilde{f}_0(\omega) + a\tilde{f}_0\left(\omega + \frac{\pi}{m}\right), \text{ waarbij } \tilde{f}_0(\omega) = \frac{1}{2\pi} \sum_{s=-m}^m \gamma(s) e^{-i\omega s} \quad (3.4.11)$$

Hoewel het periodogram de natuurlijke basisschatter van het spectrum vormt, is deze schatter echter statistisch gezien een slechte schatter voor de spectraalfunctie, omdat de schattingen een grillige en erg fluctuerende vorm kennen, zoals blijkt uit figuur 3.13. De grilligheid van deze schattingen wordt veroorzaakt doordat de variantie van de schatter niet nul convergeert en de fluctuatie wordt veroorzaakt doordat de covariantieschatter op de afzonderlijke discrete frequenties wel naar nul convergeert, waardoor de schattingen behorende bij nabije frequenties geen (cor)relatie kennen. Consistentie van het

periodogram wordt afgedwongen door het geschatte periodogram glad te strijken met een gewogen integraal. Deze bewerking blijkt het spectraalequivalent te zijn van een kernelschatter, die een weegfunctie introduceert en ver weg gelegen covarianties niet meeweegt. De intuïtieve uitleg is dat ver weg gelegen covarianties ongeveer gelijk aan nul zijn voor stationaire tijdreeksen, omdat stationariteit grosso modo een absoluut sommeerbare covariantiereeks impliceert. Een bepaald type kernelschatter is die met een Tuckey-Hanning weegfunctie en deze wordt gebruikt in het MATLAB-programma, waarmee de spectraalschattingen berekend worden en welke is weergegeven in appendix C.

Hoofdstuk 4: De conjunctuur voorspeld

Met de tijdreeksdecompositie en de spectraalanalyse van respectievelijk de theoretische hoofdstukken 2 en 3 is zowel een methodologie als het gereedschap beschikbaar voor de constructie van een huizenprijstvoorspeller. Het aanvangspunt is de decompositie van de huizenprijs in vier componenten, te weten trendmatig, cyclisch, seizoen en transitair. De focus is in hoofdstuk 2 gelegd op twee typen van trendmatige bewegingen, te weten de deterministische en de stochastische. De ontwikkeling van de huizenprijstvoorspeller verloopt langs de twee lijnen, die worden bepaald door de keuze van het type trendcomponent. Een veronderstelde stochastische trend wordt uit een tijdreeks gefilterd door de differentiereeks te bepalen en levert een conjunctuurmaat op, die statistisch gezien stationair is. De stationaire conjunctuurreeks is hierdoor geschikt voor toepassing van spectraalanalyse. Het wegfilteren uit een tijdreeks van een veronderstelde deterministische trend levert een conjunctuurcomponent op, die niet voldoet aan de statistische stationariteitseigenschap, maar wel door economische of marktspecifieke ontwikkelingen te interpreteren is. De analyse met deterministische trend staat in de traditie van de barometrische conjunctuurindicatoren voor de algemeen economische bedrijvigheid. Nadat in paragraaf 4.1. de initiële samenstelling en beschrijving van de data is gegeven wordt in paragraaf 4.2. de conjunctuur van de tijdreeksen bepaald aan de hand van beide typen trendvormen. De relevante methoden uit de barometrische traditie zullen worden toegelicht. In paragraaf 4.3 worden de barometrische en de spectraalvoorspeller voor de huizenprijconjunctuur geconstrueerd op basis van een lineaire leidende cyclische indicatorenmodel. In deze paragraaf komen ook de praktische complicaties van het toepassen van spectraalanalyse aan bod.

4.1: Beschrijving data

Voordat het mogelijk is om leidende indicatoren voor de Nederlandse huizenprijs te selecteren wordt er een verzameling tijdreeksvariabelen bepaald die potentieel invloed uitoefenen op de prijsvorming van de huizen. De initiële samenstelling van deze verzameling sluit aan bij de bestaande ideeën over het functioneren van de huizenmarkt.

In een studie van Van Assenbergh (1999) wordt vooral vanuit een macro-economische perspectief verklaringen gegeven voor de te onderscheiden ontwikkelingen op de Nederlandse woning- en hypotheekmarkt, inclusief een balans van de bestaande en de te verwachten onevenwichtigheden. De ontwikkelingen op de koopwoningenmarkt worden beschreven aan de hand van een drietal interacties, te weten de vraag- en aanbodfactoren en de door overheidsbeleid gevormde institutionele omgeving. Het overheidsbeleid komt tot uitdrukking in het volkshuisvestingsbeleid (sociale woningbouw en vaststelling van de huurprijs), ruimtelijke ordening (toewijzing van de VINEX²³-locaties) en fiscaliteit (waarbij de hypotheekrenteaftrek en het woonwaardeforfait de bekendste beleidsuitingen zijn). De vraagontwikkeling wordt ontleed in demografie (aantal en samenstelling huishoudens) en de financieringsmogelijkheden van potentiële huizenkopers, die afhangen van de rente- en inkomensontwikkelingen. In 1993 is er een forse impuls gegeven aan de financieringsmogelijkheden door de invoering van de tweeverdienersregeling, waarbij het inkomen van de partner ook volledig mag worden meegerekend bij de bepaling van het maximale hypotheekbedrag. De analyse van de aanbodontwikkeling concentreert zich op de bestaande woningvoorraad en de nieuwbouwproductie in kwantitatieve en kwalitatieve zin.

Complementair aan deze studie die zich voornamelijk concentreert op de Nederlandse situatie is het onderzoek van Van Rooij & Vos (1999) over de samenhang tussen Europese huizenprijzen. Zij stellen dat de aankoop van een huis in zekere zin een belegging is en identificeren om die reden sentimentsvariabelen als een expliciete klasse van huizenprijsdeterminanten, inclusief de werkloosheidsvoet en de aandelenkoersen. Voor aandelenkoersen bestaat er echter een meer

²³ VINEX staat voor vierde nota ruimtelijke ordening extra en biedt het kader voor de ruimtelijke ordening tussen 1995 en 2005. Kern van deze nota is dat de nieuwbouwlocaties vooral in het stedelijk gebied dienen te liggen.

structurele samenhang doordat bij een sterke huizenprijsstijging de overwaarde²⁴ kan worden verzilverd en belegd in aandelen, leidend tot hogere aandelenkoersen. Omgekeerd levert een hogere waarde van de effectenportefeuille meer financiële ruimte bij de aankoop van een huis, waarmee de vraag naar huizen en de huizenprijs gestimuleerd worden. De causaliteit van huizenprijzen naar aandelenkoersen is voor een aantal landen, Nederland uitgezonderd, statistische vastgesteld.

Aanvullende op werkloosheidsvoet en de aandelenkoersen worden de lange en korte rente, het beschikbaar inkomen, prijspeil van investeringen in woningen en de huurprijs beschouwd als determinanten van de huizenprijs. Tevens wordt gekeken naar de correlaties van de huizenprijsstijgingen in de afzonderlijke landen. De conclusie luidt dat er een zekere doch beperkte directe internationale samenhang van huizenprijsontwikkelingen bestaat, die min of meer een afspiegeling vormt van een parallelle economische ontwikkeling in de betreffende landen. Van Rooij & Vos formuleren de conclusie als volgt: "Ook al zouden rente- en conjunctuurbewegingen in Europa als gevolg van de monetaire eenwording sterk convergeren, dan nog zal de onderlinge samenhang van huizenprijsontwikkeling beperkt blijven vanwege de specifieke nationale kenmerken van vraag en aanbod op de huizenmarkt en, vooral, het ontbreken van arbeidsmobiliteit over de grenzen heen."

Tabel 4.1: Initiële data

Woningmarkt gerelateerd	Monetair	Economisch
Huizenprijsindex (<i>Januari 1994=100</i>)	Lange rente (<i>10-jaars NL-staatsobligaties</i>)	Kredietproductie Rabobank
# transacties woningmarkt	Korte rente (<i>AIBOR/EURIBOR</i>)	Gdp (<i>beschikbaar inkomen</i>)
Verkooptijd woningen	Hypotheekrente* ⁶	Leading indicator
Marktstemmingsindex (<i>=omzet_t/omzet_{t-1}</i>) <i>omzet=huizenprijsindex*#transacties</i>	Rentemarge (<i>=hypotheekrente -/- lange rente</i>)	AEX-index (<i>beurskoersen</i>)
Krapte-indicator woningmarkt (<i>=aanbod_t/verkooptijd_{t-1}</i>) <i>aanbod=verkooptijd*#transacties</i>	Cpi-index* ⁴ (<i>inflatie</i>)	Grondprijs (<i>landbouwgrond</i>)
# bouwvergunningen* ¹ (<i>bron datastream</i>)	Cpi-joj* ⁴ (<i>groeipercentage</i>)	Industrievertrouwen
# bouwvergunningen* ¹ (<i>bron Rabobank</i>)	M3* ⁵ (<i>geldhoeveelheid</i>)	Zakenvertrouwen
# bouwvergunningen jaartotalen* ¹	M3-joj* ⁵ (<i>groeipercentage</i>)	Persoonlijke besparingen
Huurprijs	Hypotheekrente reeël* ⁶	Werkloosheid
# hypotheeklen* ²	Yield (<i>lange rente -/- korte rente</i>)	Werkgelegenheid
# hypotheeklen Rabobank* ²		Consumentenvertrouwen
Bedrag hypotheeklen* ³		Economisch sentiment
Bedrag hypotheeklen Rabobank* ³		
Ordervoorraad constructie		
Kosten constructie		

Voor de variabelen met superscript *¹ =1,..,6 geldt de restrictie dat er per i 1 variabele mag worden geselecteerd

²⁴ De overwaarde is het verschil tussen de actuele waarde van de woning en de daarop gevestigde hypotheek.

Op basis van bovenstaande studies is een initiële verzameling van aan de nationale huizenprijs gerelateerde tijdreeksvariabelen bepaald en weergegeven in tabel 4.1, onderverdeeld naar hun aard in economisch, monetair of woningmarkt gerelateerd. Deze indeling is niet strict en dient louter om de variabelen overzichtelijk te presenteren. Voor de verzameling variabelen die zijn aangegeven met het superscript ⁱ, $i=1,2,\dots,6$ geldt de restrictie dat er slechts één variabele van geselecteerd mag worden. Zo is er de beschikking over twee tijdreeksen van het aantal bouwvergunningen uit verschillende bronnen en deze reeksen worden geacht substituten van elkaar te zijn in plaats van complementen. De reeksen van de huizenprijsindex en het aantal transacties in de woningmarkt komen van de NVM en het Kadaster. De overige reeksen komen uit Datastream of interne Rabobank-bronnen. De reeksen zijn op basis van maandelijkse cijfers en lopen maximaal van mei 1973 tot september 1999²⁵.

4.2: Conjunctuurbepaling

De huizenprijs is als macro-economische grootheid onderhevig aan de macro-economische conjunctuurecycli. De conjuncturele beweging van de huizenprijs en van de overige tijdreeksen wordt voor een groot gedeelte bepaald door het type trendvorm dat wordt verondersteld. In deze paragraaf wordt de conjunctuur bepaald aan de hand van beide in hoofdstuk 2 beschreven typen trendvormen. Een deterministisch veronderstelde trendcomponent levert een interpreteerbare conjunctuur op in termen van economische en marktspecifieke ontwikkelingen. Deze conjunctuurreeks is echter niet exact eenduidig en bovendien hangt de interpretatie mede af van de toegepaste statistische technieken.

4.2.1: Conjunctuurbepaling met deterministische trend, de barometrische methode

In de traditie van de conjunctuurindicatoren, die een barometrische indicatie geven voor de conjuncturele toestand van de economie, wordt er een indicator voor de cyclische beweging van de huizenprijs geconstrueerd. De op- en neergang van de algemeen economische bedrijvigheid uit zich in min of meer systematische bewegingen van economische tijdreeksen. Een deel van deze systematiek zijn de golfbewegingen, waarvoor een periode van twee á zeven jaar karakteristiek is en volgens Fase & Bikker (1985) conjunctuur genoemd wordt. Voor het analyseren van de conjunctuur bestaan twee methoden van onderzoek: de causale en de barometrische. Voor toepassing van de causale methode kan in het bijzonder gedacht worden aan de macro-econometrische modellen, die de structuur van een economie beschrijven en inzicht geven in de feitelijke oorzaken van de conjuncturele bewegingen. Door de interactieve causale mechanismen is deze methode minder geschikt voor het beschrijven van een eenduidig conjunctuurbelooop. De barometrische methode genereert daarentegen een conjunctuurdiagnose op basis van een niet-causale, mechanische wijze. De technische ontleding van een economische tijdreeks in zijn afzonderlijke componenten²⁶ geschiedt op basis van methoden en algoritmen die geen specifiek economisch theoretische grondslag kennen. Een analyse van het algemeen economische conjunctuurbelooop voor Nederland in deze traditie is uitgevoerd door onder andere Fase & Bikker en Van Duijn (1978).

De doelstelling van deze subparagraaf is om de methodiek van de barometrische traditie, die succesvol is gebleken in een algemeen economische omgeving, toe te passen in een marktspecifieke context met betrekking tot de Nederlandse huizenprijs. Deze marktspecifieke toepassing is ingegeven door de vaststelling dat de Nederlandse huizenmarkt door zijn omvang een macro-economische factor van betekenis is. Volgens de bestaande opvattingen over het functioneren van de woningmarkt, zoals samengevat in paragraaf 4.1, oefenen verscheidene macro-economische grootheden, waaronder nationaal inkomensontwikkeling en rente een belangrijke invloed uit op de huizenmarkt. Vanuit dit oogpunt wordt de methodiek van de barometrische methode toegepast op de huizenmarkt.

4.2.1.1: Trendalgoritme

De deterministische trend is in hoofdstuk 2 gedefinieerd als de structurele beweging van een tijdreeks, die wordt beschreven door een van de tijdsvariabele t afhankende bepaalde functie. Deze definitie

²⁵ De individuele looptijden van de tijdreeksvariabelen staan weergegeven in tabel B.3 in appendix B.

²⁶ De klassieke componenten zijn trendmatig, cyclisch, seizoens en transitoir, zie hiervoor hoofdstuk 2.

reflecteert het verschijnsel dat deze trendvorm een gepredetermineerde ontwikkeling in de loop van de tijd kent, die qua vorm min of meer onafhankelijk is van de onderliggende tijdreeks. Voor het theoretische hoofdstuk 2 is deze definitie geschikt om het conceptuele verschil met een stochastische trendvorm te verduidelijken. In deze praktische subparagraaf wordt er met een algemenere vorm van deterministische trendbeweging gewerkt. De operationele definitie luidt dat een deterministische trend een algoritme is dat de structurele beweging van een tijdreeks bepaald aan de hand van de realisaties van de tijdreeks en (eventueel indirect) van de variabele t op een zodanige wijze dat de tijdreeks grosso modo even vaak een hogere als een lagere waarde dan de trendwaarde heeft.

Het startpunt van de trendanalyse is het gepredetermineerde tijdtrendmodel (2.3.1) inclusief hogere orden van t^n . Voor de tijdreeks van de Nederlandse huizenprijs levert zowel de kleinste kwadraten montage van een derdegraads als van een vierdegraads polynomiale tijdtrend een verklarend vermogen van de reeks, R^2 , op van respectievelijk 0.949 en 0.933. Deze polynomiale vorm van tijdtrendbepaling is feitelijk een kleinste kwadraten fit van een a priori bepaalde vorm over een a priori bepaalde periode. De keuze van een gepredetermineerde vorm wordt op deze manier niet direct door de data bepaald, waardoor er onzekerheid blijft bestaan over het juiste aantal buigpunten en over het feit of een periode tussen twee buigpunten wel representatief is voor een conjunctuurbeweging.

Een veelgebruikte methode die niet gekenmerkt wordt door een gepredetermineerde vorm is het bepalen van een 75-maands voortschrijdend gemiddelde als maat voor de trendmatige beweging. De trend in t_0 wordt bepaald door de realisaties van de tijdreeks uit te middelen voor $t \in [t_0-37, t_0+37]$ van maandelijkse waarnemingen, waarmee dit filter voldoet aan de operationele definitie van een deterministische trend. Het achterliggende idee is volgens Van Duijn (1978) dat de gemiddelde lengte van een conjunctuurcyclus 6 jaar bedraagt en dat het voortschrijdend gemiddelde deze cyclische beweging uitvlakt.

Volgens Bikker & Fase (1985) bestaat er in werkelijkheid grote variëteit in de duur van de golfbeweging, waardoor de met het voortschrijdend gemiddelde verkregen trendschatting resten van conjuncturele patronen bevat. De fasegemiddelden-trend (PAT²⁷-trend) is een niet-parametrische methode en ontworpen om deze onzuiverheid te ontlopen, zonder de flexibiliteit van de trendschatting te reduceren.

De PAT-trend wordt berekend door een niet-parametrisch algoritme dat als volgt globaal verloopt voor een seizoensvrije tijdreeks:

1. de verhouding van de reeks en zijn (instelbaar) voortschrijdend gemiddelde bepaalt de initiële expansie- en contractieperioden, die fasen genoemd worden;
2. bereken de fase-gemiddelden van de originele data over de afwisselende expansie- en contractieperioden (fasen), bepaald volgens 1;
3. bereken de drie-perioden-voortschrijdende gemiddelden over de fasegemiddelden;
4. verbind de voortschrijdende gemiddelden door rechte lijnstukken. Hiermee wordt voor het midden van de periode een trendwaarde vastgesteld;
5. vervolgens wordt het niveau van de trend aangepast aan het niveau van de oorspronkelijke gegevens;
6. bepaal de trend voor begin en eind van de reeks via extrapolatie.

In plaats van het bepalen van een relatief onafhankelijke trendlijn in de originele data, onafhankelijk van enige cyclische beweging, is de trendmatige beweging juist bepaald aan de hand van de onderliggende cyclische beweging. De methode geeft daardoor een goede beschrijving van het conjunctuurgedrag van een reeks in het verleden, maar is minder geschikt voor voorspeldoeleinden. Dit wordt veroorzaakt doordat de procedure stoelt op cyclische omslagpunten en extrapolatie voorbij het laatste omslagpunt lastig is, omdat onbekend is hoe lang de huidige conjunctuurgolf nog blijft

²⁷ PAT is een afkorting van Phase Average Trend. Voor een meer gedetailleerde en technische beschrijving van dit algoritme wordt verwezen naar appendix B. Als referentie dient Boschan & Ebanks (1978).

voortduren. Het laatste stukje trend is daardoor niet volledig bepaald en juist de laatste waarnemingen zijn voor voorspeldoeleinden belangrijk. Boschan & Ebanks formuleren dit als volgt: "As far as the trend is concerned, additional observations, before a new turning point is reached, can change the level of the trend, but usually not by very much. However, if a new turning point is experienced, the entire slope is changed and the last segment of the trend can change drastically."

In figuur 4.2. zijn vier tijdreeksen uitgezet samen met hun PAT-trend (linkeras). In deze figuur zijn tevens gestandaardiseerde reeksen (rechteras) weergegeven, die een maat vormen voor de conjuncturele beweging van deze reeksen. De wijze waarop deze reeksen bepaald worden staat beschreven in de volgende paragraaf.

4.2.1.2: Conjunctuurbepaling en interpretatie

Conjuncturele beweging is in zijn meest algemene vorm het gedeelte van de beweging van een tijdreeks dat niet door de trendmatige, seizoensmatige en ruisachtige beweging bepaald wordt. Conjunctuur is hiermee synoniem aan cyclische beweging en volgens de barometrische methode gelijk aan golfbewegingen, waarvoor een periode van twee á zeven jaar karakteristiek is. De barometrische maatstaf voor conjunctuur wordt op de volgende operationele wijze verkregen. Bepaal de verhouding van een seizoensgecorrigeerde²⁸ tijdreeks en de bijbehorende met het PAT-algoritme gedetermineerde trendlijn. Deze verhouding is een maat voor de expansieve en contractieve perioden van een tijdreeks en representeert daarmee de cyclische beweging en de ruisbeweging. Het verloop van deze multiplicatieve representatie is een maat voor de cyclische beweging. De grilligheid van het verloop is een maat voor de ruisbeweging en wordt weggefilterd met behulp van een voortschrijdend gemiddelde.

Deze maat voor de cyclische beweging hangt zowel af van de mate van de schommeling van een reeks om zijn trend als van de eenheid van de onderliggende tijdreeksvariabele. Zo wordt onder meer in figuur 4.2. dat de kredietproductie van de Rabobank gemeten in guldens, de beide rentes in procentpunten en is de huizenprijs een indexreeks. De korte rente kent een schommeling van een half maal de trend tot tweemaal de trend, terwijl voor de huizenprijs een schommeling geldt van driekwart tot anderhalf maal de trend. De operationele conjunctuurmaat is gedefinieerd als de gestandaardiseerde maat voor de cyclische beweging en voldoet daarmee aan de eis van vergelijkbaarheid tussen afzonderlijke conjuncturele bewegingen. Voor de mate van schommeling en de eenheid van een tijdreeksvariabele wordt dus gecorrigeerd door standaardisatie, waarbij van de cyclische beweging zijn gemiddelde wordt afgetrokken en vervolgens gedeeld door zijn standaarddeviatie. Deze gestandaardiseerde conjunctuurreeksen worden kortweg conjunctuurreeksen genoemd en staan voor de vier reeksen weergegeven in figuur 4.2. (rechteras).

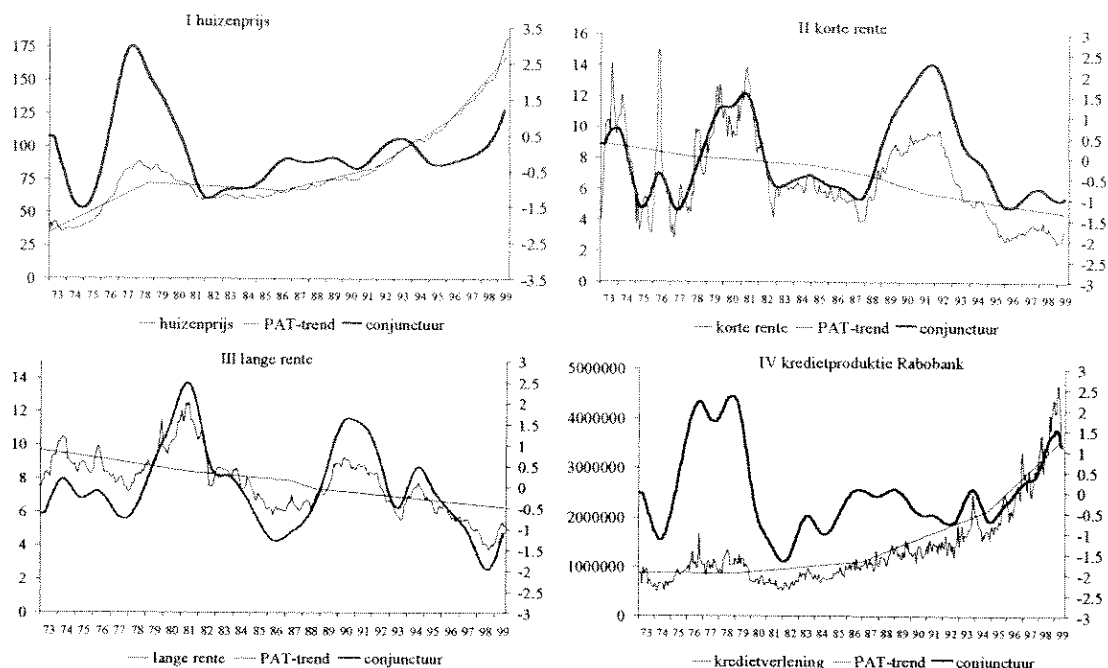
De gestandaardiseerde conjunctuurreeksen geven de beweging weer ten opzichte van de trend, die samenvalt met de nul-as. Een positieve conjunctuurreeks betekent dat de oorspronkelijke tijdreeks een hogere dan trendmatige waarde heeft en wordt hoogconjunctuur genoemd. Een positieve helling van de conjunctuurreeks betekent een hogere dan trendmatige groei in geval de trend stijgend is of een lager dan trendmatige daling in geval de trend dalend is. Als de conjunctuurreeks zich boven de nul-as bevindt is er sprake van hoogconjunctuur. Op het moment dat de conjunctuurreeks een piek passeert, heeft de conjunctuur haar hoogtepunt bereikt. De daarop volgende dalende beweging duidt op een minder dan trendmatige groei en (hoog)conjuncturele afzwakking. Als de conjunctuurreeks de nul-as passeert is er sprake van laagconjunctuur. De richting van de wijzer (richtingscoëfficiënt) is een maat voor de intensiteit van de beweging: hoe hoger die waarde, hoe sneller de groei.

4.2.1.2.1: Huizenprijsconjunctuur

De huizenprijsconjunctuur is de volgens de barometrische methodiek geconstrueerde conjunctuurreeks van de Nederlandse huizenprijs. De barometrische methode stelt met betrekking tot de conjunctuurbeweging van de algemeen economische bedrijvigheid dat conjunctuur zich uit in meerdere tijdreeksen, zoals nationaal inkomen, industriële produktie, werkloosheidsvoet, etc.

²⁸ seizoenscorrectie geschiedt volgens een standaardmethode van SPSS, de zogenaamde Census Method I. Zie Makridakis et al. (1983).

Figuur 4.2: Tijdreeksen met trend en conjunctuur



Een conjunctuurindicator wordt vanuit deze visie geconstrueerd als een gewogen gemiddelde van de conjunctuur van meerdere reeksen, waarmee de verstoringe invloed van tijdreeks specifieke schommelingen minimaal is. De huizenprijsconjunctuur wordt echter slechts bepaald aan de hand van enkel de huizenprijsreeks en wordt daardoor wel beïnvloed door marktspecifieke factoren. Zoals zal worden betoogd in de inleiding van subparagraaf 4.2.3, is de verwachting echter dat huizenmarktspecifieke ontwikkelingen zich (voornamelijk) manifesteren in de structurele trendmatige beweging.

Het PAT-algoritme kent een aantal instelbare paramaters, waarmee het aantal trendomslagpunten gemanipuleerd kan worden en er dus meerdere trendreeksen en conjunctuurreeksen gecreëerd kunnen worden. Exercities met het algoritme²⁹ wijzen uit, zoals gedocumenteerd in appendix B.1.1 en tabel B.1, dat ongeacht de parameterinstellingen en het bijbehorende aantal omslagpunten van de trend er een stabiele verzameling van trendomslagpunten blijkt te bestaan. Het zijn oktober 1975, januari 1979, december 1983, oktober 1986 en januari 1992.

Aanvankelijk zijn er drie verschillende conjunctuurreeksen voor de huizenprijs geconstrueerd, waarvan de beste is weergegeven is figuur 4.2.1. Het kwaliteitsoordeel over de conjunctuurreeksen wordt gevormd aan de hand van de praktische overweging in hoeverre de conjunctuurreeks voorspeld kan worden door vooruitlopende conjunctuurreeksen van andere variabelen. Dit komt nog uitgebreid aan bod in paragraaf 4.3. Een tweede kwaliteitsoordeel is in hoeverre de geconstrueerde conjunctuurreeksen overeenkomen in termen van uitslag en omslagpunten met een conjunctuurreeks voor de huizenprijs die is geconstrueerd middels een andere methode van trendbepaling. In figuur 4.3.1 is een verkorte huizenprijsconjunctuurreeks uitgezet samen met een conjunctuurreeks van de huizenprijs, die is verkregen door een 2^e graads polynomiale trendbepaling, genaamd polynoomreeks. Qua vorm en qua omslagpunten vertoont de polynoomreeks sterke overeenkomsten met de verkorte conjunctuurreeks, alleen qua uitslag bestaan er zijn verschillen. In de volgende subparagraaf wordt een interpretatie gegeven over de uitslagwijdte van de huizenprijsconjunctuurreeks.

4.2.1.2.2: Interpretatie van de invloed van standaardisatie op de huizenprijsconjunctuur

In het verloop van de volledige huizenprijsconjunctuur in figuur 4.2 komt duidelijk de explosie en implosie eind jaren '70 van de huizenprijzen tot uitdrukking, terwijl er voor de laatste waarneming in

²⁹ het PAT-algoritme inclusief conjunctuurbepaling is in Visual Basic geprogrammeerd met toepassing in Excel. De syntax wordt om redenen van compactheid niet opgenomen in deze scriptie. Het programma is echter verkrijgbaar bij de auteur.

1999 geen sprake is van (conjuncturele) oververhitting. De conjunctuurmaat geeft voor deze periode een waarde van ongeveer 1.5, terwijl de waarde eind jaren '70 ruim 3 was. De suggestie dat de huizenprijsconjunctuur destijds twee maal zo hoog was is onjuist als gevolg van een restrictie op de interpretatie van de gestandaardiseerde conjunctuurreeks. De conjunctuurmaat, die is gedefinieerd als de gestandaardiseerde en gladgestreken verhouding van een tijdreeks en de bijbehorende PAT-trend, is afkomstig uit de barometrische traditie en wordt toegepast door onder meer Bikker & Fase (1985) en Van Duijn (1978). Deze conjunctuurmaat is multiplicatief en daarmee alleen stationair indien de schommeling rond de trend verhoudingsgewijs toe- en afnemen als de trend toe- en afneemt. Van de vier reeksen die in figuur 4.2 zijn weergegeven is te zien dat voor beide rentes zowel de trendlijn, als de schommeling rond de trendlijn licht afnemen. De trend van de kredietproductie van de Rabobank daarentegen stijgt sterk, terwijl ook de schommelingen in de loop van de tijd sterker zijn geworden. Uit de grafiek van de huizenprijs blijkt grafisch dat de trend sterk is toegenomen en de schommeling rond de trend juist sterk is gereduceerd. Voor een multiplicatieve conjunctuurmaat betekent een in absolute termen stabiele schommeling rondom een stijgende trendmatige beweging feitelijk dat de gemeten verhoudingsgewijze conjunctuurmaat uitdempt. De multiplicatieve conjunctuurmaat is geschikt voor reeksen die geen sterke op- of neerwaartse trendmatige bewegingen kennen of voor reeksen die een steeds volatielere beweging laten zien als er een hogere trendwaarde bereikt wordt.

De invloed van standaardisatie van een niet-stationaire convergerende reeks, zoals de initiële conjunctuurmaat van de huizenprijs, wordt als volgt geanalyseerd. Zij X_t de tijdreeksvariabele van een multiplicatieve conjunctuurreeks en x_t zijn realisatie. In deze studie wordt met initiële conjunctuurmaat de verhouding van de huizenprijsindexreeks en de (PAT-)trendreeks.

De definitie van standaardisatie is:

$$Z(t, T) = \frac{X_t - \mu(X_1, \dots, X_T)}{\sigma(X_1, \dots, X_T)} \quad (4.2.1)$$

In geval de oorspronkelijke tijdreeks, zoals de huizenprijsreeks, een stabiele absolute schommeling om een stijgende trend kent, geldt voor de initiële conjunctuurreeks de volgende relaties:

$$X_t \xrightarrow{t \rightarrow \infty} \mu \quad (4.2.2)$$

$$\hat{\sigma}(x_1, \dots, x_t) \xrightarrow{t \rightarrow \infty} = 0 \quad (4.2.3)$$

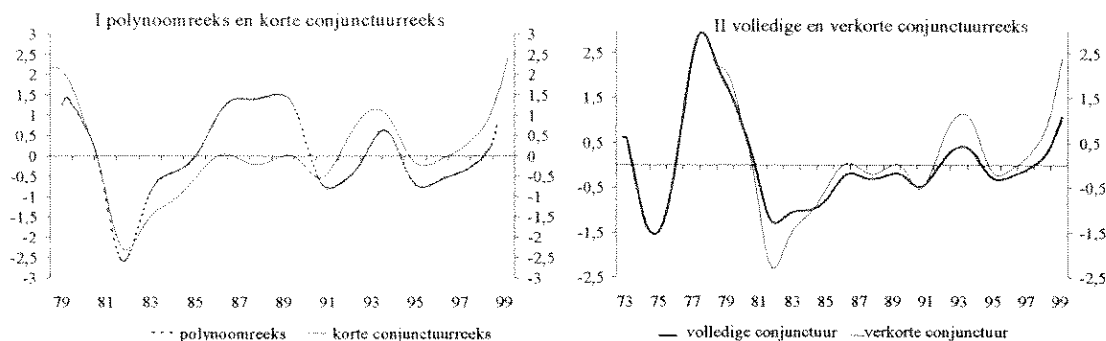
Hierbij is:

- X_t = de te standaardiseren initiële conjunctuurreeks;
- Z_t = de gestandaardiseerde conjunctuurreeks;
- σ = de standaardafwijking van de initiële conjunctuurreeks.

Vergelijking (4.2.3) symboliseert het uitdempfenomeen, dat zich voordoet bij een multiplicatieve conjunctuurmaat indien een tijdreeks een stationaire absolute schommeling rond een opwaartse trend kent. In dat geval is de X_t -reeks immers niet stationair en convergeert naar zijn gemiddelde. De gemiddelde waarde van één waarnaar een multiplicatieve conjunctuurreeks convergeert wordt theoretisch bereikt door de operationele definitie van een deterministische trend, waaronder de PAT-trend, die luidt dat een deterministische trend zich even vaak boven als onder een tijdreeks bevindt.

De interpretatie van de grafische vorm van de conjunctuurmaat, zoals weergegeven in figuur 4.2 en figuur 4.3, dient genuanceerd te worden, omdat de waarden die deze maat genereert een functie zijn van zowel het tijdstip t als van het totaal aantal waarnemingen T , zoals weergegeven in (4.2.1). In het geval dat $T=t$ heeft de geschatte standaarddeviatie door het uitdempfenomeen een hogere waarde dan in het geval $T=2t$, dat wil zeggen: $\hat{\sigma}(x_1, \dots, x_t) > \hat{\sigma}(x_1, \dots, x_{2t})$. Met de definitie van standaardisatie (4.2.1) is in te zien dat met de geschatte standaarddeviaties als gevolg geldt: $Z(t, t) < Z(t, 2t)$. Ondanks dat de initiële conjunctuurrealisatie x_t dezelfde blijft, verandert zijn gestandaardiseerde equivalent $Z(t, T)$ met T en daarmee de grafische beeldvorming omtrent het conjunctuurbeloop. De consequentie van deze mechanische eigenschap is dat de conjuncturele huizenprijsbubble van eind jaren '70 qua uitslag de conjunctuur met het toenemend aantal waarnemingen steeds meer gaat domineren.

Figuur 4.3: Verkorte huizenprijsconjunctuur en polynoomwijzer



Vooruitlopend op paragraaf 4.3. wordt er voor de constructie van de voorspeller van de huizenprijsconjunctuur gewerkt met correlaties en kruislingse correlaties om samenhang en verschuiving van twee tijdreeksen te meten. De kruislingse correlatie is een gestandaardiseerde maatstaf gedefinieerd als de covariantie tussen twee reeksen gedeeld door het produkt van de standaardafwijkingen van beide reeksen. De invloed van het uitdempfenomeen (4.2.3) op de correlatie manifesteert zich zowel in de teller als in de noemer van de correlatiedefinitie, waardoor het geen invloed uitoefent op de correlatie.

De correlatiefunctie corrigeert voor het uitdempfenomeen en blijft daardoor geldig. Voor de grafische presentatie en interpretatie van de conjunctuurmaat is daarentegen voorzichtigheid geboden. De huizenprijsbubble van eind jaren '70 is dominant aanwezig in de conjunctuurreeks en zal, zoals betoogd, steeds dominanter worden naarmate de tijd vordert en er meer waarnemingen beschikbaar komen. Vanwege zijn versturende invloed is de bubble weggelaten en de huizenprijsconjunctuur opnieuw bepaald vanaf het trendomslagpunt van januari 1979. Zowel de volledige als de verkorte huizenprijsconjunctuurreeksen zijn grafisch weergegeven in figuur 4.3.II. De verhouding van de conjunctuuruitslag van eind '99 ten opzichte van die van januari 1979 is door de verkorting drastisch gewijzigd. Voor de volledige conjunctuurreeks geldt een waarde van één eind '99 en van ongeveer twee begin '79, wat een verhouding impliceert van een half. De verkorte conjunctuurmaat geeft voor beide tijdstippen echter een equivalente uitslag wat in een verhouding van één resulteert. Onafhankelijk van de onderliggende conjuncturele verhouding veroorzaakt louter standaardisatie in dit geval een schaling van een factor twee.

Dit voorbeeld illustreert dat de mechanica van standaardisatie voor niet-stationaire convergerende reeksen een interpretatiestoornis op de grafische beeldvorming van conjunctuur veroorzaakt.

4.2.1.2.3: Conjunctuur van de overige reeksen

De barometrische methode van conjunctuurbepaling, bestaande uit het determineren van een PAT-trend, het berekenen van de initiële multiplicatieve cyclische beweging en de finale, gladgestreken en gestandaardiseerde conjunctuurmaat, wordt tevens toegepast op alle reeksen van de dataverzameling van paragraaf 4.1. De ingestelde parameters van het PAT-algoritme en het aantal trendomslagpunten zijn voor alle reeksen uit de verzameling gedocumenteerd in tabel B.3 van appendix B. Voor iedere tijdreeks is de conjunctuur bepaald en drie van deze conjunctuurreeksen zijn samen met de huizenprijsconjunctuurreeks uitgezet in figuur 4.2. De bepaalde deterministische conjunctuurreeksen dienen potentieel als leidende indicator voor de huizenprijsconjunctuur als ze in voldoende mate samenhangen met de huizenprijsconjunctuur en als ze in voldoende mate voorlopen op de huizenprijsconjunctuur. Op basis van geïdentificeerde leidende indicatoren wordt er een lineair voorspelmodel voor de huizenprijsconjunctuur geschat. Het identificeren en schatten zijn onderwerpen van paragraaf 4.3. In de volgende subparagraaf 4.2.2 worden de tijdreeksen vanuit een stochastische uitgangspunt van hun trendmatige beweging ontdaan.

4.2.2: Conjunctuurbepaling met stochastische trend, de differentiemethode

In het theoretische hoofdstuk 2 is de decompositie van een tijdreeks in zijn afzonderlijke componenten beschreven. De trendmatige component kent zowel een deterministische als een stochastische verschijningsvorm, die beide door differentie uit een reeks verwijderd kunnen worden. Voor praktische toepassing wordt ook de stochastische trendvorm algemener omschreven. De operationele definitie van een stochastische trend is de structurele beweging van een tijdreeks, die weggefilterd wordt door een differentiereeks van een zodanige orde dat deze differentiereeks stationair is. Een differentiereeks wordt verkregen door toepassing van de differentiebewerker (2.1.2) en impliciet van de vertragsbewerker (2.1.1) op de oorspronkelijke tijdreeks en wordt stochastische conjunctuurreeks genoemd. Dergelijke conjunctuurreksen zijn hoofdstuk 2 reeds berekend voor onder andere het random walk (plus drift)-model en het kwadratische tijdtrend-model. De bij deze gesimuleerde reeksen behorende differentiereksen zijn weergegeven in figuur 2.1. De stochastische conjunctuurmaat van de huizenprijs wordt gevormd door de tweede orde differentiereeks van de huizenprijs en is tevens weergegeven in figuur 2.1.IV. Uit deze figuur blijkt grafisch dat de stochastische conjunctuurmaat in tegenstelling tot zijn deterministische pendant niet interpreteerbaar is in de zin dat hoog- en laagconjunctuur en conjuncturele beweging niet identificeerbaar zijn. De afwijkingen in de uitslag van de stochastische huizenprijsconjunctuur eind jaren '70 zijn hoogstens een indicatie dat de onderliggende tijdreeks van de huizenprijs in die periode ook afwijkend is geweest.

De orde van de differentiereeks die nodig is om een tijdreeks stationair te maken verschilt per reeks. In paragraaf 2.4 is de Dickey-Fuller toetsingsprocedure beschreven, waarmee kan worden getoetst of een tijdreeks stationair is³⁰. De toetsingsprocedure is uitgevoerd op de initiële verzameling tijdreeksen en als aanvulling is voor een aantal variabelen de seizoensgecorrigeerde tijdreeks bepaald op de volgende wijze: $\Delta_{12}y_t = y_t - y_{t-12}$. Het gaat hier om de volgende variabelen: #*Transacties, Kredietproductie Rabobank, #Hypotheke, #Hypotheke Rabobank, Bedrag hypotheke* en *Bedrag hypotheke Rabobank*. Deze seizoensgecorrigeerde reeksen worden meegenomen als potentiële leidende indicatoren onder de restrictie dat ze niet tegelijk met de bijbehorende seizoensongecorrigeerde tijdreeks geselecteerd mogen worden.

De resultaten van de toetsingsprocedure staan volledig gedocumenteerd in tabel B.5 van appendix B.2.1. Van de oorspronkelijke tijdreeksen die zijn getoetst op stationariteit is de reeks *Zakenvertrouwen* van nature stationair. Voor de overige reeksen wordt de trend door een eerste orde differentie weggefilterd en de resulterende differentiereksen blijken op een aantal na voor alle reeksen stationair te zijn. Voor het aantal overgebleven reeksen is wederom een differentie bepaald en de tweede orde differentiereksen van deze variabelen blijken wel stationair te zijn. De verzameling tweede orde stationaire reeksen zijn: *Huizenprijsindex, Cpi-index, M3 (geldhoeveelheid), Grondprijs (landbouwgrond)* en de *Huurprijs*.

Met het bepalen van de differentiereksen van geschikte orde is voldaan aan de stationariteitseis voor toepassing van de spectraalanalyse en deze methode zal in de volgende paragraaf gebruikt worden om leidende indicatoren te identificeren en een lineair voorspelmodel voor de stochastische huizenprijsconjunctuur te schatten. Zoals beschreven in hoofdstuk 3 wordt in de spectraalanalyse met (3.1.2) een tijdreeksvariabele gerepresenteerd als functie van de frequentie en daarmee beschouwd als een gewogen som van golven met verschillende frequenties. De conjunctuurcomponent van een tijdreeks is in het frequentiedomein dan gelijk aan de opbouwende golven met frequenties die overeenkomen met golflengtes tussen 2 en 7 jaar. Een stochastische conjunctuurreeks is echter een differentiereeks van de oorspronkelijke tijdreeks en de vraag resteert in hoeverre de opbouwende golven van de differentiereeks overeenkomen met de opbouwende golven van de oorspronkelijke reeks. Het zijn niet de golfbewegingen van de oorspronkelijke tijdreeks $y(t)$, die geanalyseerd worden, maar de golfbewegingen van de eerste-orde verschilreeksen. Het continue equivalent van de eerste orde

³⁰ Merk op dat de Dickey-Fuller test in deze studie niet gebruikt wordt om te discrimineren tussen trendstationair versus differentiestationair, ofwel dat er sprake is van een deterministische trend of een (formele) stochastische trend. De enige restrictie in deze studie is dat de differentiereeks stationair is, zodat spectraalanalyse mogelijk wordt.

verschilreeks is de afgeleide $\partial y(t)/\partial t$ en aangetoond wordt dat de golfbewegingen van de afgeleide hetzelfde zijn in termen van frequentie en golflengte als van de oorspronkelijke tijdreeks $y(t)$. De afgeleide luidt als volgt:

$$y(t) = \sum_{m=0}^{L(N)} (a_m \cos(\omega_m t) + b_m \sin(\omega_m t)) \quad (3.1.2)$$

$$\frac{\partial y(t)}{\partial t} = \omega_m \sum_{m=0}^{L(N)} (b_m \cos(\omega_m t) - a_m \sin(\omega_m t)) \quad (4.2.4)$$

Hieruit volgt dat de amplitude van de golven van een differentiereeks een met m variërende factor ω_m anders zijn dan van de oorspronkelijke reeks, maar dat de frequentie en golflengte niet veranderen. Het minteken voor de sinusfunctie veroorzaakt wel een verschuiving van de golven met een halve periode, maar gemeten in discrete tijdsintervallen is deze verschuiving verwaarloosbaar. De opbouwende golven van een tijdreeks ondergaan in termen van frequentie geen verandering als gevolg van het wegfilteren van een stochastische trend. De stochastische conjunctuurcomponent, gedefinieerd als opbouwende golven met een lengte groter dan één en kleiner dan zeven jaar, is identiek voor een tijdreeks en de differentiereeks. In voorbeeld 3.7 over het differentiefilter wordt dezelfde conclusie op een andere wijze is aangetoond.

De stochastische conjunctuurreksen dienen potentieel als leidende indicator voor de stochastische huizenprijsconjunctuur als ze in voldoende mate samenhangen met de huizenprijsconjunctuur en als ze in voldoende mate voorlopen op de huizenprijsconjunctuur. Op basis van geïdentificeerde leidende indicatoren wordt er een lineair voorspelmodel voor de huizenprijsconjunctuur geschat. Zowel de selectie aan de hand van beide criteria als het schatten van het voorspelmodel gebeurt voor de stochastische conjunctuurreksen in de volgende paragraaf met de in hoofdstuk 3 opgebouwde spectraalanalyse.

4.3: Voorspellen van de conjunctuur

Volgens de bestaande opvattingen over het functioneren van de woningmarkt, zoals samengevat in paragraaf 4.1, wordt de prijsvorming op deze markt bepaald door de interactie van inelastische marktspecifieke aanbodfactoren, demografische en monetair macro-economische vraagfactoren in een beleidsmatige institutionele omgeving.

Over de interactie tussen vraag en aanbod in de institutionele context merkt Van Assenbergh (1999) het volgende op: "De woningmarkt wordt gekenmerkt door een relatief sterk fluctuerende elastische vraag, terwijl het aanbod zich tamelijk inelastisch toont en in ieder geval vertraagd reageert op vraagveranderingen. De onevenwichtigheden tussen vraag en aanbod uiteten zich in de prijsontwikkeling." Deze constatering impliceert de mogelijkheid dat de cyclische beweging van de huizenmarkt samenhangt met de fluctuatie van macro-economische en monetaire vraagfactoren. Bovendien loopt de macro-economische conjunctuurontwikkeling mogelijk voor op de cyclische beweging van de huizenprijs als gevolg van de vertraagde reactie van de inelastische aanbodfactoren. De potentiële samenhang tussen de cyclische beweging van de huizenprijs en de macro-economische conjunctuur stoelt op de aanname dat de trendmatige component in de huizenprijs de invloed van de inelastische aanbodfactoren, de demografische vraagfactoren en de ontwikkeling van de institutionele omgeving incorporeert. De aanname oogt valide voor de als inelastisch gekarakteriseerde aanbodfactoren, de per definitie niet conjunctuurgevoelige demografie en een stabiele en structurele beleidsvoering, waarbij ingrijpende beleidsveranderingen min of meer eens in de zeven jaar voorkomen. Voorbeelden van ingrijpende beleidsverandering zijn onder andere de afschaffing van de hypotheekrente, de invoering van de tweeverdienersregeling en de toewijzing van VINEX-locaties voor het uitbreiden van de woningvoorraad.

Bovenstaande constatering en aannames bieden ruimte voor het bestaan van een conjunctureel leidende indicatorenmodel. In combinatie met een per definitie starre trendmatige ontwikkeling is door terugkoppeling tevens een puntschatting voor de Nederlandse huizenprijs op korte tot middellange

termijn te bepalen. Het lange termijn evenwichtsniveau wordt in deze visie meer gevormd door een afweging van structurele invloeden als demografie, bouwproductie en institutionele context, die door de trendmatige component beschreven worden.

Deze paragraaf beoogt enerzijds uit de verzameling tijdreeksen geschikte leidende indicatoren te selecteren en anderzijds een leidende indicatorenmodel te schatten. De selectie van de leidende indicatoren geschiedt aan de hand van twee selectiecriteria, te weten de mate van samenhang van een conjunctuurreeks ten opzichte van de huizenprijsconjunctuurreeks en de voorlooptijd van een conjunctuurreeks ten opzichte van de huizenprijsconjunctuurreeks.

Uiteindelijk is de doelstelling om met de geselecteerde leidende indicatoren invulling te geven aan het volgende multivariate lineaire model:

$$HPC_{t+s} = g_0 + \sum_{u=-(s_1-s)}^T g_1(u)LI_{1,t-u} + \sum_{u=-(s_2-s)}^T g_2(u)LI_{2,t-u} + \dots + \sum_{u=-(s_q-s)}^T g_q(u)LI_{q,t-u} + N_{t+s}, \quad (4.3.1)$$

waarbij $N_t \sim N(0, \sigma^2)$ o.i.v. en ongecorreleerd met iedere $LI_{j,t}$ en $s = \min\{s_1, \dots, s_q\}$

Hierbij is HPC de vector met de data van de huizenprijsconjunctuur, g_0 een constante en zijn de LI_j 's vectoren met de leidende indicatoren die zijn geselecteerd uit de initiële dataset. De veronderstelling is dat de storingstermen homoskedastisch en ongecorreleerd met elke leidende indicator. De maximale voorspelperiode van de huizenprijsconjunctuur is s en is het minimum van de individuele voorlooptijden van de leidende indicatoren. Het aantal te schatten parameters $g_j(u)$ hangt af van het aantal perioden T uit het verleden die meegenomen worden.

De barometrische methode en de spectraalanalyse geven een verschillende invulling aan de selectiecriteria en de schattingsmethode van het lineaire voorspelmodel. De specifieke invulling door beide methodieken worden respectievelijk in subparagraaf 4.3.1 en subparagraaf 4.3.2 uitgewerkt. Met de conjunctuurvoorspelling van de huizenprijs als basis wordt in het volgende hoofdstuk een puntvoorspelling van de nominale huizenprijsindex gegenereerd.

4.3.1: Barometrische voorspeller van de conjunctuur

De huizenprijsconjunctuur geeft de huidige huizenprijsontwikkeling weer, geschoond voor trendontwikkeling, seizoenseffecten en toevallige schokken. De voorspeller van de huizenprijsconjunctuur zal op basis van een aantal geselecteerde deelreeksen een goede voorspellende waarde moeten geven voor de huizenprijsconjunctuur. Om als leidende indicator geselecteerd te worden moet een conjunctuurreeks in de barometrische traditie voldoen aan een aantal criteria, waaronder conjunctuurgevoeligheid, databeschikbaarheid, theoretische plausibiliteit en consistentie met het historisch conjunctuurbeloop.

De conjunctuurreeksen geven de conjuncturele ontwikkeling weer van economische, monetaire en woningmarkt gerelateerde tijdreeksen en worden als leidende conjunctuurindicator geselecteerd op basis van de volgende criteria:

1. Een voldoende correlatie met de wijzer ($>0,5$);
2. Overeenkomstige omslagpunten;
3. Een zo lang mogelijke voorlooptijd;
4. Theoretische plausibiliteit, een logisch beredeneerbaar verband met de huizenprijs.

De in deze inleiding van de paragraaf geformuleerde selectie-eis van samenhang komt tot uitdrukking in de eerste twee criteria. De selectie-eis van voorlooptijd is gelijk aan het derde criterium. Het vierde criterium is afkomstig uit de barometrische traditie en is voornamelijk gebruikt in paragraaf 4.1 om de initiële tijdreeksenverzameling vast te leggen.

Voor elke tijdreeks is met het PAT-algoritme en de standaardisatieprocedure de conjunctuurreeks bepaald. Het conjuncturele beloop van verschillende reeksen is in veel gevallen niet synchroon;

sommige reeksen lopen vooruit en andere reeksen ijlen na ten opzichte van een bepaalde referentiereeks, waarbij de conjunctuurreeks behorende bij de huizenprijsreeks als referentiereeks dient. De mate van synchroniteit of timing kan op twee manieren worden vastgesteld. Ten eerste door het lokaliseren van de toppen en dalen door grafische inspectie en het vergelijken van de keerpunten van een reeks ten opzichte van de referentiereeks. Een tweede wijze waarop inzicht in de timing wordt verkregen is met de kruislingse correlatie, $\rho(l)$, dat voor verschillende waarden van l , hier de mate van vooruitlopen of najlen uitgedrukt in maanden, de correlatie bepaalt tussen de referentiereeks en de overige beschouwde reeksen. De waarde van l waarvoor de correlatiefunctie zijn hoogste waarde bereikt geldt als de maat voor najling of vooruitlopen.

De berekening van de maximale kruislingse correlatie, de bij deze maximale kruislingse correlatie behorende verschuiving en de grafische inspectie omtrent de bij deze verschuiving behorende synchroniteit tussen toppen en dalen zijn voor elke conjunctuurreeks ten opzichte van de huizenprijsconjunctuur uitgevoerd en gedocumenteerd in tabel B.3 van appendix B. Met de verkregen resultaten zijn aan de hand van de geformuleerde criteria de volgende variabelen geselecteerd, die ieder zes maanden voorlopen op de wijzer: *Lange rente*, *Korte rente* en *Kredietproductie Rabobank*. Deze tijdreeksen (linkeras) en hun conjunctuurreeksen (rechtteras) zijn uitgezet in figuur 4.2.

De conjunctuurvoorspeller is een lineaire combinatie van deze geselecteerde conjunctuurreeksen en dient de huizenprijsconjunctuur zo goed mogelijk te repliceren. Als lineaire combinatie van reeksen die minimaal zes maanden voorlopen is de conjunctuurvoorspeller een replicator van de huizenprijsconjunctuur en tevens een zesmaands voorspeller. Het lineaire voorspelmodel is een toepassing van (4.3.1) met voorspelperiode $s=6$ en een viertal te schatten parameters $g_i(0)$, $i=0,1,2,3$. Deze toepassing is hiermee een lineaire regressie van de huizenprijsconjunctuur op de geselecteerde en verschoven conjunctuurreeksen.

In subparagraaf 4.2.1.2.1 over de huizenprijsconjunctuur zijn, afhankelijk van de instellingen van het PAT-algoritme, een drietal PAT-trends en bijbehorende huizenprijsconjunctuurreeksen geconstrueerd. De beste van de drie reeksen van de huizenprijsconjunctuur is weergegeven in figuur 4.3.1. Het kwaliteitsoordeel over de conjunctuurreeksen van de huizenprijs wordt mede gevormd aan de hand van de praktische overweging in hoeverre de conjunctuurreeks voorspeld kan worden door vooruitlopende conjunctuurreeksen van andere variabelen. Het verklarend vermogen van de regressie-analyse (R^2) en de correlatie tussen de huizenprijsconjunctuur en zijn voorspeller vormen twee kwaliteitsmaatstaven voor de voorspeller en bovendien selectie maatstaven voor de keuze tussen de drie geconstrueerde reeksen van de huizenprijsconjunctuur.

In tabel B.4. van appendix B staan de uitkomsten van de regressies gedocumenteerd met betrekking tot de drie geconstrueerde reeksen van de huizenprijsconjunctuur en de drie geselecteerde leidende indicatoren. De regressie van de beste conjunctuurreeks van de huizenprijs op de leidende indicatoren leverde een R^2 op van 0.88 en een correlatie tussen huizenprijsconjunctuur en zijn voorspeller van 0.89. Zowel de beste conjunctuurreeks als zijn voorspeller zijn grafisch weergegeven in figuur 4.4.1.

Uit de figuur blijkt dat de conjunctuur en zijn voorspeller voor de laatste waarnemingen uiteen gaan lopen. Behalve een reductie in de voorspelkwaliteit kan deze divergentie verklaard worden door een naderende trendomslag in de huizenprijsreeks. De door de trendomslag veroorzaakte storing is een directe consequentie van de uiteindenproblematiek van het PAT-algoritme. Bikker & Fase (1985) formuleren de voorspelkwaliteiten van de gevolgde procedure uit de barometrische traditie als volgt: "De hier gehanteerde methode is in het algemeen minder geschikt voor het maken van voorspellingen omtrent het actuele conjunctuurbeloop. Hiervoor is de techniek der macro-economische structuurmodellen meer geschikt, al biedt de gevonden 'leading indicator' wel een bescheiden mogelijkheid globale uitspraken te doen over het te verwachten conjunctuurbeloop in de nabije toekomst".

4.3.2: Spectraalvoorspeller van de conjunctuur

Het spectrum is de uitsplitsing van een tijdreeksvariabele naar frequentie en geeft de relatieve importantie van de bij de afzonderlijke frequenties behorende golven weer. Op deze manier is het mogelijk te onderzoeken of lange termijn golven in een tijdreeks relatief dominant zijn ten opzichte van

korte termijn golven. De betekenis van een piek in een bepaald frequentiegebied van het spectrum is dat de bij deze frequenties behorende golven dominant zijn in de beweeglijkheid van de tijdreeks.

De mogelijkheid tot uitsplitsing naar frequenties geldt ook voor de in hoofdstuk 3 gedefinieerde spectraalmaatstaven als (partiële) coherentie (3.3.12) en (3.3.31) en verschuiving (3.3.14). De uitsplitsing maakt het mogelijk om vast te stellen of de samenhang en verschuiving tussen twee tijdreeksvariabelen bepaald wordt door korte termijn golven of juist door (middel)lange termijn golven. De toepassing van de spectraalanalyse concentreert zich op de kwantificering met de spectraalmaatstaven van de conjuncturele relaties tussen twee tijdreeksvariabelen. Conjunctuur is synoniem aan de cyclische component van een tijdreeks en bestaat in spectraalterminologie uit de opbouwende golven van een tijdreeks, die een frequentie kennen behorende bij golflengtes groter dan 1 jaar en kleiner dan 7 jaar. De spectraalmaatstaven worden toegepast op het frequentieinterval behorende bij de conjuncturele beweging van een tijdreeks. De gemiddelde waarde van de maatstaven voor dit interval vormen de kritische waarden, die ten opzicht van instelbare referentiewaarden bepalend zijn voor de selectie van de variabele als leidende indicator. De twee grensfrequenties van het frequentieinterval zijn de frequenties die behoren bij golven met een lengte van 1 jaar en met een lengte van 7 jaar. De berekening van de bij een frequentie behorende golflengte is uitgevoerd in voorbeeld 3.1 voor een interval van het in figuur 3.2 geïllustreerde spectrum. De gebruikte relatie tussen de golflengte l in maanden en de frequentie ω_k luidt als volgt: ($l=N/k=2\pi/\omega_k$), waarbij N het aantal waarnemingen en k een discrete variabele die loopt vanaf nul tot een bovengrens van $N/2$.

De spectraalmaatstaven worden toegepast op het conjuncturele frequentieinterval van een tijdreeksvariabele. In hun kern zijn de spectraalmaatstaven wiskundige bewerkingen van de beide spectra van twee tijdreeksvariabelen en hun kruislingse spectrum. De schatters van de spectraalmaatstaven zijn daarmee ook wiskundige bewerkingen van de spectrumschatters. De eerste stap is het toepassen van de kernschatter (3.4.4) van paragraaf 3.4 op de initiële dataverzameling. De toepassing met als specificatie van de weegfunctie (3.4.5) de Tuckey-Hanning weegfunctie (3.4.9) vereist een waarde voor de bandbreedte van deze weegfunctie. De keuze voor de bandbreedte levert enige praktische complicaties op.

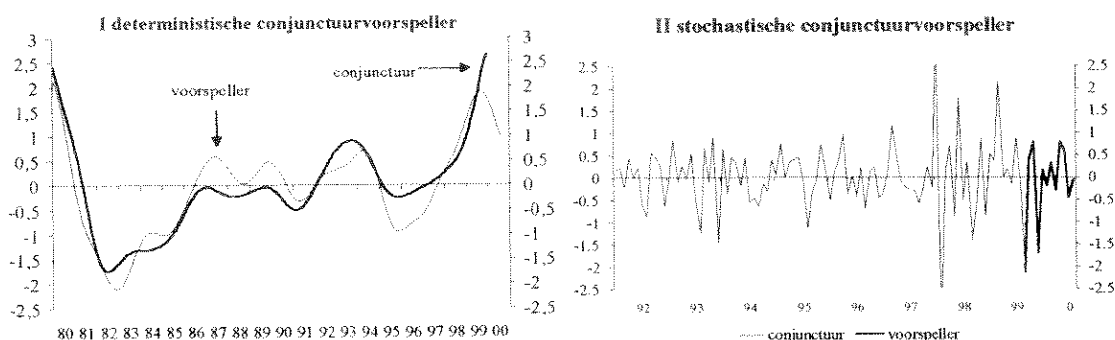
4.3.2.1: De bandbreedte van de spectraalschatter

De praktische toepassing van de spectraalschatter vereist een specificatie van de bandbreedte van de raamschatter, de parameter m . In subparagraaf 3.4.2 is reeds aangetoond dat voor de onzuiverheid van de schatter met (3.4.6) geldt: $b(\omega, m) = O(1/m^2)$ en met (3.4.7) is reeds aangetoond dat de schatter consistent is met $O(m/\sqrt{n})$. Volgens Priestley (1981) convergeert de variantie zelfs met $O(m/n)$. Bovenstaande relaties impliceren dat de bandbreedte m en de variantie van de schatter omgekeerd evenredig aan elkaar zijn, terwijl een hogere bandbreedte de zuiverheid van de schatter wel ten goede komt. Met het relateren van "resolvability" aan bandbreedte en "reliability" aan variantie formuleert Grenander de *Grenander Onzekerheidsrelatie*³¹ als volgt: "reliability and resolvability are antagonistic". Als het echter mogelijk is om zowel de hoeveelheid data n als de bandbreedte m te bepalen kan de schatter een geëiste mate van precisie bereiken. De bandbreedteparameter m is met (3.4.6) bepalend voor de zuiverheid van de spectraalschatter en de geëiste convergentie van de variantie legt gegeven de waarde voor m een claim op de benodigde hoeveelheid data n . Echter, in de praktijk gelden afhankelijk van databeschikbaarheid restricties voor n en voor de in deze studie gebruikte reeksen ligt deze restrictie meestal rond de 310.

De mate van zuiverheid is gegeven de restrictie nog steeds te bepalen door een geschikte waarde voor de parameter $m(<n)$ te kiezen, maar dit gaat wel koste van de mate van precisie, de variantie. Het ligt meer voor de hand de waarde m te kiezen als resultaat van een compromis tussen (on)zuiverheid en variantie.

³¹ Grenander bemerkte de sterke overeenkomst tussen de aard van dit resultaat en Heisenberg's befaamde 'onzekerheidsrelatie' in de quantum fysica. Zie voor de quote Priestley (1981).

Figuur 4.4: Conjunctuurvoorspellers



In fysische experimenten bestaan vaak a priori aanwijzingen, die bepaalde waarden van de bandbreedte uitsluiten. Deze priorinformatie bestaat echter niet voor economische of financiële tijdreeksen. Priestley (1981) formuleert dat als volgt: "...many situations in which we have no prior knowledge whatsoever on the spectral bandwidth. Unfortunately, these occur only too frequently in practice; for example, in the analysis of economic time series where the concept of bandwidth does not have any physical interpretation. In such cases one may still attempt to estimate spectral density functions but now we have to adopt more of a 'trial and error' approach to the selection of the estimation parameters."

De toepassing van de kernel in de kernelschatter (3.4.4) betekent dat de geschatte autocovarianties gewogen gesommeerd worden. De weegfunctie van de kernelschatter is een functie die verderaf gelegen covarianties een kleiner gewicht geeft en heel veraf gelegen covarianties niet meeweegt. De Tuckey-Hanning kernel is weergegeven in figuur 3.14 en hieruit blijkt dat autocovarianties met een vertraging die groter is dan de bandbreedteparameter m niet meetellen, omdat de weegfunctie hiervoor de waarde nul genereert. Een redelijke en natuurlijk keuze voor m is dan de waarde van waaraf de autocovariantiefunctie ongeveer gelijk is aan nul. In de praktijk betekent dit een keuze van m op basis van de geschatte covariantiefunctie op de volgende manier:

$$\hat{\gamma}(s) \sim 0, \quad |s| > m \quad (4.3.2)$$

De geschatte autocovarianties zijn echter zelf geautocorreleerd en zullen in het algemeen langzamer naar nul convergeren dan de theoretische autocovarianties. Op de tweede plaats is het voor een autocovariantiefunctie die geleidelijk naar nul convergeert een redelijk arbitraire bepaling wanneer deze functie nog wel en wanneer deze functie niet meer ongeveer nul is.

Hoewel (4.3.2) een praktische oplossing biedt voor de keuze van de bandbreedteparameter m geeft het geen inzicht in wat precies de invloed is van deze parameter op het berekende spectrum. Het spectrum van een tijdreeks wordt berekend door de geschatte autocovarianties gewogen te sommeren, waarbij deze vanaf grenswaarde m de waarde nul krijgen en niet meer meedoen. Het op deze wijze verkregen spectrum is volgens de theoretische relatie (3.4.8) identiek aan het spectrum dat wordt verkregen uit een gewogen integraal van het periodogram. De weegfunctie van de integraal vormt hierbij een Fourier-transformatiepaar met de weegfunctie van de autocovarianties. Toepassing van deze relatie (3.4.8) op de Tuckey-Hanning kernel (3.4.9) genereert deze weegfunctie van de integraal (3.4.10), die tevens met (3.4.11) de relatie weergeeft tussen het berekende spectrum en het ruwe spectrum. Met (3.4.11) kan het berekende spectrum worden geïnterpreteerd als zijnde een gladstrijking van het ruwe spectrum over een frequentieinterval, waarbij de lengte van het frequentieinterval omgekeerd evenredig is met de waarde van de bandbreedteparameter m . Een lage waarde voor m betekent dan een groot interval waarop het spectrum gladgestreken wordt en resulteert in een relatief effen en vlak spectrum.

4.3.2.2: De selectie van de reeksen

De selectie van de leidende cyclische indicatoren geschiedt aan de hand van de gemiddelde waarden van de verschuiving s met (3.3.14) en de samenhang K^2 met (3.3.12) voor de opbouwende golven met lengtes groter dan 1 jaar en kleiner dan zeven jaar.

De definitieve keuze voor de bandbreedteparameter m van de weegfunctie is voor de selectie van de leidende cyclische indicatoren uit de initiële dataset genuanceerd door de spectraalmaatstaven voor meerdere waarde van deze bandbreedteparameter te berekenen. De kwantificering van de spectraalmaatstaven onder een variërende bandbreedteparameter zijn in appendix B.2.2 en tabel B.6 gedocumenteerd. Op basis van deze berekeningen worden de volgende leidende indicatoren voor de cyclische beweging van de huizenprijs geselecteerd, te weten *Korte rente*, *Hypotheekrente*, *Verkooptijd* en *CPI-index*.

Gezien het feit dat de *hypotheekrente* een sterk synchroon verloop kent met de *lange rente* worden deze rente en de korte rente door zowel de spectraalmethode als de barometrische methode geselecteerd. De reeks *verkooptijd* kent een sterke theoretische link met het functioneren van de huizenmarkt. Als huizen kort te koop staan is dat een signaal dat de markt overspannen is en dat zal een opwaartse beweging van de huizenprijs tot gevolg hebben. De selectie van de *CPI-index* als maat voor de ontwikkeling van de inflatie is enigszins verrassend, maar kennelijk fluctueert het nominale prijspeil van woningen vertraagd op de ontwikkeling van de consumentenprijzen.

4.3.2.3: Het schatten van het leidende cyclische indicatorenmodel

De spectraalvoorspeller van de cyclische beweging van de huizenprijs bestaat uit een toepassing van het lineaire leidende cyclische indicatorenmodel (4.3.1) met behulp van spectraalanalyse. De reeksen van de cyclische bewegingen zijn de differentiereeksen van de oorspronkelijke reeksen, zoals is gesteld in subparagraaf 4.2.2. De huizenprijsconjunctuur HPC is de stationaire 2^e-orde differentiereeks van de huizenprijsreeks. De leidende indicatoren LI_i , $i=1,2,3,4$ zijn de vier met behulp van de gekwantificeerde spectraalmaatstaven geselecteerde leidende cyclische indicatoren³². Afgezien van de constante term g_0 is het leidende indicatorenmodel (4.3.1) een toepassing van het algemene multivariate lineaire model (3.3.24) uit paragraaf 3.3.4. De optimale lineaire relatie tussen de huizenprijsconjunctuur en de leidende cyclische indicatoren wordt in spectraalterminologie weergegeven door een transferfunctie, die met (3.3.27) wordt berekend. Deze transferfunctie is het spectraalequivalent van de optimale coëfficiënten $g(u)$, die als gevolg berekend kunnen worden met de inverse Fourier-transformatie. Een dergelijke berekening wordt voor het bivariate lineaire model gegeven met (3.3.23). Zoals reeds in subparagraaf 3.3.4 is gesteld kunnen op deze wijze de optimale coëfficiënten worden berekend ongeacht de restricties voor T en s , zoals die in het leidende indicatorenmodel worden opgelegd.

De equivalentie tussen het algemeen lineaire model (3.3.24) en het leidende indicatorenmodel (4.3.1) kent echter één belangrijk manco, namelijk dat (3.3.24) slechts geldt voor tijdreeksen met een verwachting van nul. Dit manco wordt op de volgende manier voor (4.3.1) gecorrigeerd:

$$\begin{aligned}
 HPC_{t+s} - \mu_{HPC} &= \sum_{u=-(s_1-s)}^T g_1(u) \{LI_{1,t-u} - \mu_{LI_1}\} + \dots + \sum_{u=-(s_4-s)}^T g_4(u) \{LI_{4,t-u} - \mu_{LI_4}\} + N_{t+s} \\
 \Leftrightarrow HPC_{t+s} &= g_0 + \sum_{u=-(s_1-s)}^T g_1(u) LI_{1,t-u} + \dots + \sum_{u=-(s_4-s)}^T g_4(u) LI_{4,t-u} + N_{t+s},
 \end{aligned}$$

waarbij $g_0 = \left\{ \mu_{HPC} - \mu_{LI_1} \sum_{u=-(s_1-s)}^T g_1(u) - \dots - \sum_{u=-(s_4-s)}^T g_4(u) \right\}$ (4.3.3)

en $N_t \sim N(0, \sigma^2)$ o.i.v. en ongecorrleerd met iedere LI_{jt} . Bovendien geldt $s = \min\{s_1, \dots, s_q\}$

Met de constante term (4.3.3) is het leidende cyclische indicatorenmodel (4.3.1) met de spectraalschatter te schatten. Resteert nog een invulling voor de waarde van T in (4.3.1). Deze parameter specificceert het aantal perioden in het verleden, waarvan de realisaties meegenomen worden in het leidende indicatorenmodel. De meest fundamentele abstractie van de spectraalanalyse is periodiciteit. In spectraaltermen worden niet-periodieke datareeksen geïnterpreteerd als periodiek met een oneindige lange periode. In de praktische toepassing is een dataverzameling echter beperkt in zijn

³² Merk op dat de geselecteerde leidende indicator *CPI-index* net als de huizenprijs 2^e-orde stationair is.

lengte.. Het algemene lineaire model (3.3.24) berekent de coëfficiënten $g(u)$ over een periode van $-\infty < u < \infty$. Door de de beperkte waarnemingshorizon geldt in de praktijk de restrictie dat u niet groter kan zijn dan de lengte van de geobserveerde datareeks. Met berekening (3.3.23) van de $g(u)$'s voor het bivariate algemene lineaire model blijkt bovendien dat de $g(u)$'s symmetrisch zijn rond nul. Hierdoor wordt de restrictie verzaamd dat u niet groter kan zijn dan de helft van de lengte van de geobserveerde datareeks, waarmee een invulling voor de parameter T is verkregen.

Met de geselecteerde leidende cyclische indicatoren is model (4.3.1) en de constante term (4.3.3) geschat. Op basis hiervan is de cyclische beweging van de huizenprijs voorspeld voor een periode van 11 maanden. Deze voorspelling is grafische weergegeven in figuur 4.4.II en de berekende verschuiving s en samenhang K^2 zijn, evenals het aantal beschikbare observaties n en de gebruikte bandbreedteparameter m , weergegeven in tabel 4.5. Merk hierbij op dat de waarden van de coherentie maatstaf K^2 niet moeten worden beschouwd als een absolute maatstaf voor samenhang, analoog aan de correlatie, maar als een relatieve samenhangskwantificering voor de prestatie van een variabele ten opzichte van de overige variabelen. De kwantificering van de coherentie varieert namelijk met de bandbreedteparameter m^{33} , wat zich nadrukkelijk manifesteert bij de selectie van de leidende indicatoren in Appendix B.2.2.

Tabel 4.5: Kwantificeringen behorende bij cyclische huizenprijsvoorspeller

K^2	s	variabele	m	n	
		<i>Huizenprijs</i>	45	314	
0.35	11	<i>Verkooptijd</i>	30	316	$s=11$
0.33	14	<i>Korte rente</i>	35	317	
0.28	13	<i>Hypotheekrente</i>	43	175	$T=87$
0.23	16	<i>CPI-index</i>	48	314	

³³ De K^2 berekend over de periodogrammen van twee reeksen is zelfs per definitie gelijk aan 1. Dat komt omdat deze maatstaf voor de betreffende frequentie slechts op één enkele waarneming gebaseerd is. Door toepassing van een kernel wordt impliciet verondersteld dat de coherentie dezelfde waarde heeft voor een interval van frequenties, waardoor deze berekend kan worden over meerdere observaties. Deze redenering verklaart de afhankelijkheid van K^2 en M . Zie Priestley (1981) blz 708.

Hoofdstuk 5: De huizenprijs voorspeld, een evaluatie

De uiteindelijke onderzoeksvraag voor de huizenmarkt is het genereren van een puntvoorspelling voor de huizenprijs. Een puntvoorspeller van de huizenprijs wordt verkregen door de geconstrueerde conjunctuurvoorspellers uit het vorige hoofdstuk met een voorspelling voor de trend terug te koppelen naar de oorspronkelijke huizenprijs. Voorspellingen voor de trendmatige beweging zijn gebaseerd op de per definitie starre en onveranderlijke ontwikkeling van de trendmatige component. Een meer precieze invulling van de trendvoorspeller hangt af van het gekozen type trend, deterministisch of stochastisch en wordt in paragraaf 5.1 uitgewerkt. De op basis van beide type van trendvormen geconstrueerde voorspellers voor de huizenprijs worden in paragraaf 5.2 geëvalueerd door het zogenaamde 'out-of-sample'-terugtesten. De huizenprijzen van de nabije historie worden voorspeld en deze voorspellingen worden vergeleken met de realisaties van de huizenprijs om de kwaliteit van de voorspeller te meten.

5.1: Van conjunctuurvoorspelling naar puntvoorspelling

Het genereren van een puntvoorspelling voor de nominale huizenprijsindexreeks vereist een terugkoppeling van de conjunctuurvoorspeller met een voorspelling van de trendmatige ontwikkeling. De aanvankelijke decompositie van een tijdreeks in onder meer trend en conjunctuur werd ingegeven door de vaststelling dat de huizenmarkt een macro-economische factor van betekenis is en daardoor beïnvloed wordt door de algemeen economische schommelingen. Voor het voorspellen op korte tot middellange termijn van de huizenprijs herbergen deze schommelingen gezien hun tijdsduur belangrijke informatie voor deze periode. Met een starre en stabiele trendmatige ontwikkeling is de terugkoppeling feitelijk een synthese van componenten, die weer terugleidt tot de oorspronkelijke tijdreeks.

5.1.1: De barometrische voorspeller; deterministische trend

De deterministische huizenprijsvoorspeller is de voorspeller van de nominale huizenprijsindex die wordt verkregen door de deterministische voorspeller van de huizenprijsconjunctuur terug te koppelen met een lineaire extrapolatie van de deterministische PAT-trend. De lineaire extrapolatie van de trend als voorspeller van de trendmatige ontwikkeling is ingegeven door de overweging dat een trend per definitie star en onveranderlijk is. Het PAT-algoritme heeft voor de huizenprijsreeks van de afgelopen 25 jaar 3 trendomslagpunten bepaald en dat impliceert eens in de 6 jaar een trendomslag. Vanwege het feit dat het laatste trendomslagpunt begin 1992 heeft plaatsgevonden, is de verwachting dat er op korte termijn een omslag in de trend gaat plaatsvinden. De discrepantie tussen huizenprijsconjunctuur en zijn voorspeller in de tweede helft van 1999 kan behalve door een geringe voorspelkwaliteit van de conjunctuur tevens verklaard worden uit de onnauwkeurigheid van de trend als gevolg van het naderende trendomslagpunt en de bijbehorende trendcorrectie. Dit laatste fenomeen is een rechtstreeks gevolg van de onbepaaldheid van het PAT-algoritme, zoals beschreven in subparagraaf 4.2.1.2.

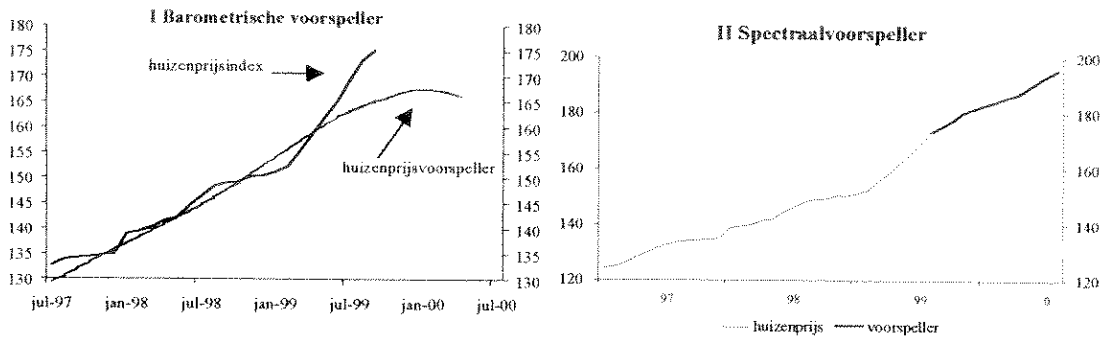
De deterministische huizenprijsconjunctuur is een verhouding van de beweging van de huizenprijs ten opzichte van de trendmatige beweging. De barometrische huizenprijsvoorspeller is dan een vermenigvuldiger van de deterministische conjunctuurvoorspeller en de lineaire extrapolatie van de PAT-trend. De barometrische huizenprijsvoorspeller staat grafisch weergegeven in figuur 5.1.1.

5.1.2: De spectraalvoorspeller; stochastische trend

De feitelijke spectraalvoorspellingen van de cyclische beweging van de huizenprijs, zoals grafisch weergegeven in figuur 4.4.11 impliceren tevens een voorspelling voor de stochastische trend. Zoals in subparagraaf 2.3.2 over stochastische trends is gesteld is de verwachting van een stochastische trend gelijk aan nul en manifesteert deze trendvorm zich ex post als realisaties van de stochasten. Voor de constructie van de spectraalvoorspeller van de huizenprijs is dan ook geen expliciete voorspelling voor de stochastische trend nodig. De huizenprijsvoorspeller wordt feitelijk berekend door een omgekeerde differentiebewerking van de gegenereerde cyclische voorspellingen en gaat als volgt:



Figuur 5.1.: De huizenprijs voorspeld



$$\Delta^2 HP_{t+1} = HP_{t+1} - 2 * HP_t + HP_{t-1}$$

$$\Leftrightarrow HP_{t+1} = \Delta^2 HP_{t+1} + 2 * HP_t - HP_{t-1} \quad (5.1.1)$$

Hierbij is:

$\Delta^2 HP_{t+1}$ = de voorspelling op t+1 van de 2^e-orde stationaire huizenprijsconjunctuur,
 HP_{t+1} = de voorspelling van de nominale huizenprijsindexreeks.

Met (5.1.1) is een voorspelling voor de eerste periode aan de hand van de voorspelling van de huizenprijsconjunctuur. Deze berekening is tevens gebaseerd op de laatste twee realisaties van de huizenprijs, te weten HP_{t-1} en HP_{t-2} . Het genereren van meerdere voorspelling is simpelweg een iteratie van (5.1.1) voor $t+2$, $t+3$, etcetera, afhankelijk van het aantal conjunctuurvoorspellingen. De op deze wijze verkregen spectraalvoorspeller van de huizenprijs is grafisch weergegeven in figuur 5.1.II.

5.2: De kwaliteit van de voorspellers, een terugtestevaluatie

Het meten van de kwaliteit van beide voorspellers geschiedt op basis van het 'out-of-sample' voorspellen van de huizenprijs in het nabije verleden. De kwaliteit wordt dan gemeten aan de hand van de juistheid van de voorspellingen. Beide voorspellers voor de huizenprijs zijn in abstractie de conjunctuurvoorspellers teruggekoppeld met de trendvoorspellers. De conjunctuurvoorspeller is een specificatie van het lineaire leidende cyclische indicatorenmodel (4.3.2), waarbij bepaald is welke variabelen dienst doen als leidende indicator, de individuele voorlooptijd s van de afzonderlijke indicatoren en het totaal aantal vertragingen T die meewegen in de berekening van de voorspeller. De voorspellingen van de trend zijn in het deterministische geval een lineaire extrapolatie van de PAT-trend, die wordt bepaald door de realisaties van de onderliggende tijdreeks. In het stochastische geval hoeft de trend niet voorspeld te worden. Met de vaststelling van bovenstaande gegevens bestaat het voorspellen van de huizenprijs louter uit het schatten van de $g(u)$'s van het leidende indicatorenmodel (4.3.1). De 'out-of-sample' -term in het terugtesten staat voor het principe dat de huizenprijs van de vorige periode ($t-1$) voorspeld wordt aan de hand van de data tot periode ($t-1$).

Voor de kwaliteit van de voorspellers bestaan twee wijdverbreide maatstaven, te weten de verwachte voorspelfout, die de zuiverheid van de voorspeller meet en de verwachte kwadratische voorspelfout, die de betrouwbaarheid van de voorspeller meet.

$$\text{Verwachte voorspelfout : } VVF = E \left\{ HP_t - \hat{HP}_t \right\}, \quad (5.2.1)$$

$$\text{Verwachte kwadratische fout : } VKF = E \left\{ \left(HP_t - \hat{HP}_t \right)^2 \right\}, \quad (5.2.2)$$

waarbij \hat{HP} de voorspelde huizenprijs.

Tabel 5.2: Terugtestevaluatie van de voorspellers

voorspelperiode	Spectraalvoorspeller		Barometrisch voorspeller	
	VVF	VKF	VVF	VKF
1	-1.62	5.21	1.09	16.11
2	-1.34	9.98	1.08	15.33
3	-1.26	19.10	1.01	14.33
4	-0.98	26.89	0.94	11.75
5	-0.55	42.45	0.98	12.55
6	-0.052	63.86	1.34	17.81
7	0.56	92.54		
8	1.12	124.76		
9	1.06	152.80		
10	0.23	160.55		
11	-0.26	182.19		

De kwaliteitsmaatstaven (5.2.1) en (5.2.2) zijn voor zowel de barometrische als de spectraalvoorspeller van de huizenprijs berekend op basis van een terugtestperiode van 30 maanden. De resultaten staan in tabel 5.2 gepresenteerd en de beide kwaliteitsmaatstaven zijn de steekproefequivalenten van de definities (5.2.1) en (5.2.2). Uit de tabel blijkt dat de zuiverheid van de spectraalvoorspeller en de barometrische voorspeller aan elkaar gewaagd zijn, maar dat voor voorspelperioden groter dan drie maanden de barometrische voorspeller een factor 3.5 lagere kwadratische voorspelfout. Hiermee is de barometrische voorspeller van de huizenprijs kwalitatief beter dan de spectraalvoorspeller.

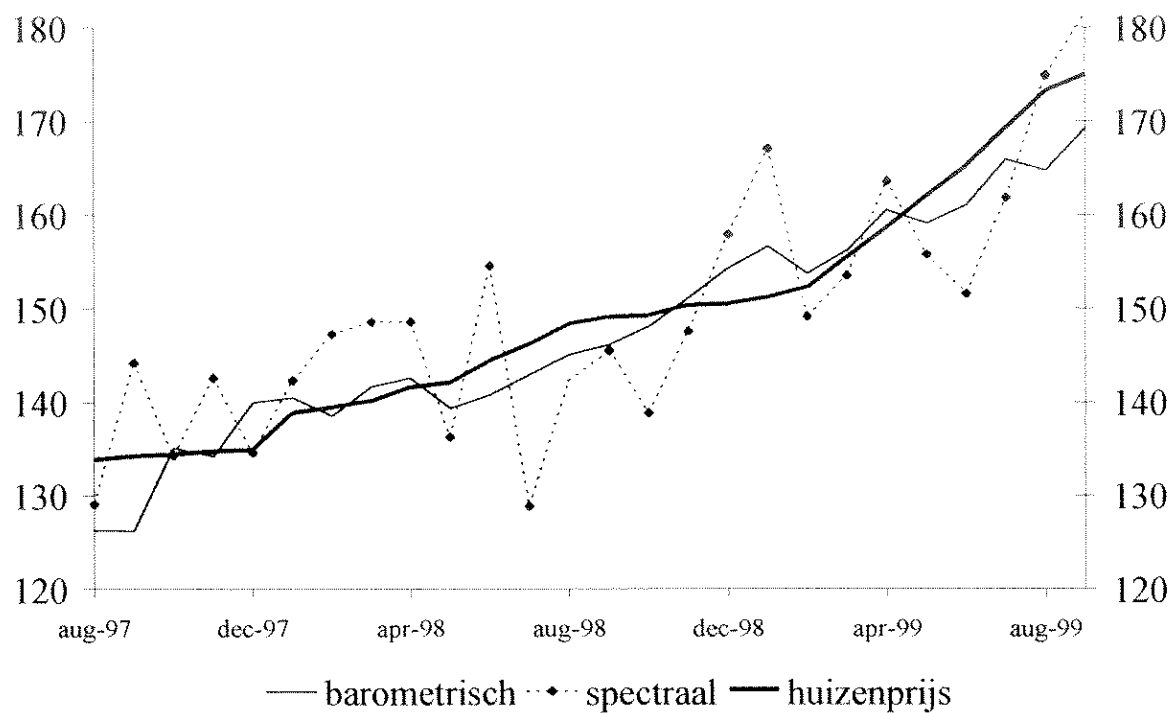
Ter illustratie zijn de zes-maands voorspellingen van zowel de barometrische als de spectraalschatter voor voor de terugtestperiode van twee jaar grafische weergegeven in figuur 5.3. In deze figuur komt de hogere *VKF* tot uitdrukking in de beweeglijkheid van de reeks met zes-maands spectraalvoorspellingen voor de afgelopen periode.

De spectraalvoorspeller bereikt wel dezelfde mate van precisie als de barometrische schatter, maar niet dezelfde mate van betrouwbaarheid. Volgens de Grenander onzekerheidsrelatie uit paragraaf 3.4. bestaat er een afruil tussen precisie en betrouwbaarheid van de spectraalschatter bij een gegeven hoeveelheid waarnemingen. Bij toeneming van het aantal waarnemingen wordt de schatter zowel preciezer als betrouwbaarder. Deze relatie geldt tevens voor de spectraalschatter van het leidende indicatorenmodel (4.3.1).

De waarnemingshorizon bedraagt in deze studie maximaal 320 maanden en dat is voor de praktijk van de economiebeoefening een redelijke lange termijn. In spectraalterminologie is dit waarnemingsaantal echter beperkt, getuige het volgende citaat uit Priestely (1981), pagina 545: "For fairly low proportional errors (e.g., 25%) the design relations lead to quite large values of N (the total number of observations). In example 1 we found that we required approximately 7000 observations, and in many practical situations it would be quite unrealistic to expect to have available time series of this length" en even verderop "it warns us that for typical lengths of series (and in this context, 1000 observations would usually be regarded as a fairly long series) we must be prepared to accept quite high proportional errors, typically of the order of 60%-80%."

Gegeven het waarnemingsaantal en dit citaat lijkt het alsof de spectraalvoorspeller zichzelf verbetert bij een toenemend aantal waarnemingen. Het is echter moeilijk een onderscheid te maken tussen de kwaliteit van de voorspeller als lineaire combinatie van gerelateerde variabelen en het kwaliteitstekort veroorzaakt door een beperkt aantal waarnemingen. Vervolgonderzoek kan zich richten op het specificeren van de betrouwbaarheid van de spectraalschatters en het daarvoor benodigde aantal waarnemingen. Een spectraalvoorspeller voor de huizenprijs op basis van betrouwbare spectraalschatters geeft pas de werkelijke kwaliteitsindicatie van de voorspellingen.

Figuur 5.3: Zes-maands voorspellingen van de huizenprijs



Bibliografie

- Anderson, T.W., 1971, 'The Statistical Analysis of Time Series', John Wiley & Sons.
- Assenbergh, W. van, 1999, 'Nederlandse woningmarkt: ontwikkelingen, onevenwichtigheden en oplossingen', Financiële & Monetaire studies, jaargang 17, nr. 2.
- Assenbergh, W. van, 2000, 'Prijs je rijk', ESB 28-7-2000, p 600-602, dit artikel is integraal opgenomen in appendix D.
- Bierens, H.J., 2000, 'EasyReg 2000', Department of Economics, Pennsylvania State University.
- Bikker, J.A. & M.M.G. Fase, 1985, 'De datering van economische fluctuaties: proeve van een conjunctuurspiegel voor Nederland 1965-1984', Maandschrift economie, jaargang 49, p 299-332.
- Boschan, Ch. & W.W. Ebanks, 1978, 'The Phase-Average trend: a new way of measuring economic growth', Proceedings of the Business and economic statistics section, American Statistical Association, p 332-335.
- Duijn, J.J. van, 1978, 'Dating postwar business cycles in the Netherlands, 1948-1976', De economist, Nr. 4, p 474-504.
- Fuller, W. A., 1996, 'Introduction to Statistical Time Series', John Wiley & Sons (2nd edition).
- Hamilton, J.D., 1990, 'Time Series Analysis, Princeton University Press.
- Hannan, E.J., 1970, 'Multiple Time Series, John Wiley & Sons.
- Koopmans, L.H., 1974, 'The Spectral Analysis of Time Series', Academic Press (Londen).
- Kruijt, B. & J. Janssen, 1991, 'Een marktstemmingsindex voor vastgoed', ESB 23-1-1991, p 92-95.
- Krauss, T.P., J.N. Little & L. Shure, 1994, 'Signal Processing Toolbox User's Guide', The MathWorks.
- Makridakis, S. & S.C. Wheelwright, 1985, 'Forecasting methods for management', John Wiley & Sons.
- Ljung, L., 1995, 'System Identification Toolbox User's Guide', The MathWorks.
- Priestly, M.B., 1981, 'Spectral Analysis and Time Series: Volume 1 Univariate Series', Academic Press.
- Priestly, M.B., 1981, 'Spectral Analysis and Time Series: Volume 2 Multivariate Series, Prediction and Control', Academic Press.
- Rooij, M.C.J. van & A.C.J. Stokman, 2000, "Verzilvering overwaarde huis: een statistische analyse van besteding en risico's", Onderzoeksrapport WO&E nr. 623, te downloaden op <http://www.dnb.nl/research/woe/pdf/623.pdf>.
- Rooij, M.C.J. van & L. Vos, 1999, 'De samenhang tussen Europese huizenprijzen!', Maandschrift economie, juni 1999, p224-229.
- Tolstov, G.P., 1962, 'Fourier Series', Prentice Hall, vertaald uit het Russisch door: R.A. Silverman.

Appendix A: Uitwerkingen en bewijzen spectraalanalyse

Appendix A.3.1.1: Orthogonaliteit goniometrisch stelsel

Onder de aanname dat m en r element zijn van de verzameling $\{0, 1, \dots, L(N)\}$ geldt :

$$\sum_{t=0}^{N-1} \cos \frac{2\pi m}{N} t \cos \frac{2\pi r}{N} t = \begin{cases} N, & m = r = 0 \text{ or } N/2 \\ N/2, & m = r \neq 0 \text{ or } N/2 \\ 0, & m \neq r \end{cases}$$

$$\sum_{t=0}^{N-1} \sin \frac{2\pi m}{N} t \cos \frac{2\pi r}{N} t = 0 \quad \forall m, r$$

$$\sum_{t=0}^{N-1} \sin \frac{2\pi m}{N} t \sin \frac{2\pi r}{N} t = \begin{cases} N/2, & m = r \neq 0 \text{ or } N/2 \\ 0, & m \neq r \end{cases}$$

bewijs (gedeeltelijk) :

Beschouw ten eerste de som van de produkten van twee cosinus functies,

$$S(m, r) \equiv \sum_{t=0}^{N-1} \cos \frac{2\pi m}{N} t \cos \frac{2\pi r}{N} t = \frac{1}{2} \sum_{t=0}^{N-1} \left[\cos \frac{2\pi t}{N} (m+r) + \cos \frac{2\pi t}{N} (m-r) \right]$$

Nu is eenvoudig in te zien

$$(1) m = r = 0 \text{ of } m = r = N/2 \quad (N \text{ even}) : S(m, r) = \frac{1}{2} \sum_{t=0}^{N-1} (1+1) = N,$$

$$(2) m = r \neq 0 \text{ of } m = r \neq N/2 \quad (N \text{ even}) : S(m, r) = \frac{1}{2} \sum_{t=0}^{N-1} \left[\cos \frac{2\pi t}{N} 2m \right] + \frac{1}{2} N,$$

Omdat $e^{i(2m)2\pi} = e^{i2\pi} = 1$ geldt :

$$\sum_{t=0}^{N-1} \cos \frac{2\pi t}{N} 2m = \frac{1}{2} \sum_{t=0}^{N-1} \left\{ e^{i2m2\pi t/N} + e^{-i2m2\pi t/N} \right\} = \frac{1 - (e^{i2m2\pi/N})^N}{1 - e^{i2m2\pi/N}} + \frac{1 - (e^{-i2m2\pi/N})^N}{1 - e^{-i2m2\pi/N}} = 0,$$

$$\text{dus : } S(m, r) = \frac{1}{2} N.$$

$$(3) m \neq r : S(m, r) = \frac{1}{2} \sum_{t=0}^{N-1} \left[\cos \frac{2\pi t}{N} (m+r) + \cos \frac{2\pi t}{N} (m-r) \right] =$$

$$\frac{1}{4} \sum_{t=0}^{N-1} \left[e^{i2\pi t(m+r)/N} + e^{-i2\pi t(m+r)/N} + e^{i2\pi t(m-r)/N} + e^{-i2\pi t(m-r)/N} \right] = 0.$$

De som van de produkten van een sinus en een cosinus functie is gegeven door :

$$\sum_{t=0}^{N-1} \sin \frac{2\pi m}{N} t \cos \frac{2\pi r}{N} t = \frac{1}{2} \sum_{t=0}^{N-1} \left\{ \sin(m+r) \frac{2\pi t}{N} + \sin(m-r) \frac{2\pi t}{N} \right\} = 0 \quad \forall m, r$$

$$\text{beschouw : } \sum_{t=0}^{N-1} \sin(m+r) \frac{2\pi t}{N} = \sum_{t=0}^{N-1} \cos \left[(m+r) \frac{2\pi t}{N} - \frac{1}{2} \pi \right] = \sum_{t=0}^{N-1} \left\{ e^{i(m+r) \frac{2\pi}{N} - \frac{1}{2} i\pi} + e^{-i(m+r) \frac{2\pi}{N} + \frac{1}{2} i\pi} \right\} =$$

$$e^{-\frac{1}{2} i\pi} \sum_{t=0}^{N-1} e^{i(m+r) \frac{2\pi t}{N}} + e^{\frac{1}{2} i\pi} \sum_{t=0}^{N-1} e^{-i(m+r) \frac{2\pi t}{N}} = 0 \quad \forall m, r \quad (m \neq r).$$

Het bewijs van het produkt van de sinus functies volgt uit de volgende relatie :

$$\sum_{t=0}^{N-1} \sin \frac{2\pi m}{N} t \cos \frac{2\pi r}{N} t = -\frac{1}{2} \sum_{t=0}^{N-1} \cos \frac{2\pi t}{N} (m+r) - \cos \frac{2\pi t}{N} (m-r)$$

en het bewijs volgt met een gelijke redenering als boven.

□

Appendix A.3.1.3: Orthogonaliteit van oneindige goniometrisch stelsel

Gegeven dat m en j niet - negatieve gehele getallen zijn, dan :

$$\int_{-\pi}^{\pi} \sin mx \cos jx dx = 0 \quad \forall m, j;$$

$$\int_{-\pi}^{\pi} \sin mx \sin jx dx = \begin{cases} 0, & m \neq j, \\ \pi, & m = j \neq 0, \\ 0, & m = j = 0; \end{cases}$$

$$\int_{-\pi}^{\pi} \cos mx \cos jx dx = \begin{cases} 0, & m \neq j, \\ \pi, & m = j \neq 0, \\ 2\pi, & m = j = 0. \end{cases}$$

bewijs (gedeeltelijk) :

$$\int_{-\pi}^{\pi} \sin mx \cos jx dx = \frac{1}{2} \int_{-\pi}^{\pi} \{\sin(m+j)x + \sin(m-j)x\} dx = \left[\frac{-\cos(m+j)x}{2(m+j)} - \frac{\cos(m-j)x}{2(m-j)} \right]_{-\pi}^{\pi} = 0;$$

$$\int_{-\pi}^{\pi} \sin mx \sin jx dx = \frac{1}{2} \int_{-\pi}^{\pi} \{\cos(j-m)x - \cos(j+m)x\} dx = \begin{cases} 0 & j \neq m, \\ \frac{1}{2} \int_{-\pi}^{\pi} dx - \frac{1}{2} \int_{-\pi}^{\pi} \cos(j+m)x dx = \pi, & j = m \neq 0, \\ \frac{1}{2} \int_{-\pi}^{\pi} dx - \frac{1}{2} \int_{-\pi}^{\pi} dx = 0, & j = m = 0; \end{cases}$$

$$\int_{-\pi}^{\pi} \cos mx \cos jx dx = \frac{1}{2} \int_{-\pi}^{\pi} \{\cos(j-m)x + \cos(j+m)x\} dx = \begin{cases} 0 & j \neq m, \\ \frac{1}{2} \int_{-\pi}^{\pi} dx - \frac{1}{2} \int_{-\pi}^{\pi} \cos(j+m)x dx = \pi, & j = m \neq 0, \\ \frac{1}{2} \int_{-\pi}^{\pi} dx + \frac{1}{2} \int_{-\pi}^{\pi} dx = 2\pi, & j = m = 0. \end{cases}$$

□

Appendix A.3.1.6: Bessel's ongelijkheid

Laat a_k , b_k en $S_n(x)$ gedefinieerd zijn door (3.1.5), (3.1.6) en (3.1.7) respectievelijk. Als $f(x)$ gedefinieerd op $[-\pi, \pi]$ kwadratisch integreerbaar is, geldt:

$$\frac{1}{\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx \geq \sum_{k=0}^{\infty} (a_k^2 + b_k^2)$$

Bewijs :

$$0 \leq \int_{-\pi}^{\pi} (f(x) - S_n(x))^2 dx = \int_{-\pi}^{\pi} f^2(x) dx - 2 \int_{-\pi}^{\pi} f(x) S_n(x) dx + \int_{-\pi}^{\pi} S_n^2(x) dx =$$

, door definities van a_k , b_k en orthogonaliteit van sinus en cosinus functies,

$$= \int_{-\pi}^{\pi} f^2(x) dx - 2\pi \left\{ \sum_{k=0}^n (a_k^2 + b_k^2) \right\} + \pi \left\{ \sum_{k=0}^n (a_k^2 + b_k^2) \right\} \Leftrightarrow \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(x) dx \geq \sum_{k=0}^n (a_k^2 + b_k^2)$$

waarbij n arbitrair is

□



Appendix A.1.3.7: Klasse van functies waarvan Fourier-reeks bestaat

Stel $f(x)$ een continue periodieke functie van periode 2π met een afgeleide $f'(x)$, die kwadratisch integreerbaar is. Dan convergeert de goniometrische reeks (3.1.8) absoluut en uniform naar $f(x)$ en is daarmee de Fourier-reeks.

Bewijs:

Onder de assumpties wordt er partieel geïntegreerd, waarbij met (3.1.5):

$$a_h = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(hx) dx = \frac{1}{\pi h} [f(x) \sin(hx)]_{-\pi}^{\pi} - \frac{1}{\pi h} \int_{-\pi}^{\pi} f'(x) \sin(hx) dx = -\frac{1}{\pi h} \int_{-\pi}^{\pi} f'(x) \sin(hx) dx.$$

Daarmee gelden de volgende relaties: $a_h = -b'_h/h$, $b_h = a'_h/h$, ofwel:

$$a'_h = \frac{1}{\pi} \int_{-\pi}^{\pi} f'(x) \cos(hx) dx,$$

$$b'_h = \frac{1}{\pi} \int_{-\pi}^{\pi} f'(x) \sin(hx) dx.$$

Onder de voorwaarde dat $f'(x)$ kwadratisch integreerbaar is convergeren de bestaande coëfficiënten van de afgeleide als gevolg van Bessel's ongelijkheid (theorema 3.1.6), dus geldt:

$$\sum_{h=0}^{\infty} (|a'_h|^2 + |b'_h|^2) < \infty.$$

Omdat: $\left(a'_h - \frac{1}{h}\right)^2 = |a'_h|^2 - \frac{2}{h}|a'_h| + \frac{1}{h^2} > 0$, geldt:

$$\frac{1}{h}|a'_h| + \frac{1}{h}|b'_h| \leq \frac{1}{2} (|a'_h|^2 + |b'_h|^2) + \frac{1}{h^2}, \text{ waaruit volgt dat } \sum_{h=0}^{\infty} (|a_h| + |b_h|) \text{ convergeert.}$$

Nu geldt: $|a_h \cos(hx) + b_h \sin(hx)| \leq |a_h \cos(hx)| + |b_h \sin(hx)| \leq |a_h| + |b_h|$,

en de reeks $\sum_{h=0}^{\infty} (a_h \cos(hx) + b_h \sin(hx))$

convergeert naar $f(x)$ absoluut en continu met Weierstrass' M-test (theorema A.3.1.10).

□

Appendix A.3.1.8: Fourier-transformatie stelling

Neem aan dat de autocorrelatie functie van een tijdreeks absoluut sommeerbaar is.

Dan bestaat er een continue functie $f(\omega)$ zodanig dat:

$$i) \rho(h) = \int_{-\pi}^{\pi} f(\omega) \cos(\omega h) = \int_{-\pi}^{\pi} f(\omega) e^{i\omega h} d\omega,$$

$$ii) f(\omega) \geq 0,$$

$$iii) \int_{-\pi}^{\pi} f(\omega) d\omega = 1,$$

$$iv) f(\omega) \text{ is een even functie.}$$

Bewijs:

De functie $g(\omega) = \frac{1}{2} + \sum_{h=0}^{\infty} \rho(h) \cos(\omega h)$ is een goed gedefinieerde functie (stelling A.3.1.7).

Op basis van de niet-negativiteit eigenschap van de correlatie functie (stelling A.3.2.6) geldt:

$$\sum_{m=1}^n \sum_{q=1}^m \rho(m-q) \cos(m\omega) \cos(q\omega) \geq 0, \quad \sum_{m=1}^n \sum_{q=1}^m \rho(m-q) \sin(m\omega) \sin(q\omega) \geq 0.$$

Dus :

$$\sum_{m=1}^n \sum_{q=1}^n \rho(m-q) [\cos(m\omega)\cos(q\omega) + \sin(m\omega)\sin(q\omega)] = \sum_{m=1}^n \sum_{q=1}^n \rho(m-q) \cos[(m-q)\omega] \geq 0.$$

Definieer : $m-q = h$, nu geldt :

$$\sum_{h=-(n-1)}^{(n-1)} \left(\frac{n-|h|}{n} \right) \rho(h) \cos(h\omega) \geq 0,$$

$\rho(h) \cos(h\omega)$ is absoluut sommeerbaar en met Kronecker's lemma, theorema A.3.1.12 :

$$\lim_{n \rightarrow \infty} \sum_{h=-(n-1)}^{(n-1)} \frac{|h|}{n} \rho(h) \cos(h\omega) = 0. \text{ Daarom :}$$

$$\lim_{n \rightarrow \infty} \sum_{h=-(n-1)}^{(n-1)} \frac{n-|h|}{n} \rho(h) \cos(h\omega) = \sum_{h=-\infty}^{\infty} \rho(h) \cos(h\omega) = 2g(\omega) \geq 0, \text{ waarmee (ii) bewezen is.}$$

$$\int_{-\pi}^{\pi} g(\omega) d\omega = \int_{-\pi}^{\pi} \left[\frac{1}{2} + \sum_{h=1}^{\infty} \rho(h) \cos(h\omega) \right] d\omega = \frac{1}{2} \sum_{h \neq 0}^{\infty} \left[\frac{\rho(h) \sin(h\omega)}{h} \right]_{-\pi}^{\pi} + \frac{1}{2} \int_{-\pi}^{\pi} d\omega = \pi.$$

$$\text{Definieer : } f(\omega) = \frac{1}{\pi} g(\omega) = \frac{1}{\pi} \left[\frac{1}{2} + \sum_{h=1}^{\infty} \rho(h) \cos(h\omega) \right] = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \rho(h) \cos(h\omega),$$

waarmee (iii) bewezen is.

$$\text{Merk op : } f(\omega) = \frac{1}{2} \sum_{h=-\infty}^{\infty} \rho(h) \cos(h\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \rho(h) (\cos(h\omega) - i \sin(h\omega)) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \rho(h) e^{-i\omega h}.$$

De functie $f(\omega)$ is een even functie, omdat het de uniforme limiet is van een som van even cosinus - functies (cosinus).

Bewijs 3.1.8.i: de Inverse Transformatie Stelling

$$\int_{-\pi}^{\pi} f(\omega) e^{i\omega k} d\omega \stackrel{\text{bewijs(2.2.4)}}{=} \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{j=-\infty}^{\infty} \rho(j) e^{-i\omega j} e^{i\omega k} d\omega = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \rho(j) \int_{-\pi}^{\pi} e^{i\omega(k-j)} d\omega \stackrel{\text{appendixB}}{=}$$

$$\frac{1}{2\pi} \sum_j \rho(j) \int_{-\pi}^{\pi} \{ \cos[\omega(k-j)] + i \sin[\omega(k-j)] \} d\omega.$$

$$\text{Als geldt } k = j, \text{ dan } \int_{-\pi}^{\pi} \{ \cos[\omega(k-j)] + i \sin[\omega(k-j)] \} d\omega = \int_{-\pi}^{\pi} \{ \cos(0) + i \sin(0) \} d\omega = \int_{-\pi}^{\pi} d\omega = 2\pi.$$

$$\text{Als geldt } k \neq j, \text{ dan } \int_{-\pi}^{\pi} \{ \cos[\omega(k-j)] + i \sin[\omega(k-j)] \} d\omega = \frac{\sin[\omega(k-j)]}{k-j} \Big|_{\omega=-\pi}^{\pi} - i \frac{\cos[\omega(k-j)]}{k-j} \Big|_{\omega=-\pi}^{\pi} =$$

$$(k-j)^{-1} \{ \sin[\pi(k-j)] - \sin[-\pi(k-j)] - i \cos[\pi(k-j)] + i \cos[-\pi(k-j)] \} = (k-j)^{-1} 2 \sin[\pi(k-j)] = 0.$$

Dus :

$$\int_{-\pi}^{\pi} f(\omega) e^{i\omega k} d\omega = \frac{1}{2\pi} \rho(k) \int_{-\pi}^{\pi} \{ \cos(0) + i \sin(0) \} d\omega = \rho(k).$$

Tot slot bewijzen we de gelijkheid van de verschillende representaties :

$$\int_{-\pi}^{\pi} f(\omega) e^{i\omega k} d\omega = \int_{-\pi}^0 f(\omega) e^{i\omega k} d\omega + \int_0^{\pi} f(\omega) e^{i\omega k} d\omega = \int_0^{\pi} f(-\omega) e^{-i\omega k} d\omega + \int_0^{\pi} f(\omega) e^{i\omega k} d\omega =$$

$$\int_0^{\pi} f(\omega) (e^{i\omega k} + e^{-i\omega k}) d\omega = \int_0^{\pi} f(\omega) 2 \cos(\omega k) d\omega = \int_{-\pi}^{\pi} f(\omega) \cos(\omega k) d\omega.$$

□



Appendix A.3.1.9: Uniforme convergentie

Uniforme convergentie van $S_n(x) = \sum_{k=1}^n f_k(x)$ betekent :

$\forall \varepsilon > 0 \exists N$ zodanig dat $\forall n \geq N$ en $\forall x \in [a, b]$ is $|S_n(x) - S(x)| < \varepsilon$.

□

Appendix A.3.1.10: Weierstrass' M-test

Zij $S_n(x) = \sum_{k=1}^n f_k(x)$, waarvoor geldt : $S(x) = \lim_{n \rightarrow \infty} S_n(x)$.

Als de reeks positieve getallen $M_1 + M_2 + \dots + M_k + \dots$ convergeert en voor elke $x \in [a, b]$ geldt $f_k(x) \leq M_k$ vanaf een bepaalde k , dan convergeert de rij $S_n(x)$ uniform op $[a, b]$.

Bewijs :

Door de aanname geldt , gegeven $\varepsilon > 0$, dat er een N is zodanig dat $\sum_{j=N+1}^{\infty} M_j < \varepsilon$.

Dus, voor $n \geq N$ geldt

$$|S(x) - S_n(x)| = \left| \sum_{k=1}^{\infty} f_k(x) - \sum_{k=1}^n f_k(x) \right| = \left| \sum_{k=n+1}^{\infty} f_k(x) \right| \leq \left| \sum_{k=n+1}^{\infty} M_k \right| < \varepsilon$$

en omdat ε arbitrair is, volgt het resultaat.

□

Appendix A.3.1.11: Monotone convergentie (zonder bewijs)

Als $f_k(x)$ continu is op $[a, b]$ en de reeks van f_k uniform convergent is op $[a, b]$, dan geldt dat de som van de reeks continu is en dat:

$$\int_a^b \sum_{k=1}^{\infty} f_k(x) dx = \sum_{k=1}^{\infty} \int_a^b f_k(x) dx.$$

□

Appendix A.3.1.12: Kronecker's lemma

Veronderstel een reeks $\{a_j\}$ zodanig dat $\lim_{n \rightarrow \infty} \sum_{j=0}^n |a_j| = A < \infty$, dan geldt :

$$\lim_{n \rightarrow \infty} \sum_{j=0}^{\infty} \frac{j}{n} |a_j| = 0.$$

Bewijs :

Zij $\varepsilon > 0$. Door de aanname geldt : $\exists N$ zodanig dat $\sum_{j=N+1}^{\infty} |a_j| < \varepsilon$.

Daarom, $\forall n > N$ geldt : $\sum_{j=0}^n \frac{j}{n} |a_j| < \frac{1}{n} \sum_{j=0}^N j |a_j| + \varepsilon$.

Dus, voor vaste N : $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^N j |a_j| = 0$ en het resultaat volgt uit de willekeur van ε

□

Appendix A.3.2.5: Spectraaltransformatie van een algemeen lineair filter

Veronderstel x_t een stationaire tijdreeks met een absoluut sommeerbare covariantie - functie en veronderstel dat de deterministische rij $\{a_j\}_{j=-\infty}^{\infty}$ absoluut sommeerbaar is.

$$\text{Beschouw het algemene lineaire filter } F(x_t) = \sum_{j=-\infty}^{\infty} a_j L^j(x_t). \quad (3.2.10)$$

Dan is de spectraaldichtheid van $y_t = F(x_t) = \sum_{j=-\infty}^{\infty} a_j x_{t-j}$ gelijk aan :

$$f_y(\omega) = (2\pi)^2 f_x(\omega) f_a(\omega) f_a^*(\omega) = |T(\omega)|^2 f_x(\omega), \quad (3.2.11)$$

waarbij $f_x(\omega)$ het spectrum is van x_t en

$$f_a(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} a_j e^{-i\omega j} \quad \text{de Fourier - transformatie van } a_j \quad (3.2.12)$$

$$f_a^*(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} a_j e^{i\omega j} \quad \text{is de complex geconjugeerde Fourier transformatie of } a_j,$$

$$T(\omega) = \sum_{j=-\infty}^{\infty} a_j L^j(e^{i\omega t}) \quad \text{op } t=0 \quad \text{de bij filter } F \text{ behorende transfer - functie.} \quad (3.2.13)$$

Bewijs:

Gegeven de absolute sommeerbaarheid is het mogelijk om integratie en sommatie om te wisselen, waarmee de verwachting per afzonderlijke term te berekenen is. Daarom geldt bij aanname $E(X_t)=0$:

$$E(y_t y_{t+h}) = E\left\{ \left(\sum_{j=-\infty}^{\infty} a_j x_{t-j} \right) \left(\sum_{k=-\infty}^{\infty} a_k x_{t-k+h} \right) \right\} = E\left\{ \sum_{j=-\infty}^{\infty} a_j \sum_{k=-\infty}^{\infty} a_k x_{t-j} x_{t-k+h} \right\} = \sum_{j=-\infty}^{\infty} a_j \sum_{k=-\infty}^{\infty} a_k \gamma_x(h-k+j).$$

$$\text{Definieer: } p = h-k + j, \text{ dan geldt: } \gamma_y(h) = \sum_{j=-\infty}^{\infty} a_j \sum_{p=-\infty}^{\infty} a_{j+h-p} \gamma_x(p).$$

Omdat $\gamma_y(h)$ absoluut sommeerbaar is, geldt:

$$\begin{aligned} f_y(\omega) &= \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-i\omega h} \gamma_y(h) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-i\omega h} \sum_{j=-\infty}^{\infty} a_j \sum_{p=-\infty}^{\infty} a_{j+h-p} \gamma_x(p) \\ &= \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} a_j a_{j+h-p} \gamma_x(p) e^{i\omega j} e^{-i\omega p} e^{-i\omega(j+h-p)} = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} a_j e^{i\omega j} \sum_{p=-\infty}^{\infty} \gamma_x(p) e^{-i\omega p} \sum_{s=-\infty}^{\infty} a_s e^{-i\omega s}, \end{aligned}$$

waarbij gedefinieerd is: $s = j + h - p$. Het bewijs voor de transfer - functie loopt als volgt:

$$T(\omega) = \sum_{j=-\infty}^{\infty} a_j L^j(e^{i\omega t}) = \sum_{j=-\infty}^{\infty} a_j e^{i\omega(t-j)} = e^{i\omega t} \sum_{j=-\infty}^{\infty} a_j e^{-i\omega j} \stackrel{t=0}{=} \sum_{j=-\infty}^{\infty} a_j e^{-i\omega j}.$$

□

Appendix A.3.2.6: De niet-negativiteit van de covariantiefunctie van een stationaire tijdreeks

De covariantiefunctie van een stationaire tijdreeks $\{T_t; t \in T\}$ is niet-negatief definitief.

Bewijs:

Zonder verlies van algemeenheid, stel $E\{x_t\}=0$. Zij $\{t_1, t_2, \dots, t_n\} \in T$ en zij (a_1, a_2, \dots, a_n) een willekeurige verzameling reële getallen. Stel $\gamma(t_j - t_k)$ is de covariantie tussen x_{t_j} en x_{t_k} . De variantie van een stochastische variabele is, indien gedefinieerd, niet-negatief. Daarom geldt:

$$0 \leq \text{Var} \left\{ \sum_{j=1}^n a_j x_{t_j} \right\} = E \left\{ \sum_{j=1}^n \sum_{k=1}^n a_j a_k x_{t_j} x_{t_k} \right\} = \sum_{j=1}^n \sum_{k=1}^n a_j a_k \gamma(t_j - t_k)$$

waarmee het gestelde aangetoond is.

□

Appendix A.3.2.7: Een gevolg van absoluut sommeerbare coëfficiënten voor covariantiesom

$$y_t = \mu + \psi(L)e_t,$$

met: $e_t \sim N(0, \sigma^2)$ o.i.v. en $\psi(L) = \psi_0 + \psi_1 L + \psi_2 L^2 + \dots$, waarbij $\sum_j |\psi_j| < \infty$.

Dan geldt: $\sum_{j=0}^{\infty} |\gamma(j)| \leq M < \infty$.

Bewijs:

$$\gamma(j) = E(y_t - \mu)(y_{t-j} - \mu) = \sigma^2(\psi_j \psi_0 + \psi_{j+1} \psi_1 + \psi_{j+2} \psi_2 + \dots) = \sigma^2 \sum_{k=0}^{\infty} \psi_{j+k} \psi_k.$$

En dus:

$$\sum_{j=0}^{\infty} |\gamma(j)| = \sum_{j=0}^{\infty} \sigma^2 \left| \sum_{k=0}^{\infty} \psi_{j+k} \psi_k \right| \leq \sigma^2 \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} |\psi_{j+k} \psi_k| = \sigma^2 \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} |\psi_{j+k}| |\psi_k| = \sigma^2 \sum_{k=0}^{\infty} |\psi_k| \left| \sum_{j=0}^{\infty} |\psi_{j+k}| \right|.$$

Met assumptie volgt:

$$\sigma^2 \sum_{k=0}^{\infty} |\psi_k| \left| \sum_{j=0}^{\infty} |\psi_{j+k}| \right| \leq \sigma^2 \sum_{k=0}^{\infty} |\psi_k| M \leq \sigma^2 M^2 < \infty,$$

waarmee het gestelde volgt. □

Appendix A.3.3.1: Niet-negativiteit van multivariate spectraalanalyse

$f(\omega)$ is een positief semi-definiete matrix.

Bewijs:

Zij $\underline{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_p)$ een rij met arbitraire constanten. Beschouw het univariate proces $y_t = \underline{\lambda} x_t$,

De autocovariantie-functie van y_t is: $\gamma_y(s) = E[y_{t+s} y_t^*] = \underline{\lambda} E[x_{t+s} x_t^*] \underline{\lambda}^* = \underline{\lambda} \Gamma(s) \underline{\lambda}^*$

en de spectraalfunctie van y_t :

$$f_y(\omega) = \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} \gamma_y(s) e^{-i\omega s} = \underline{\lambda} f(\omega) \underline{\lambda}^*.$$

Omdat de univariate $f_y(\omega) \geq 0$ (theorem 3.1.8) voor iedere $\underline{\lambda}$ volgt het resultaat. □

Appendix A.3.3.2: Sommatie van sinus en cosinus functies

$$\sum_{t=1}^T \cos[\omega_j(t-1)] = \sum_{t=1}^T \sin[\omega_j(t-1)] = 0$$

voor $j = 1, 2, \dots, M$ waarbij T oneven, $M = (T-1)/2$ en $\omega_j = \frac{2\pi j}{T}$

Bewijs:

De beweringen gelden als de volgende relatie bewezen is:

$$0 = \sum_{t=1}^T e^{i2\pi \left(\frac{j}{T}\right)(t-1)} = \sum_{t=1}^T \cos\left[\left(\frac{2\pi j}{T}\right)(t-1)\right] + i \sum_{t=1}^T \sin\left[\left(\frac{2\pi j}{T}\right)(t-1)\right]$$

Beschouw de volgende simplificatie:

$$\sum_{t=1}^T e^{i2\pi \left(\frac{j}{T}\right)(t-1)} = \sum_{t=1}^T z^{(t-1)} = \frac{1-z^T}{1-z}, \text{ o.v.d. } z \neq 1, \text{ wat klopt voor } 0 < |j| < T.$$

Bovendien geldt: $z^T = e^{i2\pi \left(\frac{j}{T}\right)T} = e^{i2\pi j} = \cos(2\pi j) + i \sin(2\pi j) = 1$, voor $j = \pm 1, \pm 2, \dots, \pm(T-1)$. □

Appendix A.3.3.3: afleidingen bij dit voorbeeld 3.9

Af te leiden relatie :

$$f_{yx}(\omega_j) = \frac{N}{8\pi} ((\alpha_j + i\delta_j)(\alpha_j - i\delta_j)) = \frac{N}{8\pi} (a_j\alpha_j + d_j\delta_j) + i \frac{N}{8\pi} (d_j\alpha_j - a_j\delta_j) = c_{yx}(\omega_j) + iq_{yx}(\omega_j)$$

De afleiding van het eerste '='-teken gaat als volgt :

$$\begin{aligned} (\alpha_j + i\delta_j)(\alpha_j - i\delta_j) &= \frac{4}{N^2} \left\{ \sum_{t=1}^N (x_t - \mu_x) \cos[\omega_j(t-1)] + i \sum_{t=1}^N (x_t - \mu_x) \sin[\omega_j(t-1)] \right\} \\ &\times \left\{ \sum_{\tau=1}^N (y_\tau - \mu_y) \cos[\omega_j(\tau-1)] + i \sum_{\tau=1}^N (y_\tau - \mu_y) \sin[\omega_j(\tau-1)] \right\} \\ &= \frac{4}{N^2} \left\{ \left(\sum_{t=1}^N (x_t - \mu_x) e^{i\omega_j(t-1)} \right) \left(\sum_{\tau=1}^N (y_\tau - \mu_y) e^{-i\omega_j(\tau-1)} \right) \right\} \\ &= \frac{4}{N^2} \left\{ \sum_{t=1}^N (x_t - \mu_x)(y_t - \mu_y) + \sum_{t=1}^{N-1} (x_t - \mu_x)(y_{t+1} - \mu_y) e^{-i\omega_j} + \sum_{t=1}^{N-2} (x_t - \mu_x)(y_{t+2} - \mu_y) e^{-2i\omega_j} + \dots + \right. \\ &\quad \left. (x_t - \mu_x)(y_T - \mu_y) e^{-i(N-1)\omega_j} + (x_1 - \mu_x)(y_1 - \mu_y) e^{i(N-1)\omega_j} \right\} \\ &= \frac{4}{N} \left\{ \gamma_{yx}(0) + \gamma_{yx}(1) e^{-i\omega_j} + \gamma_{yx}(-1) e^{i\omega_j} + \dots + \gamma_{yx}(-(N-1)) e^{(N-1)i\omega_j} + \gamma_{yx}(N-1) e^{-i(N-1)\omega_j} \right\} \\ &= \frac{4}{N} \sum_{k=-(T-1)}^{(T-1)} \gamma_{yx}(k) e^{-ki\omega_j} = \frac{8\pi}{N} f_{yx}(\omega_j) \text{ waarmee het gestelde volgt.} \end{aligned}$$

□

Appendix A.3.3.4: Kwadratische coherentie niet groter dan een

$$K_{xy}^2 \leq 1.$$

Bewijs :

Uit theorema 3.1.8 en het determinanten - criterium voor niet - negatieve functies volgt :

$$\det \begin{bmatrix} f_{xx}(\omega) & f_{xy}(\omega) \\ f_{yx}(\omega) & f_{yy}(\omega) \end{bmatrix} \geq 0, \text{ dat wil zeggen : } |f_{xy}(\omega)|^2 = f_{xy}(\omega) f_{xy}^*(\omega) = f_{xy}(\omega) f_{yx}(\omega) \leq f_{xx}(\omega) f_{yy}(\omega)$$

$$\text{Dus geldt : } K_{xy}^2 = \frac{|f_{xy}(\omega)|^2}{f_{xx}(\omega) f_{yy}(\omega)} \leq 1.$$

□

Appendix A.3.3.5: Transferfunctie als beste lineaire benadering

Om aan te tonen dat de transferfunctie, $\frac{1}{2\pi} \frac{f_{xy}(\omega)}{f_{yy}(\omega)}$ de minimale waarde genereert,

$$\text{beschouw een alternatieve kandidaat - oplossing : } f_h(\omega) = \frac{1}{2\pi} \left[\frac{f_{xy}(\omega)}{f_{yy}(\omega)} + \frac{f_d(\omega)}{2\pi f_{yy}(\omega)} \right],$$

waarbij $f_d(\omega)$ element uit de klasse van functies waarvan de integraal bestaat. Dan geldt :

$$\int_{-\pi}^{\pi} \left\{ f_{xx}(\omega) - 2\pi [f_h(\omega) f_{xy}(\omega) + f_h^*(\omega) f_{yx}(\omega)] + (2\pi)^2 f_h(\omega) f_h^*(\omega) f_{yy}(\omega) \right\} d\omega =$$

$$\int_{-\pi}^{\pi} \left[f_{xx}(\omega) - \frac{f_{xy}(\omega) f_{xy}^*(\omega)}{f_{yy}(\omega)} + \frac{f_d(\omega) f_d^*(\omega)}{f_{yy}(\omega)} \right] d\omega \text{ en omdat } \frac{f_d(\omega) f_d^*(\omega)}{f_{yy}(\omega)} \text{ niet - negatief is volgt dat}$$

$f_g(\omega)$ de minimale functie is.

□

Appendix A.3.4.3: Niet convergerende variantie van het periodogram

Veronderstel een tijdreeks x_t , waarbij: $x_t = \sum_{j=0}^{\infty} a_j e_{t-j}$ met $e_t \sim N(0, \sigma^2)$. o.i.v

en vierde moment $\eta\sigma^4$. Bovendien $\sum_{j=1}^{\infty} j^{1/2} |a_j| < \infty$.

Dan geldt:

$$\text{Cov}\{I_n(\omega_j), I_n(\omega_k)\} = \begin{cases} 2(4\pi)^2 f^2(0) + 0(n^{-1}) & \omega_j = \omega_k = 0 \\ (4\pi)^2 f^2(\omega_k) + 0(n^{-1}) & \omega_j = \omega_k \neq 0, \pi \\ 0(n^{-1}) & \omega_j \neq \omega_k \end{cases}$$

Bewijs:

Wegens de definitie van het periodogram geldt

$$I_n(\omega_k) = \frac{2}{n} \sum_{t=1}^n \sum_{s=1}^n x_t x_s e^{i\omega_k(t-s)} \quad \text{en}$$

$$E\{I_n(\omega_k)I_n(\omega_j)\} - E\{I_n(\omega_k)\}E\{I_n(\omega_j)\} =$$

$$E\left\{\frac{4}{n^2} \sum_{t=1}^n \sum_{s=1}^n \sum_{u=1}^n \sum_{v=1}^n x_t x_s x_u x_v e^{i\omega_k(t-s)} e^{i\omega_j(u-v)}\right\} - 4 \sum_{h=-(n-1)}^{(n-1)} \frac{n-|h|}{n} \gamma(h) e^{-i\omega_k h} \sum_{q=-(n-1)}^{(n-1)} \frac{n-|q|}{n} \gamma(q) e^{-i\omega_j q}.$$

Gebruikmakend van

$$E\{e_s e_t e_u e_v\} = \begin{cases} \eta\sigma^4, & s = t = u = v \\ \sigma^4, & \text{als de onderschriften gelijk zijn in paren, maar niet allemaal gelijk} \\ 0, & \text{anders} \end{cases}$$

en $a_r = 0, \quad r < 0$

geldt

$$= \frac{4(\eta-3)\sigma^4}{n^2} \sum_{t=1}^n \sum_{s=1}^n \sum_{u=1}^n \sum_{v=1}^n \sum_{r=-\infty}^{\infty} a_r a_{r+t-s} a_{r+t-u} a_{r+t-v} e^{i\omega_k(t-s)} e^{i\omega_j(u-v)}$$

$$+ \frac{4}{n^2} \sum_{t=1}^n \sum_{s=1}^n \sum_{u=1}^n \sum_{v=1}^n \gamma(t-u) \gamma(s-v) e^{i\omega_k(t-u)} e^{i\omega_j(s-v)} e^{i(\omega_j + \omega_k)(u-s)}$$

$$+ \frac{4}{n^2} \sum_{t=1}^n \sum_{s=1}^n \sum_{u=1}^n \sum_{v=1}^n \gamma(t-v) \gamma(u-s) e^{i\omega_k(t-v)} e^{i\omega_j(u-s)} e^{i(\omega_j - \omega_k)(s-v)}. \quad (*)$$

Bovendien:

$$\frac{4^4}{n^2} \left| \sum_{t=1}^n \sum_{s=1}^n \sum_{u=1}^n \sum_{v=1}^n \sum_{r=-\infty}^{\infty} a_r a_{r+t-s} a_{r+t-u} a_{r+t-v} e^{i\omega_k(t-s)} e^{i\omega_j(u-v)} \right| \leq \frac{4}{n} \sum_{j=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \sum_{h=-\infty}^{\infty} |a_j| |a_{j+p}| |a_{j+p+q}| |a_{j+p+q+h}|$$

$$= 0(n^{-1}) \text{ door de absolute sommeerbaarheid van } a_j.$$

Als geldt $\omega_j = \omega_k$, of als n even is en $\omega_j = \omega_k = \pi$, dan is de tweede term van (*):

$$\frac{4}{n^2} \left[\sum_{h=-(n-1)}^{(n-1)} (n-|h|) \gamma(h) \right]^2.$$

Voor $\omega_k = \omega_j \neq 0$, de tweede term van (*) is:

$$\frac{4}{n^2} \sum_{t=1}^n \sum_{s=1}^n \sum_{u=1}^n \sum_{v=1}^n \gamma(t-u) \gamma(s-v) e^{i\omega_k(t-u+ru) - i(\omega_k(s-v+rv))} \leq \left[2 \sum_{p=-(n-1)}^{(n-1)} \frac{|p|}{n} |\gamma(p)| \right]^2 = 0(n^{-1})$$

waarbij de ongelijkheid volgt uit appendix A.3.4.5 Voor $\omega_k \neq \omega_j$, de derde term vermindert tot:

$$\left[2 \sum_{p=-(n-1)}^{(n-1)} \frac{n-|p|}{n} |\gamma(p)| e^{i\omega_k p} \right]^2. \quad \text{Door appendix A.3.4.5: voor } \omega_k \neq \omega_j :$$

$$\left| \sum_{t=1}^n \sum_{u=1}^n \gamma(t-u) e^{i\omega_k(t-u)} e^{i(\omega_j+\omega_k)u} \right| \leq \sum_{p=-(n-1)}^{(n-1)} |p| |\gamma(p)|$$

en de absolute waarde van de tweede term hierboven is kleiner dan $\left[\frac{2}{n} \sum_{p=-(n-1)}^{(n-1)} |p| |\gamma(p)| \right]^2$, welke is $o(n^{-1})$ met

appendix A.3.4.6. Door eenzelfde argument geldt voor de derde term van (*) ook $o(n^{-1})$ als $\omega_k \neq \omega_j$.

Dus, als $\omega_k = \omega_j \neq 0, \pi$

$$\lim_{n \rightarrow \infty} \frac{4}{n^2} \sum_{t=1}^n \sum_{s=1}^n \sum_{u=1}^n \sum_{v=1}^n \gamma(t-v) \gamma(u-s) e^{i\omega_k(n,w)(t-v)} e^{i\omega_k(n,w)(u-s)} = \lim_{n \rightarrow \infty} \left[2 \sum_{p=-(n-1)}^{(n-1)} \frac{n-|p|}{n} |\gamma(p)| e^{-i\omega p} \right]^2 = (4\pi)^2 f^2(\omega)$$

$$\text{Als } \omega_k = \omega_j = 0, \text{ dan } \lim_{n \rightarrow \infty} \frac{4}{n^2} \sum_{t=1}^n \sum_{s=1}^n \sum_{u=1}^n \sum_{v=1}^n \gamma(t-v) \gamma(u-s) + \lim_{n \rightarrow \infty} \frac{4}{n^2} \sum_{t=1}^n \sum_{s=1}^n \sum_{u=1}^n \sum_{v=1}^n \gamma(t-u) \gamma(s-v) =$$

$$2 \lim_{n \rightarrow \infty} \left[2 \sum_{p=-(n-1)}^{(n-1)} \frac{n-|p|}{n} |\gamma(p)| e^{-i\omega p} \right]^2 = (4\pi)^2 f^2(0), \text{ waaruit de gestelde resultaten volgen.}$$

□

Appendix A.3.4.4: De variantie van covarianties

Veronderst el dat de tijdreeks $\{x_t\}$ gedefinieerd is door :

$$x_t = \sum_{j=-\infty}^{\infty} a_j e_{t-j}, \text{ waarbij de reeks } \{a_j\} \text{ absoluut sommerbaar is en de } e_t \sim \mathcal{N}(0, \sigma^2) \text{ o.i.v.,}$$

met $E\{e_t^4\} = \eta\sigma^4$. Dan geldt voor vaste h en g :

$$\text{var}\{\hat{\varphi}(h)\} \leq \frac{K}{n}, \text{ waarbij } K < \infty.$$

Bewijs :

$$E\{e_s e_t e_u e_v\} = \begin{cases} \eta\sigma^4, & s = t = u = v \\ \sigma^4, & \text{als de onderschreven gelijk zijn in paren, maar niet allemaal gelijk,} \\ 0, & \text{anders} \end{cases}$$

hieruit volgt dat :

$$E\{x_t x_{t+h} x_{t+h+p} x_{t+h+p+q}\} = (\eta-3)\sigma^4 \sum_{j=-\infty}^{\infty} a_j a_{j+h} a_{j+h+p} a_{j+h+p+q} + \gamma(h)\gamma(q) + \gamma(h+p)\gamma(q+p) + \gamma(h+p+q)\gamma(p).$$

De covariantie wordt op de volgende manier afgeleidt :

$$\text{cov}\{\hat{\varphi}(h), \hat{\varphi}(q)\} = E\{\hat{\varphi}(h)\hat{\varphi}(q)\} - E\{\hat{\varphi}(h)\}E\{\hat{\varphi}(q)\} = \frac{1}{(n-h)(n-q)} \sum_{s=1}^{n-q} \sum_{t=1}^{n-h} E\{x_t x_{t+h} x_s x_{s+q}\} - \gamma(h)\gamma(q) =$$

$$\frac{1}{(n-h)(n-q)} \sum_{s=1}^{n-q} \sum_{t=1}^{n-h} \left\{ (\eta-3)\sigma^4 \sum_{j=-\infty}^{\infty} a_j a_{j+h} a_{j+s-t} a_{j+s-t+q} + \gamma(s-t)\gamma(s-t-h+q) + \gamma(s-t+q)\gamma(s-t-h) \right\}$$

Merk op dat met $p = (s-t)$:

$$\text{Var}\{\hat{\varphi}(q)\} = \frac{1}{(n-q)} \sum_{p=-(n-q-1)}^{(n-q-1)} \left\{ 1 - \frac{|p|}{n-q} \right\} \left\{ (\eta-3)\sigma^4 \sum_{j=-\infty}^{\infty} a_j a_{j+h} a_{j+s-t} a_{j+s-t+q} + \gamma^2(p) + \gamma(p+q)\gamma(p-q) \right\} \leq \frac{K}{n}$$

omdat de coëfficiëntenreeks sommerbaar is.

□

Appendix A.3.4.5: Bovengrens sommeerbaarheid

Voor reële L geldt :

$$\left| \sum_{t=1}^n \sum_{j=1}^n \gamma(t-j) \cos\{\omega_k(L+j)\} \right| \leq \sum_{p=-(n-1)}^{(n-1)} |p| \gamma(p), \quad \left| \sum_{t=1}^n \sum_{j=1}^n \gamma(t-j) \sin\{\omega_k(L+j)\} \right| \leq \sum_{p=-(n-1)}^{(n-1)} |p| \gamma(p),$$

waarbij : $\omega_k = \frac{2\pi k}{n}$, $k = 1, 2, \dots, (n-1)$.

Bewijs :

$$\sum_{j=1}^n \cos\{\omega_k(L+j)\} = \sum_{j=1}^n \{\cos(\omega_k L) \cos(\omega_k j) - \sin(\omega_k L) \sin(\omega_k j)\} =$$

$$\cos(\omega_k L) \sum_{j=1}^n \cos(\omega_k j) - \sin(\omega_k L) \sum_{j=1}^n \sin(\omega_k j) = 0, \quad \text{voor } \omega_k \neq 0.$$

Met behulp van bovenstaande relatie geldt :

$$\left| \sum_{j=1}^{n-p} \cos\{\omega_k(L+j)\} \right| = \left| \sum_{j=1}^n \cos\{\omega_k(L+j)\} - \sum_{j=n-p+1}^n \cos\{\omega_k(L+j)\} \right| \leq \left| \sum_{i=1}^p \cos\{\omega_k(L+i+n-p)\} \right| \leq$$

$$\sum_{i=1}^p |\cos\{\omega_k(L+i+n-p)\}| \leq p.$$

Hieruit volgt het gestelde :

$$\left| \sum_{t=1}^n \sum_{j=1}^n \gamma(t-j) \cos\{\omega_k(L+j)\} \right| = \left| \sum_{p=0}^{n-1} \sum_{j=1}^{n-p} \gamma(p) \cos\{\omega_k(L+j)\} + \sum_{p=-(n-1)}^{-1} \sum_{j=-p+1}^n \gamma(p) \cos\{\omega_k(L+j)\} \right| \leq \sum_{p=-(n-1)}^{(n-1)} |p| \gamma(p).$$

□

Appendix A.3.4.6: Voldoende voorwaarde voor sommeerbaarheid van covariantiereeks

Stel als voorwaarde voor de covarianties dat:

$$\sum_{h=-n}^n |h| \gamma(h) = O(n^{-\frac{1}{2}})$$

Dit is een bescheiden aanname en er is aan voldaan door elk eindig autoregressief proces.

Een voldoende voorwaarde is als voor een tijdreeks x_t :

$$x_t = \sum_{j=0}^{\infty} a_j e_{t-j} \quad \text{waarbij } e_t \sim N(0, \sigma^2) \text{ i.i.d. geldt } \sum_{j=0}^{\infty} j^{\frac{1}{2}} |a_j| < \infty.$$

Dit komt omdat :

$$n^{-\frac{1}{2}} \sum_{h=0}^n |h| \gamma(h) \leq n^{-\frac{1}{2}} \sum_{h=0}^n \sum_{j=0}^{\infty} |h| |a_j a_{j+h}| \sigma^2 \leq \sum_{h=0}^n \sum_{j=0}^{\infty} h^{\frac{1}{2}} |a_j| |a_{j+h}| \sigma^2 \leq \sum_{h=0}^{\infty} \sum_{j=0}^{\infty} (j+h)^{\frac{1}{2}} |a_j| |a_{j+h}| \sigma^2 \leq$$

$$\sigma^2 \sum_{j=0}^{\infty} |a_j| \sum_{s=0}^{\infty} s^{\frac{1}{2}} |a_s|.$$

□

Appendix A.3.5.1: Enkele goniometrische relaties

$\forall j \in \mathbb{Z}$:

$$\sin(2\pi j + \theta) = \sin(\theta)$$

$$\cos(\theta) = \sin(\theta + \pi/2)$$

$$\sin(-\theta) = -\sin(\theta)$$

$$\cos(-\theta) = \cos(\theta)$$

$$\sin(\theta + \eta) = \sin(\theta) \cos(\eta) + \cos(\theta) \sin(\eta)$$

$$\cos(\theta + \eta) = \cos(\theta) \cos(\eta) - \sin(\theta) \sin(\eta)$$

Appendix A.3.5.2: Cosinusommatie

$$\sum_{s=-m}^m \cos(s\vartheta) = \left\{ \frac{\sin\left(m + \frac{1}{2}\right)\vartheta}{\sin\left(\frac{\vartheta}{2}\right)} \right\}$$

Bewijs:

$$\sum_{s=-m}^m \cos(s\vartheta) = 2 \left(\frac{1}{2} + \cos\vartheta + \cos(2\vartheta) + \dots + \cos(m\vartheta) \right)$$

Definieer: $S = \frac{1}{2} + \cos\vartheta + \cos(2\vartheta) + \dots + \cos(m\vartheta)$

Dan geldt:

$$2S \sin\left(\frac{\vartheta}{2}\right) = \sin\left(\frac{\vartheta}{2}\right) + 2 \cos\vartheta \sin\left(\frac{\vartheta}{2}\right) + 2 \cos(2\vartheta) \sin\left(\frac{\vartheta}{2}\right) + \dots + 2 \cos(m\vartheta) \sin\left(\frac{\vartheta}{2}\right)$$

Door toepassen van de formule $2\cos\alpha\sin\beta = \sin(\alpha + \beta) - \sin(\alpha - \beta)$ geldt:

$$2S \sin\left(\frac{\vartheta}{2}\right) = \sin\left(\frac{\vartheta}{2}\right) + \left(\sin\left(\frac{3}{2}\vartheta\right) - \sin\left(\frac{\vartheta}{2}\right)\right) + \dots + \left(\sin\left(m + \frac{1}{2}\right)\vartheta\right) - \sin\left(m - \frac{1}{2}\right)\vartheta = \sin\left(m + \frac{1}{2}\right)\vartheta$$

daarom: $S = \frac{\sin\left(m + \frac{1}{2}\right)\vartheta}{2 \sin\left(\frac{\vartheta}{2}\right)}$

□

Appendix A.3.5.3: Complexe representatie van goniometrische reeksen

Toepassen van: $e^{-i\vartheta} = \cos(\vartheta) - i \sin(\vartheta)$, waarbij: $i = \sqrt{-1}$, geeft:

$$\cos(\vartheta) = \frac{\cos(\vartheta) + i \sin(\vartheta) + \cos(\vartheta) - i \sin(\vartheta)}{2} = \frac{e^{i\vartheta} + e^{-i\vartheta}}{2},$$

$$\sin(\vartheta) = \frac{\cos(\vartheta) + i \sin(\vartheta) - (\cos(\vartheta) - i \sin(\vartheta))}{2i} = \frac{(-i)(e^{i\vartheta} - e^{-i\vartheta})}{2},$$

$$a_k \cos(kx) + b_k \sin(kx) = a_k \left(\frac{e^{ikx} + e^{-ikx}}{2} \right) - b_k \left(\frac{(-i)(e^{i\vartheta} - e^{-i\vartheta})}{2} \right),$$

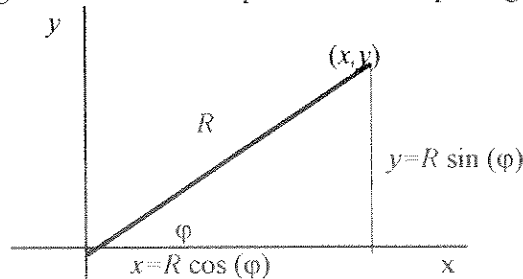
$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx)) = \sum_{k=-n}^n c_k e^{ikx},$$

waarbij: $c_k = \frac{a_k - ib_k}{2}$, $c_{-k} = c_k^* = \frac{a_k + ib_k}{2}$.

□

A.3.5.4. Poolcoördinaten representatie

Beschouw de volgende metrische interpretatie van complexe getallen in poolcoördinaten.



Hierbij geldt:

Amplitude $|r| = \sqrt{x^2 + y^2}$, fase(hoek) $\varphi = \text{Arg}(z) \Leftrightarrow \frac{\sin(\varphi)}{\cos(\varphi)} = \frac{y}{x} \Leftrightarrow \varphi = \tan^{-1}\left(\frac{y}{x}\right)$.

$$z = x + iy = r \cos(\varphi) + ir \sin(\varphi) = r(\cos(\varphi) + i \sin(\varphi)) = re^{i\varphi} \text{ en}$$

$$zz^* = (x + iy)(x - iy) = x^2 - i^2 y^2 = x^2 + y^2 = |z|^2$$



Appendix B: Documentatie procedures en resultaten

In deze appendix worden de tabellen gepresenteerd met gedetailleerde informatie over de conjunctuurbepaling en de constructie van de conjunctuurvoorspeller met deterministische trend en de resultaten van de Dickey-Fuller testprocedures voor de bepaling van de orde van de stochastische trend en de selectie van de reeksen met spectraalanalyse.

B.1: Documentatie procedure en resultaten deterministische werkwijze

In deze appendixparagraaf wordt het PAT-algoritme gedetailleerd uitgelegd inclusief de invloed van de instelbare parameters van dit algoritme op de uiteindelijke vorm van de PAT-trend. De instellingen van de uiteindelijke PAT-parameters staan voor de huizenprijs in tabel B.1 en voor de overige reeksen in tabel B.3 gedocumenteerd. In deze laatste tabel staan tevens de uitkomsten van de prestatimaatstaven getabelleerd, die een indicatie geven over de kwaliteit van de variabele als leidende indicator. In subparagraaf B.1.2 staan tenslotte de procedure en resultaten van de constructie van de conjunctuurvoorspeller neergeschreven.

B.1.1: Conjunctuur bepalen met deterministische trend

Het PAT-algoritme dat is gebruikt om de deterministische trend in de tijdreeksen te bepalen kent een aantal parameters, waarmee het aantal trendbuigpunten en de ligging van de trendbuigpunten in beperkte mate stuurbaar zijn. Het PAT-algoritme en de bijbehorende conjunctuurbepalingsmethode zijn geprogrammeerd in Visual Basic met toepassing in MS-Excel. Het doel van deze paragraaf is om de parameters van het PAT-algoritme te documenteren en anderzijds inzicht te verschaffen in de effecten op de trend van de afzonderlijke parameterwaarden. De parameters zijn voor de huizenprijsreeks in tabel B.1 en voor de overige reeksen in tabel B.3 weergegeven onder het kopje 'PAT-instellingen'. De getabelleerde parameters worden in deze paragraaf onderstreept weergegeven.

De input in het PAT-algoritme is een voor seizoensinvloeden gecorrigeerde tijdreeks. Seizoenscorrectie vindt plaats als de optie 'Seizoenscorrectie' aanstaat en geschiedt volgens de procedure van de seizoenscorrectie van het programma SPSS³⁴.

De eerste stap is het bepalen van een initiële trendschatting in de vorm van een gecentreerd voortschrijdend gemiddelde van een instelbaar MA1 aantal perioden. De verhouding van de oorspronkelijke reeks en de initiële trendschatting wordt ook met een instelbaar, MA2, voortschrijdend gemiddelde geëffend en stelt de initiële expansie- en contractieperioden vast, die fasen genoemd worden.

Met deze fasenindex kunnen initiële conjunctuuromslagpunten worden bepaald. Een lokaal extremum moet voldoende extreem zijn, minimaal zz%, om dienst te doen als conjunctuuromslagpunt. Zo is een lokaal maximumpunt pas een toppunt als er een waarde wordt gevonden die kleiner is dan $(1-zz) \cdot \text{maximumwaarde}$ onder de voorwaarde dat er tussen dit laatste punt en het toppunt geen hogere waarde dan de maximumwaarde ligt.

Met de optie gebruik beginpunt/eindpunt worden tevens de eerste en de laatste waarneming beschouwd als conjunctuuromslagpunten. Deze optie kan samen met de zz%-waarde het aantal omslagpunten enigszins manipuleren en daarmee het aantal fasen en uiteindelijk het aantal conjunctuurgolven.

Aan de hand van de fasen en de conjunctuuromslagpunten worden de PAT-omslagpunten bepaald. De waarneming waar het eerste en het laatste PAT-omslagpunt liggen zijn getabelleerd onder respectievelijk 1e plaats PAT top/dal en laatste plaats PAT top/dal. Het aantal PAT-omslagpunten is tevens weergegeven met aantal PAT top/dal.

³⁴ Dit is de zogenaamde Census Method I. Zie voor een uitgebreide beschrijving Makridakis, Wheelwright & McGee, 1983

Om de trendlijn over de gehele tijdreeks te bepalen wordt er geëxtrapoleerd vanaf deze punten, waarvoor een aantal opties beschikbaar zijn. Deze opties staan in de tabel vermeld onder de items opties beginpunt en opties eindpunt en luiden als volgt:

- 0) Geen extrapolatie.
- 1) Eerste/laatste punt reeks;
er wordt een rechte lijn getrokken tussen het middenpunt en het eerste/laatste datapunt van de tijdreeks.
- 2) Basistrend;
vanaf het middenpunt wordt een rechte lijn getrokken met een hellingshoek, die gelijk is aan de hellingshoek van een lineaire regressielijn over de hele reeks.
- 3) Gemiddelde eerste/laatste en basistrend;
een combinatie van de lijnen verkregen uit optie 1 en optie 2.
- 4) Eerste/laatste lijntrend;
met deze optie wordt de rechte lijn tussen het op een na laatste middenpunt en het laatste middenpunt doorgetrokken tot het einde van de tijdreeks.
- 5) Laatste datatrend;
voor de data vanaf het laatste middenpunt tot het einde van de tijdreeks wordt een lineaire regressielijn berekend en bovendien wordt de laatste lijntrend doorgetrokken (optie 4). Het snijpunt tussen deze twee lijnen bepaald dan het laatste omslagpunt van de trendlijn. Indien er geen snijpunt bestaat wordt de lineaire regressielijn bepaald over een datapunt extra tot een instelbaar maximum. Dit maximum wordt in de tabel aangegeven achter het '/'-teken. Als er dan nog geen snijpunt is wordt automatisch optie 4 ingeschakeld.

De uiteindelijke conjunctuurreeks is de verhouding van de originele, optioneel seizoensgecorrigeerde tijdreeks en de PAT-reeks. Deze reeks wordt geëffend door er een 12-perioden gecentreerd voortschrijdend gemiddelde van te nemen. De reeks wordt vervolgens gestandaardiseerd door van de reeks zijn geschatte gemiddelde af te trekken en vervolgens te delen door zijn geschatte standaarddeviatie.

De huizenprijsconjunctuur is de met het PAT-algoritme verkregen gestandaardiseerde conjunctuurreeks van de gemiddelde Nederlandse huizenprijs. In tabel B.1. staan de parameterwaarden van het algoritme, waarmee de afzonderlijke huizenprijsconjunctuurreeksen verkregen zijn. De correlatie tussen de huizenprijsconjunctuur, geconstrueerd over de gehele looptijd en de huizenprijsconjunctuur geconstrueerd over de verkorte periode vanaf 1979 is weergegeven onder het kopje 'correlatie lang-kort'. In figuur B.2.1. zijn de drie korte huizenprijsconjunctuurreeksen grafisch uitgezet.

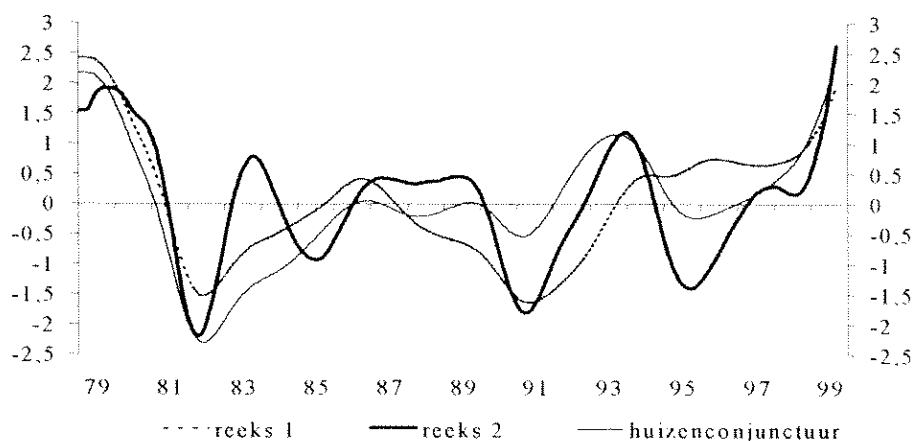
De conjunctuur van 37 tijdreeksen uit de initiële dataset, die een potentiële leidende indicator zijn en de initiële dataverzameling vormen, zijn onderzocht en de resultaten hiervan zijn samengevat in tabel B.3. Onder het kopje naam staan de variabelen die gepresenteerd zijn in tabel 4.1 in cryptische vorm omschreven. In de regel eronder staat als bronvermelding de mnemonic behorende bij Datastream, ofwel de formule waarmee de reeks berekend is uit andere reeksen, ofwel Rabo als de reeks uit interne Rabobankbronnen komt. Het item 'classificatie' deelt de beschouwde tijdreeks in als monetaire tijdreeks (m), economische tijdreeks (e) of woningmarkt tijdreeks (w).

De prestaties van de reeksen zijn samengevat in kruislingse correlaties. De kruislingse correlatie is de correlatie tussen de conjunctuurreeksen van twee variabelen, waarbij één reeks een aantal perioden verschoven is. Deze correlatie wordt dan berekend over de overlappende periode van beide reeksen. In de tabel is de maximale kruislingse correlatie weergegeven voor het symbool '/'. Achter dit symbool staat het aantal maanden verschuiving van de tijdreeks ten opzichte van de huizenprijsreeks wat deze maximale kruislingse correlatie veroorzaakt. De correlaties zijn berekend ten opzichte van huizenprijsconjunctuurreeks 1 en huizenprijsconjunctuurreeks 2, zoals die zijn weergegeven in tabel B.1.

Tabel B.1: De drie conjunctuurreksen van de huizenprijs

	Huizenprijs- conjunctuur 1	Huizenprijs- conjunctuur 2	Huizenprijs- conjunctuur
begindatum	mei-73	mei-73	mei-73
einddatum	sep-99	Sep-99	Sep-99
PAT-instellingen:			
MA-berekening	3/10	3/10	3/10
Opties beginpunt	4	4	4
opties eindpunt	5/0	5/0	5/0
opties algemene berekening:			
MA1	36	36	24
MA2	12	12	12
ZZ%	3,7	1,5	2
Seizoenscorrectie	j	j	j
gebruik beginpunt/eindpunt	j	j	n
MA1 ipv originele data	n	n	n
correlatie kort/lang	0,960553	0,798616	0,939741
Omslagpunten			
aantal PAT top/dal lange reeks	5	9	5
plaats 1e PAT top/dal	okt-75	okt-75	okt-75
plaats 2e PAT top/dal	jan-79	jan-79	dec-78
plaats 3e PAT top/dal	dec-83	nov-80	jan-84
plaats 4e PAT top/dal	okt-86	okt-83	aug-86
plaats 5e PAT top/dal	okt-94	okt-86	feb-92
plaats 6e PAT top/dal		apr-88	
plaats 7e PAT top/dal		jan-92	
plaats 8e PAT top/dal		mei-93	
plaats 9e PAT top/dal		dec-96	
aantal PAT top/dal korte reeks	4	8	4
plaats 1e PAT top/dal	aug-80	aug-80	aug-80
plaats 2e PAT top/dal	apr-85	feb-82	aug-85
plaats 3e PAT top/dal	okt-86	okt-83	aug-86
plaats 4e PAT top/dal	okt-94	okt-86	feb-92
plaats 5e PAT top/dal		apr-88	
plaats 6e PAT top/dal		feb-92	
plaats 7e PAT top/dal		apr-93	
plaats 8e PAT top/dal		dec-96	

Figuur B.2:Korte conjunctuurreksen van de huizenprijs



Tabel B.3: Conjunctuurreeksen

Naam	verkiptijd	Mktstmind	kraptelnd	lange rente	hyp. Rente	# transacties	korre rente	Kredprod Rabo	rentemarg e	Bouwver gunning	gdp	leading indicator	hpnndelt
bron	NVM	pt/pt(-1) (zie 1a)	v ^t /tr(-1) (zie 1b)	nibry/d	n/mortgg	Rabo	neuro3.	Rabo	hrent-krent	nicoconstrb	nlgdp...d	nloidding	hp(-1)-hp(-1)
Classificati	w	w	w	m	m/w	w	m	m/w	m/w	w	e	e	w
beginndatum	jan-85	feb-85	feb-85	mei-73	mei-73	mei-73	mei-73	mei-73	mei-73	jan-81	jan-77	mei-73	jun-73
einddatum	sep-99	sep-99	sep-99	nov-99	nov-99	sep-99	okt-99	okt-99	nov-99	jul-99	jun-99	jun-99	sep-99
PAT-instellingen:													
Opties beginpunt	4	4	4	2	2	2	2	4	2	4	2	4	2
opties eindpunt	4	4	4	5/36	2	4	2	5/0	5/0	5/24	5/0	5/0	5/0
opties algemene berekening:													
MA1	12	12	12	24	24	24	48	24	36	24	12	12	12
MA2	6	6	6	12	12	12	24	12	12	6	6	6	6
ZZ%	5	9	5	12	10	6	20	15	50	17	0.7	1	9
Seizoenscorrectie	j	j	j	j	j	j	j	j	j	j	j	j	n
gebruik beginpunt/eindpunt	j	j	j	n	j	j	j	j	j	j	j	j	j
Prestaties:													
plaats 1e PAT top/dal	feb-87	nov-86	feb-88	dec-81	jul-77	okt-75	jun-78	dec-77	feb-77	dec-82	okt-78	okt-74	mei-75
plaats laatste PAT top/dal	feb-89	jan-99	aug-98	sep-94	okt-90	nov-91	sep-91	feb-91	feb-88	apr-94	apr-91	feb-91	apr-97
aantal PAT top/dal	7	9	6	4	4	7	4	4	3	6	6	7	9
correlatie/verschuiving:													
wijzer 1, PAT-looptijd	0.94/31	0.25/7	0.87/27	0.86/3	0.44/18	0.46/6	0.31/21	0.87/5	0.5/25	0.8/31	0.26/0	0.52/4	0.67/17
wijzer 1, gehele looptijd	0.85/2	0.28/6	0.81/23	0.41/13	0.45/12	0.41/31	0.45/19	0.89/7	0.38/15	0.75/31	0.42/31	0.53/0	0.6/15
wijzer 2, PAT-looptijd	0.84/31	0.35/8	0.84/22	0.7/4	0.5/14	0.42/5	0.42/17	0.91/0	0.55/16	0.75/31	0.59/16	0.58/0	0.76/15
wijzer 2, gehele looptijd	0.83/3	0.28/7	0.8/22	0.37/10	0.42/11	0.43/31	0.44/21	0.85/7	0.44/17	0.57/22	0.41/0	0.55/0	0.76/15
aantal verschuiving:	2	6	21	6	9	9	6	6	12	18	30	6	12
selectie	n	n	j	j	n	n	j	j	n	j	n	n	n
verwijzing		1a	1b	2	3	4				5	6		7/8

1a) berekening van de marktstemmingsindex: $huizenprijs_t * transacties_t / huizenprijs_{t-1} * transacties_{t-1}$ bron: Kruijt & Janssen (1991)

1b) berekening van de krapteindex: $AANBOD_t / VERKOOP_t = (VRKPTD_t * TRANS_t) / VRRKPTD_t$

2) Vanaf april 1978 datastrom. De periode ervoor: lange rente interne Rabobankbron

3) conjunctuurmaat niet tegelijk gebruiken met lange rente (nibry/d)

4) variabele #Transacties samengesteld uit NVM-reeks en Kadaster-reeks

Naam	cp-iindex	cp-i-mutat	hyprentreel	m3	m3(jo)	aex	gmdprs	huur	#hyp	#hyrabo	hypbedra	hypbedra	Induventr
bron	nipc...f	nipcanni	nilmortgg-nipcanni	jitske	jitske	amsteoe	Gerben	wim	wim	wim	wim	wim	nleusicq
classificatie	m	m	m/w	m	m3(jo)	m	w	w	w	w	w	w	e
begindatum	mei-73	mei-73	mei-73	jan-74	jan-75	jan-83	mei-73	mei-73	jan-85	jan-85	jan-85	jan-85	mei-73
einddatum	sep-99	sep-99	sep-99	aug-99	aug-99	nov-99	dec-98	jan-99	sep-99	sep-99	sep-99	sep-99	okt-99
PAT-instellingen:													
Opties beginpunt	4	4	2	4	4	4	4	2	4	4	4	4	4
opties eindpunt	5/0	5/65	4	5/0	4	4	4	2	4	4	4	5/0	5/0
opties berekening:	algemene												
MA1	18	48	30	24	12	12	24	6	12	24	12	12	24
MA2	9	24	18	12	6	6	12	6	6	6	9	9	12
ZZ%	0,6	27	13	1,1	30	8	5	0,7	10	21	8	7	6
Seizoenscorrectie	n	n	n	j	n	n	j	j	j	j	j	j	j
gebruik beginpunt/eindpunt	j	j	j	j	j	j	j	j	j	j	j	j	j
Prestaties:													
plaats 1e PAT top/dal	okt-76	feb-76	apr-75	jul-75	apr-79	okt-83	sep-75	okt-75	dec-87	dec-88	dec-87	okt-85	jun-74
plaats laatste PAT top/dal	jun-91	jun-92	jun-93	dec-96	feb-97	feb-99	mrt-94	jun-91	dec-97	apr-96	jan-97	jan-96	nov-89
aantal PAT top/dal	7	4	5	11	5	10	6	4	6	3	4	7	4
correlatie/verschuiving:													
wijzer 1, PAT-looptijd	0,57/30	0,24/31	0,81/31	0,34/10	0,55/22	0,39/1	0,82/2	0,16/31	0,97/0	0,89/0	0,93/31	0,58/0	0,2/0
wijzer 1, gehele looptijd	0,58/19	0,31/31	0,49/22	0,34/9	0,16/0	0,39/1	0,81/3	0,27/31	0,93/0	0,9/0	0,89/0	0,55/0	0,52/31
wijzer 2, PAT-looptijd	0,45/25	0,18/0	0,89/31	0,37/8	0,22/19	0,31/20	0,85/1	0,13/0	0,8/0	0,7/0	0,88/0	0,48/19	0,29/0
wijzer 2, gehele looptijd	0,5/19	0,24/0	0,61/22	0,37/8	0,1/11	0,31/19	0,84/2	0,19/31	0,8/0	0,66/0	0,73/1	0,44/18	0,58/30
aantal verschuiving:	20	-4(!)	22	0	12	18	2	0	4	0	3	3	0
selectie verwijzing	n	n	n	n	n	n	n	n	n	n	n	n	n

5) vanaf 81 en 30 maanden verschoven: erg korte reeks

6) oorspronkelijk kwartaaldata. Transformatie als volgt: middelste maand van het kwartaal krijgt de waarde en de waarden voor de tussenvolgende 2 maanden wordt linear geïnterpoleerd.

7) geanalyseerde reeks is oorspronkelijke tijdsreeks +10. Hiermee wordt de de nulpuntproblematiek vermeden. Het PAT-algoritme werkt met twee reeksen van verhoudingen, nl Input/MA(1) en Input/PAT. Als een van de variabelen in de noemer (PAT of MA(1)) heel dicht bij nul licht, dan wordt er door bijna nul gedeeld en komen er hele extreme getallen uit. Het verschuiven van een reeks heeft geen invloed op het conjunctuurpatroon van die reeks. Bovendien zijn alle verschoven reeksen zogenaamde kwalitatieve reeksen, dat wil zeggen reeksen die het vertrouwen van mensen in de economie weerspiegelen.

8) Hoge correlaties zijn misleidend. Ze worden louter veroorzaakt door de 'bubble'.



Naam	zakenventr	persobespa	werkloosh	leadind	werkgelege nheid	Orders bouw	consument vertrouwen	economien timent	bouwverg	bouwvergi(jf	Kostconstr uctie	yield
bron	nlocbsfp	nlnpersva	wim	nloodng	nlempalf	nleusbodq	nleuscicq	nleuseisg	wim	wim	nlocconsf	lange rente-korte rente
classificatie	e	e	e	e	e	w	e	e	w	w	w	m
begindatum	mei-73	jan-74	mei-73	mei-73	jan-84	mei-73	mei-73	mei-73	jan-88	dec-88	mei-73	mei-73
einddatum	jul-99	sep-99	okt-99	jun-99	mt-99	okt-99	sep-99	sep-99	jun-99	jun-99	jun-99	nov-99
PAT-instellingen:												
Opties beginpunt	5	2	4	2	2	2	4	4	4	2	4	4
opties eindpunt	5/50	5/50	4	2	2	2	4	4	4	4	4	5/0
opties algemene berekening:												
MA1	24	24	24	9	12	24	12	24	24	12	12	24
MA2	12	12	9	6	12	28	9	18	12	12	6	12
ZZ%	5	20	8	1	1	1	4,4	1	10	2	1	15
Seizoenscorrectie	j	j	n	j	n	n	n	j	j	n	j	j
gebruik beginpunt/eindpunt	n	n	j	j	j	j	j	j	j	j	j	j
Prestaties:												
plaats 1e PAT top/dal	okt-74	okt-78	sep-75	apr-75	jun-84	jun-76	jun-76	jun-76	feb-90	dec-91	dec-77	aug-74
plaats laatste PAT top/dal	jul-85	jan-93	apr-96	apr-87	apr-92	aug-87	apr-96	sep-96	jun-97	mt-97	jul-93	apr-88
aantal PAT top/dal	4	5	6	2	4	2	5	6	5	4	8	4
correlatie/verschuiving:												
wijzer 1, PAT-looptijd	0,21/22	0,29/0	0,32/0	0,71/8	0,55/31	0,38/1	0,36/5	0,23/4	0,94/0	0,99/31	0,6/0	0,53/31
wijzer 1, gehele looptijd	0,2/4	0,32/0	0,26/0	0,52/2	0,51/14	0,42/0	0,39/4	0,28/2	0,85/31	0,99/29	0,68/31	0,4/30
wijzer 2, PAT-looptijd	0,21/0	0,25/0	0,29/0	0,64/5	0,34/10	0,45/0	0,4/4	0,3/1	0,76/0	0,93/13	0,49/29	0,54/28
wijzer 2, gehele looptijd	0,3/0	0,26/0	0,19/0	0,42/0	0,56/9	0,48/0	0,42/5	0,36/0	0,73/0	0,86/19	0,68/31	0,45/29
aantal verschuiving:	0	0	0	3	0	0	0	0	0	12	18	24
selectie	n	n	n	n	n	n	n	n	n	n	n	n
verwijzing	13		14		6			13	11	15/16	6	7

9) overeenkomst van reeks met huizenprijsreeks frappant. Ook hier transformatie naar maanddata

10) lineaire interpolatie van jaarcijfers naar maandcijfers

11) geen voorloop, maar goede wijzer!

12) geen voorloop, beperkte data, maar goede wijzer!

13) transformatie: oorspronkelijke reeks +100: nulpuntenproblematiek (zie 8)

14) transformatie: oorspronkelijke reeks +1000: nulpuntenproblematiek (zie 8)

15) Let op: hele goede reeks, maar te weinig data

16) niet tegelijk gebruiken met Bouwverg (nlocconstrb)

Onder PAT-looptijd wordt verstaan de periode van een tijdreeks tussen het eerste PAT-punt en het laatste PAT-punt. Vanwege de onzekerheidsproblematiek van het PAT-algoritme ten aanzien van de trend aan de uiteinden van een tijdreeks zijn ook de kruislingse correlaties berekend over de periode van de overlap van de PAT-looptijden van beide reeksen. Het idee erachter is dat over deze perioden een juiste trend en daarmee een juiste conjunctuurmaat is berekend. Deze prestaties zijn weergegeven in de tabel onder het item 'PAT-looptijd'.

De optimale verschuiving wordt bepaald aan de hand van de verschuiving behorende bij een maximale kruislingse -correlatie en de verschuiving behorende bij een optimale overeenkomst van de omslagpunten van de conjunctuurreeks en de huizenprijsconjunctuurreeks.

Onder het item 'verwijzing' staat een verwijzing naar opmerkingen die onder aan de tabel zijn geplaatst. Het uiteindelijke aantal perioden verschuiving wordt bepaald aan de hand van een afweging tussen het aantal perioden verschuiving behorende bij de maximale kruislingse correlatie en de visuele inspectie over de mate van synchroniteit van de omslagpunten van beide tijdreeksen.

B.1.2: Conjunctuur voorspellen

Onder het kopje 'selectie' in tabel B.3 staat weergegeven of een reeks al dan niet is geselecteerd om opgenomen te worden in de conjunctuurvoorspeller. De selectie geschiedt aan die hand van de criteria, die in subparagraaf 4.3.1 geformuleerd staan. De geselecteerde variabelen zijn: *Lange rente*, *Korte rente* en *Kredietproductie Rabobank*.

De conjunctuurvoorspeller is als toepassing van (4.3.1) een lineaire combinatie van de geselecteerde reeksen en dient de huizenprijsconjunctuur te repliceren en te voorspellen. De lineaire wegingscoëfficiënten worden bepaald met lineaire regressie. De regressie op de leidende indicatoren is verricht voor de drie geconstrueerde reeksen van de huizenprijsconjunctuur, zoals weergegeven in tabel B.1. De wegingscoëfficiënten van de lineaire regressie van de drie korte reeksen van de huizenprijsconjunctuur op de drie geselecteerde conjunctuurreeksen zijn weergegeven in tabel B.4. Op basis van de prestatie maatstaven wordt de selectie van de conjunctuurreeks in de derde kolom verantwoord en zijn tevens de *t*-statistische gegevens van de coëfficiënten behorende bij deze conjunctuurreeks opgenomen.

Tabel B.4: Lineaire regressie van de huizenprijsconjunctuur op de geselecteerde reeksen

Coëfficiënten	Reeks 1	Reeks 2	huizenconjunctuur	t-waarden (huizenconjunctuur)
Snijpunt	0,200769339	0,13308671	0,120570945	5,166917577
LRENT	-0,001190077	0,474156515	0,780688853	19,04997464
KRENT	0,305261646	-0,274296726	-0,721698727	-18,87838473
LENAB	0,837999461	0,601610117	0,666561454	20,83092622
Prestatiemaatstaven				
R ²	0,70	0,56	0,88	
Correlatie	0,83	0,75	0,89	

B.2: Documentatie procedure en resultaten stochastische werkwijze

In deze paragraaf wordt alle documentatie met betrekking tot de spectraalmethode voor de constructie van de huizenprijs weergegeven. In subparagraaf B.2.1 staan de resultaten van de toepassing van de Dickey-Fuller testprocedure ter wegfiltering van de stochastische trend. In subparagraaf B.2.2 staat de documentatie van de selectie van de leidende cyclische indicatoren.

B.2.1: Conjunctuur bepalen met stochastische trend

De stochastische trend wordt operationeel verwijderd uit een tijdreeks door de differentiereeks van deze tijdreeks te bepalen. De oorspronkelijke tijdreeks en de geconstrueerde differentiereeks worden met de in paragraaf 2.4 beschreven Dickey-Fuller procedures getoetst op stationariteit. De testprocedure is geautomatiseerd in het softwarepakket van Bierens (2000) en de resultaten van de

toepassing van deze procedure op tijdreeksen van de initiële dataverzameling staan getabelleerd in tabel B.5. Met (2.4.3) wordt de nulhypothese van een eenheidswortel getoetst tegen de alternatieve hypothese van stationariteit. Indien geldt $\gamma=0$ en $a_0=0$, is het geschikter te testen met (2.4.2). Het aantal vertragingen p van (2.4.2) en (2.4.3) staan in de laatste kolom gedocumenteerd. Indien voor beide reeksen de nulhypothese van stationariteit verworpen is, wordt er een hogere orde differentiereeks bepaald. In de praktijk van economische tijdreeksen komen tweede orde differentiereeksen niet vaak voor en hogere orden differentiereeksen zijn zelfs zeer zeldzaam. In de derde kolom staat de conclusie vermeldt: n betekent geen eenheidswortel en dus stationariteit en j het tegenovergestelde.

Als aanvulling op de initiële dataset is voor een aantal variabelen de seizoensgecorrigeerde tijdreeks bepaald op de volgende wijze: $\Delta_{12}y_t = y_t - y_{t-12}$. Het gaat hier om de volgende variabelen: #transacties, Kredietproductie Rabobank, #hypotheek, #hypotheek Rabobank, Bedrag hypotheek, Bedrag hypotheek Rabobank. Deze seizoensgecorrigeerde reeksen worden meegenomen als potentiële leidende indicatoren onder de restrictie dat ze niet tegelijk met de bijbehorende seizoensongecorrigeerde tijdreeks geselecteerd mogen worden.

Tabel B.5: Toets op stationariteit data

Volledige naam	Afkorting	Vergelijking 2.4.3.				Vergelijking 2.4.2.			
		γ	$a_0=0 \gamma=0$		OLS	t	p	$a_0=0 \gamma=0$	
			OLS	t				OLS	T
Huizenprijsindex	Hpi	y	0.0015	0.9513	-0.0562	-0.4488	0.0009	1.4977	11
	$\Delta(\text{hpi})$	y	-0.0508	-0.8422	0.0551	1.2385	-0.0121	-0.2348	10
	$\Delta^2(\text{hpi})$	n	-3.9674	-10.7428					9
lange rente (10-jaars Nederlandse staatsobligaties)	Lrent	y	-0.015	-1.496	0.1092	1.3665	-0.0016	-0.7523	1
	$\Delta(\text{lrent})$	n	-0.8572	-1.527					0
korte rente (3-maands AIBOR/EURIBOR)	Krent	y	-0.0394	-1.8497	0.2358	1.5711	-0.0075	-1.1671	19
	$\Delta(\text{krent})$	n	-1.4358	-5.0433					18
# transacties woningmarkt	Trans	y	0.0147	1.5012	0.2664	0.3211	0.0173	3.1973	20
	$\Delta(\text{trans})$	n	-1.4787	-2.9548					19
	$\Delta_{12}(\text{trans})$	n	-0.4798	-67.386					11
Kredietproductie Rabobank	kredprod	y	0.0919	3.9448	-73.3071	-2.6507	0.0362	3.5703	18
	$\Delta(\text{kredprod})$	n	-2.359	-3.4613					17
	$\Delta_{12}(\text{kredprod})$	y	-0.0991	-1.612	17744,704 5	1,3909	-0.0559	-1.052	16
	$\Delta\Delta_{12}(\text{kredprod})$	n	-3.674	-7.661					
Hypotheekrente	hyprent	y	-0.0153	-2.0255	0.1263	1.8759	-0.0014	-0.9682	10
	$\Delta(\text{hyprent})$	n	-0.562	-5.3353					9
Verkooptijd	vkptd	y	-0.0402	-2.673	1.630	2.168	-0.0093	-1.910	15
	$\Delta(\text{vkptd})$	n	-1.734	-4.182					
Marktstemmingsindex woningmarkt	msi	y	0.0689	5.1158	-307.7123	-1.6217	0.0495	7.9888	10
	$\Delta(\text{msi})$	n	-4.3332	-13.5305					9
krapte-indicator woningmarkt	krind	y	0.0524	1.694	-2.8076	-0.8317	0.0277	3.1238	14
	$\Delta(\text{krind})$	n	-7.4962	-4.3777					18
Rentemarge (=hyre-lrent)	rentmrg	y	-0.3347	-7.9279	0.3009	7.1948	-0.0582	-3.0897	0
	$\Delta(\text{rentmrg})$	n	-2.008	-11.545					4
Bouwvergunningen (bron Rabo)	bvgr	y	-0.0314	-0.8881	38.3429	1.0084	0.0033	0.4077	2
	$\Delta(\text{bvgr})$	n	-4.067	-3.791					18
gdp	gdp	y	0.0014	1.613	-51.9229	-1.0018	0.0006	2.6899	13
	$\Delta(\text{gdp})$	n	-0.3437	-3.4753					12
leading indicator	lind	y	0.0011	0.9535	-0.0012	0.0126	0.0011	4.3335	20
	$\Delta(\text{lind})$	n	-0.5606	-5.5002					19
cpi-index	cpiind	y	-0.0007	-0.9148	0.1048	1.3073	0.0003	1.1542	12
	$\Delta(\text{cpiind})$	y	-0.1859	-1.8505	0.0358	1.4862	-0.0549	-1.1364	11

	$\Delta^2(\text{cpiind})$	n	-8.8978	-12.8823					10
cpi-joj (groeipercentage)	cpijoj	y	-0.0117	-1.9556	0.0288	1.0388	-0.0067	-1.8892	12
	$\Delta(\text{cpijoj})$	n	-0.7431	-5.4921					11
hypotheekrente reeel	hyrere	y	-0.0155	-1.5096	0.0809	1.4734	-0.0015	-0.3804	12
	$\Delta(\text{hyrere})$	n	-1.2773	-6.4177					11
geidhoeveelheid	m3	y	0.0043	2.62	0.1113	0.9378	0.005	3.3054	19
	$\Delta(\text{m3})$	y	-0.2371	-0.9894	0.1457	1.2688	0.0437	0.4752	23
	$\Delta^2(\text{m3})$	n	-13.1033	-14.6815					11
geidhoeveelheid (groeipercentage)	m3joj	y	-0.0496	-2.6032	0.3731	2.3229	-0.0081	-1.2163	12
	$\Delta(\text{m3joj})$	n	-1.2908	-6.5725					11
AEX-index	aex	y	0.0201	1.8061	-0.5672	-0.3087	0.0175	2.4307	14
	$\Delta(\text{aex})$	n	-0.6844	-2.4904	1.9606	1.6358	-0.4313	-1.890	13
grondprijs	grdprs	y	-0.0198	-2.446	765.401	2.613	0.0007	0.3504	12
	$\Delta(\text{grdprs})$	y	-0.5233	-2.839	70.626	0.975	-0.4612	-2.666	11
	$\Delta^2(\text{grdprs})$	n	-9.420	-15.415					10
huur	huur	y	0	0.6782	0.0332	1.5387	0.0001	1.2464	1
	$\Delta(\text{huur})$	y	-0.0199	-1.7863	0.038	1.8639	-0.0001	-0.0328	0
	$\Delta^2(\text{huur})$	n	-1.001	-17.42					0
aantal afgesloten hypotheek	#hyp	y	0.0302	0.9991	-292.3203	-0.3528	0.0202	1.892	12
	$\Delta(\text{\#hyp})$	n	-2.044	-3.649					11
	$\Delta_{12}(\text{\#hyp})$	y	-0.2287	-2.560	600.115	1.809	-0.1347	-1.838	15
	$\Delta\Delta_{12}(\text{\#hyp})$	n	-2.479	-6.381					12
aantal afgesloten hypotheek bij Rabobank	#hyprab	y	0.0333	0.815	-91.0812	-0.3283	0.0204	1.7761	12
	$\Delta(\text{\#hyprab})$	n	-3.081	-4.505					11
	$\Delta_{12}(\text{\#hyprab})$	y	-0.2747	-25.713	137.665	1.647	-0.1807	-1.941	13
	$\Delta\Delta_{12}(\text{\#hyprab})$	n	-2.761	-6.085					12
totale bedrag uitstaande hypotheek	hypbed	y	0.1011	3.3482	-197.2692	-1.821	0.0557	3.245	17
	$\Delta(\text{hypbed})$	n	-3.754	-6.643					10
	$\Delta_{12}(\text{hypbed})$	y	-0.0294	-0.3991	70.619	0.989	0.0177	0.3155	13
	$\Delta\Delta_{12}(\text{hypbed})$	n	-2.265	-3.865					13
totale bedrag uitstaande hypotheek bij Rabobank	hypbedrabo	y	0.1343	3.7696	-60.6855	-2.2153	0.068	3.4669	18
	$\Delta(\text{hypbedrabo})$	n	-1.715	-2.822					11
	$\Delta_{12}(\text{hypbedrabo})$	y	-0.1148	-1.809	24.779	1.602	-0.0539	-1.055	6
	$\Delta\Delta_{12}(\text{hypbedrabo})$	n	-2.455	-4.760					11
Industrievertrouwen	induver	y	-0.039	-2.511	3.659	2.507	-0.0002	-0.1557	15
	$\Delta(\text{induver})$	y	-0.7374	-6.575					3
Zakenvertrouwen	bussconf	n							13
persoonlijke besparingen	persosav	y	-0.3032	-2.4101	354.8401	2.479	0.0065	0.4325	12
	$\Delta(\text{persosav})$	n	-4.6563	-6.5583					11
werkloosheid	ww	y	-0.0062	-1.5758	0.0386	1.4474	-0.0008	-0.6451	20
	$\Delta(\text{ww})$	n	-0.3517	-3.6571					19
leading indicator	leadind	y	0.0011	0.9535	-0.0012	-0.0126	0.0011	4.3335	20
dleading indicator	$\Delta(\text{leadind})$	n	-0.5606	-5.5002					19
werkgelegenheid	empl	y	0.0014	0.5923	-4.1937	-0.3316	0.0006	2.9035	7
	$\Delta(\text{empl})$	n	-0.3343	-3.1302					9
ordervoorraad constructie	ordvoconst	y	-0.0061	-0.6568	0.0478	0.1693	-0.0074	-1.2879	1
	$\Delta(\text{ordvoconst})$	n	-1.256	-22.8438					0
consumentenvertrouwen	consconf	y	-0.0107	-1.1046	1.0655	1.1578	0.0004	0.2704	0
	$\Delta(\text{consconf})$	n	-0.9045	-16.0237					0
economisch sentiment	ecosent	y	-0.016	-1.693	1.615	1.710	0.0002	0.7077	21
	$\Delta(\text{ecosent})$	n	-0.6548	-3.796					21
bouwvergunning	bvg	y	-0.195	-1.3983	1102.4078	1.3755	-0.0042	-0.29	11
	$\Delta(\text{bvg})$	n	-6.0764	-9.7379					10
kosten constructie	kostconst	y	-0.0004	-0.559	0.069	1.1584	0.0004	1.929	10

	$\Delta(\text{kostconst})$	n	-0.1645	-2.786	0.0371	2.1832	-0.0649	-1.720	9
yield (bron: datastream)	yield	y	-0.1038	-2.878	1.159	2.891	-0.0005	-0.1191	17
	$\Delta(\text{yield})$	n	-1.722	-5.498					16
bouwvergunning (jaartotalen)	bvgjt	y	-0.0283	-18.806	1925,9831	1.871	-0.0003	-0.1845	15
	$\Delta(\text{bvgjt})$	n	-0.651	-4.119					11

B.2.2: Selectie van de leidende cyclische indicatoren

De selectie van de leidende cyclische indicatoren geschiedt aan de hand van de gemiddelde waarden van de verschuiving s en de samenhang K^2 voor de tijdreeks opbouwende golven met lengtes groter dan 1 jaar en kleiner dan zeven jaar. De samenhang K^2 is gedefinieerd met (3.3.12), de verschuiving s is gedefinieerd met (3.3.14) en het frequentieinterval behorende bij de conjunctuurgolven met $\omega_p = 2\pi \cdot p/n$, waarbij $l = n/p$ de golfte in aantal maanden per cyclus. Door de keuze van de golflengtes voor de conjunctuur worden de grenzen berekend met de waarden $l=12$ en $l=84$ in maanden. De maatstaven K^2 en s worden gemiddeld over het frequentieinterval behorende bij deze golflengten.

In subparagraaf 4.3.2.1 is de praktische complicatie beschreven van toepassing van de spectraalanalyse betreffende de keuze voor de bandbreedteparameter m van de weegfunctie. De definitieve keuze voor deze parameter is voor de selectie van de leidende cyclische indicatoren uit de initiële dataset genuanceerd door de spectraalmaatstaven voor meerdere waarde van deze bandbreedteparameter te berekenen. De resultaten staan weergegeven in onderstaande tabel B.6. In de eerste kolom staat de cryptische naam van de variabele en in de tweede kolom het aantal beschikbare data voor deze variabele. De derde, vierde en vijfde kolom geven respectievelijk de berekende waarden per variabele weer van K^2 , s en de gebruikte parameter m . Deze trits kolomen herhalen zich zeven keer, waarbij er telkens lagere waarden voor de bandbreedteparameter m bepaald zijn. In de laatste kolom staan het aantal maal dat de spectraalmaatstaf s een lagere waarde dan -6 getoond heeft. Indien dat voor de acht berekeningen vaker dan zes maal het geval is, wordt deze variabele geselecteerd als leidende indicator voor de cyclische beweging. De enige uitzondering is de reeks *Industrievertrouwen* (in de tabel cryptisch omschreven als *inducon*: industrial confidence), omdat juiste voorlooptijd van minimaal zes perioden vergezeld wordt met een samenhang van meestal kleiner dan 0.1.

De als leidende indicatoren geselecteerde reeksen zijn: *Korte rente*, *Hypotheekrente*, *Verkooptijd* en *CPI-index*.

Opvallend in tabel B.6 is dat de samenhangsmaatstaf K^2 afneemt met de afname van de bandbreedteparameter m . Volgens Priestley (1981) blz 708 is de K^2 berekend over de periodogrammen van twee reeksen zelfs per definitie gelijk aan 1. Dat komt omdat deze maatstaf voor de betreffende frequentie slechts op één enkele waarneming gebaseerd is. Door toepassing van een kernel wordt impliciet verondersteld dat de coherentie dezelfde waarde heeft voor een interval van frequenties, waardoor deze berekend kan worden over meerdere observaties. Deze redenering verklaart de afhankelijkheid van K^2 en M .



Appendix C: Matlab-applicatie

```
function [vspl,vsplz,nummer,mwin,begineindselect,ecf]=indicato(z,critK2,crits);
```

```
%Dit is het hoofdprogramma van de spectraalanalyse. De schuingedrukte regels zijn commando's die naar andere M-files  
% verwijzen.
```

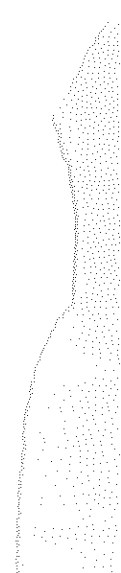
```
%output: vspl,som,nummer,mwin,begineind,ecf  
%vspl=voorspeller huizenprijsreeks  
%som=voorspeller 2e orde differentie van huizenprijsreeks  
%nummer=label van variabele dat is geselecteerd  
%mwin=de ingestelde parameter m die is gebruikt voor de bandbreedte van de raamschatter  
%begineind=positie eerste en positie laatste waarneming. Hieruit is het totaal aantal waarnemingen te berekenen.  
%ecf=empirische correctiefactor. Dit is het getal waarmee de voorspeller 'som' moet worden vermenigvuldigd om aan de  
% stationariteitsvoorwaarde te voldoen.  
%  
%input: z,critK2,crits  
%z=matrix met in de eerste kolom de referentietijdreeks (=huizenprijs) en in de overige kolommen de potentiële leidende  
% indicatoren. Indien een tijdreeks later begint of eerder stopt kan de default-waarde 123.45 ingevuld worden.
```

```
Ytmineen=173.2590;  
Yt=175.0379;  
[n,m]=size(z);  
MWIN2=[];  
[begineind]=begeind(z);  
for i=1:m  
    [mwin,zrw]=covmaxs(z(:,i),begineind(1,i),begineind(2,i));  
    MWIN2=[MWIN2 round(mwin)];  
end;  
[K2gem,sgemvooraftek]=hpi(z,MWIN2,begineind);  
[sgem]=safirek(sgemvooraftek,begineind);  
[nummer]=selectie(K2gem,sgem,critK2,crits);  
if length(nummer)>1  
    [zselect,mwinselect,smin,smax]=schuif(z,MWIN2,nummer,sgem);  
    [begineindselect]=begeind(zselect);  
    [coefmatruw]=mulspan(zselect,mwinselect,smin,smax,begineindselect);  
    coefmatr=(real(coefmatruw)); %afroundingsfout  
    [vspl,vsplz,ecf]=voorspel(zselect,coefmatr,smin,smax,ytmineen,yt,begineindselect);  
else 'geen selectie'  
end;
```

```
function [begineind]=begeind(z);
```

```
%De inputmatrix z bestaat behalve uit waarnemingen voor een niet volledige tijdreeks ook aan het begin en het eind van een  
%tijdreeks uit default-waarde 123.45 Deze functie genereert een matrix met de positie (=datum) van de eerste waarneming  
%en van de laatste waarneming
```

```
[n,m]=size(z);  
for teller2=1:m  
    beginwaardeteller=0;  
    beginwaarde=123.45;  
    eindwaarde=123.45;  
    for teller=1:n  
        if z(teller,teller2)==123.45  
            beginwaardeteller=beginwaardeteller+1;  
        else  
            beginwaarde=beginwaardeteller+1;  
        end;  
        if ((beginwaarde==123.45) & (z(teller,teller2)==123.45) & (eindwaarde==123.45))  
            eindwaarde=teller-1;  
        end;  
    end;  
end;
```



```

end;
end;
if (eindwaarde==123.45) eindwaarde=teller;
end;
begin eind(1,teller2)=beginwaarde;
begin eind(2,teller2)=eindwaarde;
end;

```

```

function [teller4,z1]=covmaxs(z,beginwaarde,eindwaarde);

```

% bepaald aan de hand van de covariantiefunctie de plaats waarachter de
 % autocovariantie ongeveer nul is voor de differentie-stationaire reeksen

```

z1=z(beginwaarde+10:eindwaarde);
covvec=covf((z1-mean(z1)),(eindwaarde-beginwaarde));
covvec3=abs(covvec)-mean(covvec)-std(covvec);
for i=3:length(covvec3)
    if covvec3(i)>0
        teller2=round(i/4);
    end;
end;
teller3=max((round((eindwaarde-beginwaarde)/20),teller2);
teller4=min(teller3,round((eindwaarde-beginwaarde)*1/2));

```

```

function [K2gem,sgem]=hpi(z,mwin,begin eind);

```

%deze functie bepaalt de samenhang (K2) en de verschuiving (s) van twee spectraalfuncties.
 %door gemiddelden te berekenen van deze maatstaven over middellange golven, dat wil zeggen golven met een lengte van
 %groter dan 1 jaar en kleiner dan 7 jaar. Deze gemiddelden (K2gem en sgem) zijn de prestatiewaarden die bepalen of deze
 %variabele geselecteerd gaat worden als leidende indicator.

```

[n,m]=size(z);
K2=[];
s=[];
K2gem=zeros(m,1);
sgem=zeros(m,1);
Fs=2*pi;
nfft=length(z);
for j=2:m,
    beginwaarde=max(begin eind(1,1),begin eind(1,j));
    eindwaarde=min(begin eind(2,1),begin eind(2,j));
    lengte=eindwaarde-beginwaarde+1;
    fmax=(2*pi)/lengte;
    if lengte<=84
        fmin=(2*pi)/(lengte-1);
    else
        fmin=(2*pi)/84;
    end;
    [Pxy,f]=csd(z(beginwaarde:eindwaarde,1)-mean(z(beginwaarde:eindwaarde,1)),z(beginwaarde:eindwaarde,j)-
    mean(z(beginwaarde:eindwaarde,j),lengte,Fs,mwin(j),'none'));
    [K2,f]=cohere(z(beginwaarde:eindwaarde,1)-mean(z(beginwaarde:eindwaarde,1)),z(beginwaarde:eindwaarde,j)-
    mean(z(beginwaarde:eindwaarde,j),lengte,Fs,mwin(j),'none'));
    f(1)=f(2);
    s=-(unwrap(angle(Pxy)))/f;
    beginwaardegem=0;
    eindwaardegem=0;
    for i=1:length(f)
        if f(i)<fmin
            beginwaardegem=beginwaardegem+1;
        end;
        if f(i)<fmax
            eindwaardegem=eindwaardegem+1;
        end;
    end;
end;

```



```

end;
end;
beginwaardegem=beginwaardegem+1;
K2gem(j)=(sum(K2(beginwaardegem:eindwaardegem))/(eindwaardegem-beginwaardegem+1));
sgem(j)=round(sum(s(beginwaardegem:eindwaardegem))/(eindwaardegem-beginwaardegem+1));
end;

```

function [sgem]=saftrek(sgemvooraftrek,begineind);

%Deze functie corrigeert de voorlooptijd van een variabele voor beschikbaarheid
 %van deze variabele. Indien een variabele s perioden voorloopt, maar data 2 perioden
 %later beschikbaar is, dan is zijn effectieve voorlooperperiode (s-2) perioden. Indien een variabele qua beschikbaarheid
 %2 perioden voorloopt is de effectieve voorlooptijd (s+2) perioden.

```

for i=2:length(sgemvooraftrek)
    verschil=begineind(2,i)-begineind(2,i);
    sgem(i)=sgemvooraftrek(i)+verschil;
end;
sgem=sgem';

```

function [nummer]=selectie(K2,s,critK2,crits);

%deze functie bepaalt aan de hand van criteriumwaarden voor K2 en s welke
 %tijdreksen geselecteerd worden als leidende indicator

```

nummer=[1];
n=length(K2);
m=length(s);
if n~=m
    error;
end;
for i=1:n
    if K2(i)>=critK2 & s(i)<=-crits
        nummer=[nummer i];
    end;
end;

```

function [zselect2,mwinselect,smin,smax]=schuif(z,MWIN2,nummer,sgem);

%deze functie verschuift de geselecteerde variabelen met hun gemiddelde
 %voorlooptijd en maakt een nieuwe matrix met daarin de verschoven tijdreksen.

```

[n,m]=size(z);
for i=1:length(nummer)
    zselect1(:,i)=z(:,nummer(i));
    if i==1
        sselect(i)=0;
    else
        sselect(i)=sgem(nummer(i));
    end;
    mwinselect(i)=MWIN2(i);
end;
smax=max(sselect(2:length(sselect))); %merk op: sselect negatieve waarden
smin=min(sselect);
zselect2=123.45*ones([n-smin length(nummer)]);
zselect1=[zselect1;123.45*ones([abs(smin) length(nummer)])];
for i=1:(n-smin)
    for j=1:length(nummer)
        if(i+sselect(j))>0
            zselect2(i,j)=zselect1(i+sselect(j),j);
        end;
    end;
end;

```



```
end;
```

```
function [coefmatr]=mulspan(z,mwin,smin,smax,begineind);
```

```
%Deze functie is bedoeld om het lineaire leidende indicatorenmodel te schatten met spectraalanalyse.  
%De functie bepaalt eerst de multivariate spectraalfunctie en vervolgens de transferfunctie, die de spectraalrelatie tussen  
%de twee tijdreeksvariabelen beschrijft. Uiteindelijk wordt met de transferfunctie de coëfficiëntenmatrix berekend.  
%De coëfficiëntenmatrix gaat overigens maximaal 7 jaar terug in de tijd.
```

```
[n,m]=size(z);  
Fs=2*pi;  
minlengte=1000;  
for i=1:m  
    lengte=begineind(2,i)-begineind(1,i);  
    if lengte<minlengte  
        minlengte=lengte;  
    end;  
end;  
nfft=length(z);  
msmatryy=zeros(size3d([m-1 m-1 (floor(nfft/2)+1)]));  
msmatrxy=zeros(size3d([1 m-1 (floor(nfft/2)+1)]));  
for i=1:m,  
    for j=2:m,  
        beginwaarde=max(begineind(1,i),begineind(1,j));  
        eindwaarde=min(begineind(2,i),begineind(2,j));  
        lengte=eindwaarde-beginwaarde+1;  
        [Pxyij,f]=csd(z(eindwaarde-minlengte:eindwaarde,i)-mean(z(eindwaarde-minlengte:eindwaarde,i)),z(eindwaarde-  
minlengte:eindwaarde,j)-mean(z(eindwaarde-minlengte:eindwaarde,j)),nfft,Fs,(mwin(i)+mwin(j)/2),'none');for  
k=1:floor(nfft/2)+1  
        if i==1  
            msmatrxy(elem3d([1 m-1 floor(nfft/2)+1],i,j-1,k))=Pxyij(k);  
        else  
            msmatryy(elem3d([m-1 m-1 floor(nfft/2)+1],i-1,j-1,k))=Pxyij(k);  
        end;  
    end;  
end;  
end;
```

```
%berekenen van de multivariate transferfunctie
```

```
mtransmatr=[];  
for k=1:floor(nfft/2)+1  
    fxy=msmatrxy(:,k)';  
    fyy=[];  
    for i=1:(m-1)  
        fyy=[fyy (msmatryy((m-1)*(i-1)+1:i*(m-1),k))];  
    end;  
    mtransmatr=[mtransmatr (fxy*inv(fyy))];  
end;
```

```
%berekenen van coëfficiënten uit transferfunctie
```

```
coefmatr=[];  
imaggetal=sqrt(-1);  
for k=1:(m-1)  
    trans=mtransmatr(k,:);  
    for L=(smin-smax):minlengte/2  
        coefmatr(k,L-(smin-smax)+1)=(1/pi)*trapz(f,trans.*exp(imaggetal*f*L));  
    end;  
end;
```

```
function [yt,som]=voorspel(zselect,coefmatr,smin,smax,ytmineen,ytminnul,sennietz,begineind);
```

```
%Deze functie is simpelweg een algoritme dat de geschatte coëfficiënten van het lineaire leidende indicatorenmodel
```

%koppelt aan de bijbehorende realisaties van de leidende indicatoren en daarmee aanvankelijk de voorspelling van de %stochastische huizenprijsconjunctuur genereert en door de terugkoppeling een voorspelling voor de huizenprijs zelf.

```

som=zeros(size(1:abs(smax)));
hulpsom=zeros(size(1:abs(smax)));
for i=1:length(beginneind);
    gem(i)=mean(zselect(beginneind(1,i):beginneind(2,i),i));
end;
[m,n]=size(coefmatr);
o=beginneind(2,1);
for s=1:abs(smax)
    for i=1:m
        for j=1:n
            if zselect(o-smin-j+1,i+1)~=123.45
                hulpsom(s)=(hulpsom(s)+coefmatr(i,j)*zselect(o-smin+(smax+s)-j+1,i+1));
            end;
        end;
    end;
end;
som=((hulpsom'+gem(1)-(gem(2:length(gem)))*(sum(coefmatr'))));
yt=zeros(size(1:abs(smax)));
yt(1)=ytmineen;
yt(2)=ytminnul;
for i=1:abs(smax)
    yt(i+2)=som(i)+2*yt(i+1)-yt(i);
end;

```



Prijs je rijk

Zal de hausse op de huizenmarkt aanhouden, of wordt het eigen huis de kat in de zak van de jaren nul? Zover kan men niet vooruit kijken, maar over de nabije toekomst kan wel degelijk iets worden gezegd.

W. van Assenbergh*

De koopwoningmarkt in Nederland heeft in de tweede helft van de jaren negentig een onstuimige ontwikkeling doorgemaakt, gekenmerkt door een toenemende krapte en forse prijsstijgingen. Interessant is om te onderzoeken hoe ver de huizenprijzen nog kunnen stijgen en of tijdstip en omvang van een omslag in de prijsontwikkeling kunnen worden voorspeld. In dit artikel wordt eerst ingegaan op de beeldvorming omtrent de huizenprijsontwikkeling. Vervolgens wordt de Rabobank-huizenprijsindicator gepresenteerd en tot slot volgt onze visie op de toekomstige huizenprijsontwikkeling.

Voorspellingen

In het voorjaar zijn er in de media enkele voorspellingen verschenen over de Nederlandse huizenprijs. De Nederlandse Vereniging van Makelaars (NVM) verwacht voor 2000 een gemiddelde prijsstijging van zo'n acht procent. De Nederlandse Vereniging van Bouwondernemers en Ontwikkelaars voorziet tot 2004 een prijsstijging van 35 procent. In

het rapport *Thermometer koopwoningen* van de bouwers komt een berekening van het onderzoeksinstituut OTB evenwel uit op een gemiddelde prijsstijging van circa 25 procent (uitgaande van een gemiddelde hypotheekrente van zes procent en een gemiddelde welvaartsgroei van 2,75 procent per jaar). De bouwers zijn zeer optimistisch over de welvaartsontwikkeling in de komende jaren en schrijven hieraan de voortgaande huizenprijsstijging toe.

ABN AMRO komt in haar *Economisch perspectief* van maart jongstleden tot veel voorzichtiger prijsverwachtingen. De zeer forse prijsstijging in de afgelopen jaren en de opgelopen hypotheekrente hebben de betaalbaarheid van de gemiddelde koopwoning onder druk gezet. In scenarioberekeningen wordt uitgegaan van een geannualiseerde prijsstijging van acht procent in 2000, zes procent in 2001 en gemiddeld 3,2 procent in 2002-2005.

Uit een enquête van DNB die in maart werd gehouden, bleek dat zestig procent van de ondervraagden een verdere stijging van de huizenprijzen voorziet. 33 procent verwacht een stabilisatie. Opmerkelijk is dat tegelijkertijd 57 procent de huizenprijzen te hoog vindt.

Regressiemodellen ontoereikend

Om gevoel te krijgen voor de toekomstige huizenprijsontwikkeling wordt vaak gebruik gemaakt van een regressiemodel voor de huizenprijs. Vervolgens wordt met behulp van scenario's voor de verklarende variabelen iets gezegd over de toekomstige ontwikkeling hiervan. Nadeel van deze aanpak is onder andere dat de verklarende variabelen (zoals de rente) slecht te voorspellen zijn, zodat er moeilijk een waarschijnlijkheidsgehalte aan de scenario's valt toe te rekenen. Bovendien is het niet gemakkelijk om een goed regressiemodel te

ontwikkelen. Uit eigen berekeningen blijkt dat bepaalde verklarende variabelen, zoals economische groei en toename van het aantal huishoudens, statistisch een negatieve bijdrage leveren aan de verandering in de huizenprijzen in de periode 1981-1999. Dit valt moeilijk te interpreteren. Al met al lijkt een regressiemodel niet het geïchte instrument om de huizenprijs te voorspellen. Ook is het moeilijk hiermee omslagen in de prijsontwikkeling op tijd te voorzien.

De Rabobank huizenprijsindicator

Omdat goede ervaringen zijn opgedaan met conjunctuurindicatoren als voorspeller voor omslagen in de conjunctuur, is analoog aan deze methodiek een huizenprijsindicator (zie figuur 1) ontwikkeld. Uitgangspunt hierbij is, dat de huizenprijzen rondom een trend fluctueren. Conjunctuur wordt dan gedefinieerd als datgene wat overblijft na verwijdering van seizoensinvloeden, trendontwikkeling en toevallige schokken. De Rabobank-huizenprijsindicator bestaat uit twee lijnen. De *wijzer* weerspiegelt de feitelijke conjunctuur in de huizenprijsindex (die op basis van NVM- en Kadastergegevens is samengesteld). De *voorspeller* is opgebouwd uit een aantal variabelen, waarvan de conjunctuur een goede samenhang vertoont met de wijzer. Doordat de voorspeller enkele maanden voorloopt op de wijzer, kunnen uitspraken gedaan worden over de toekomstige ontwikkeling van de 'conjunctuur' in de huizenprijsindex.

De voorspeller van de huizenprijsindicator loopt zes maanden voor op de wijzer en toont eind 1999 een omslag. Ook de wijzer lijkt zijn top te gaan bereiken in het eerste halfjaar van 2000. Dit betekent dat de conjunctuur in de huizenprijs een hoogtepunt heeft bereikt. De hierop volgende dalende beweging duidt overigens niet

* De auteur is werkzaam bij de Stafgroep Economisch Onderzoek van de Rabobank. Met dank aan Ard den Reijer.

Constructie van de wijzer

De eerste stap die is gezet om tot een wijzer voor een huizenprijsindicator te komen bestaat uit het verwijderen van eventueel aanwezige seizoenseffecten. Vervolgens wordt de trend uit de tijdreeks verwijderd. Evenals bij de Rabobank-conjunctuur-indicatoren is hierbij gebruik gemaakt van de PAT-methode (PAT = phase average trend), waarbij per cyclus (= periode van versnelling en vertraging) de trend wordt bepaald door een lineaire lijn. De tijdsduur en de richtingscoëfficiënt van de trend verschillen per cyclus. De trendbreuken (1974-'75, 1978, 1982-'83, 1985-'86, 1991) vallen goed te interpreteren als gekeken wordt naar de belangrijkste factoren die de koopwoningmarkt beïnvloeden, zoals de hypotheekrente, het verstrekingsbeleid van financiers, financiële innovaties, de inkomensontwikkeling, de huurontwikkeling (als alternatief voor kopen) en de nieuwbouwproductie.

Na verwijdering van de trend wordt de overgebleven reeks van de huizenprijsindex afgevlakt door het berekenen van voortschrijdende gemiddelden. Hiermee wordt voorkomen dat kleine schommelingen tot een verkeerde interpretatie van de richting van de ontwikkeling van de conjunctuur in de huizenprijzen leiden.

op een daling van de huizenprijs, maar op een minder dan trendmatige prijsstijging. Op dezelfde wijze komt een dal in het verloop van de indicator overeen met een conjunctureel dieptepunt. Een daaropvolgende stijging duidt op een meer dan trendmatige prijsstijging. Een zijwaartse beweging wijst dus op een trendmatige stijging van de huizenprijzen.

De positie van de indicator ten opzichte van de nullijn geeft een beeld van de cumulatieve ontwikkeling van de huizenprijs. Een positieve waarde geeft aan, dat de gemiddelde huizenprijs (sinds het jongste snijpunt met de x-as)

gemiddeld genomen meer dan trendmatig is gegroeid.

Uit figuur 1 kan worden afgeleid, dat de cyclische component in de huizenprijsontwikkeling zijn top recentelijk lijkt te hebben bereikt. Dit betekent dat de huizenprijzen vanaf nu minder dan trendmatig zullen stijgen. De huizenprijzen hoeven overigens niet te gaan dalen. Dat hangt immers grotendeels van de trendontwikkeling in de huizenprijs af.

Trend

Zoals gezegd bij de constructie van de wijzer, blijkt de trend in bepaalde jaren om te slaan naar een ander groeipad (sneller/langzamer stijgen of van stijging naar daling). Een andere manier om de trend te bepalen dan volgens de PAT-methode levert figuur 2, waarbij tussen de gevonden breekpunten in de trend de constante prijsstijging berekend is op basis van lineaire regressie. Volgens deze berekening bedroeg de trendmatige stijging van de gemiddelde huizenprijs in de jaren negentig 9,55 procent. Opmerkelijk is dat in 1999 de gemiddelde huizenprijs fors boven deze trendgroei is uitgestegen. Er lijkt dan ook sprake van doorschieten van de huizenprijzen. Een breekpunt in de trend naar een nog hogere trendgroei lijkt ons gegeven het te verwachten verloop van de onderliggende verklarende variabelen onwaarschijnlijk. Ook dit bevestigt ons beeld, dat de huizenprijzen binnen de afzienbare termijn van circa een jaar minder dan trendmatig zullen kunnen gaan stijgen.

Breekpunten in de trend

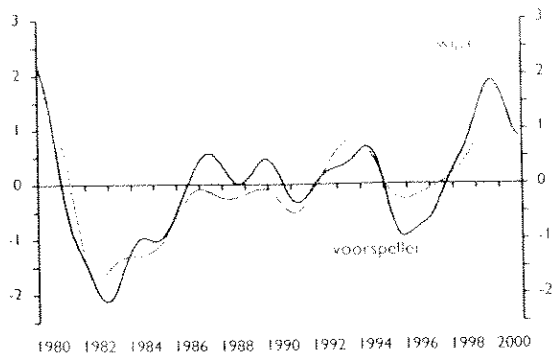
Het is goed mogelijk dat zich momenteel ook een omslag in de trendontwikkeling voordoet. In de jaren negentig stonden voor de vraagfactoren op de koopwoningmarkt immers alle lichten op groen en dat lijkt nu niet meer het geval. De economische groei zat mee, de werkloosheid daalde fors en de inkomensontwikkeling verliep gunstig. De rente daalde voortdu-

rend (met uitzondering van een hobbel in 1994) tot een historisch laagtepunt begin 1999. Demografische ontwikkelingen waren gunstig, want de toename van de groep huishoudens van middelbare leeftijd stimuleerde de vraag naar koopwoningen. De forse huurstijgingen in het begin van de jaren negentig maakten kopen (mede gegeven de rente-ontwikkeling) aantrekkelijker dan huren. De tweeverdienersregeling van begin jaren negentig, waarbij het inkomen van de partner ook volledig mag worden meegerekend bij de bepaling van het maximale hypotheekbedrag, verruimde de financieringsmogelijkheden van huizenkopers, terwijl bovendien het aantal tweeverdieners sterk toenam. Tegenover al deze vraagstimuli kon het aanbod niet meeko-

Constructie van de voorspeller

Gegeven het verloop van de wijzer is de conjunctuurcomponent van 38 variabelen (waarvan enige samenhang met de huizenprijs te verwachten valt) getest op samenhang met de wijzer. De criteria die hierbij werden gehanteerd waren een voldoende samenhang met de wijzer (correlatie van minimaal 0,5), een voorlooptijd van minimaal zes maanden op de wijzer en voldoende samenhang in de omslagpunten. Uiteindelijk bleken slechts vijf variabelen aan deze criteria te voldoen: de lange rente (tienjaars Nederlandse staatsobligaties), de korte rente (driemaands AIBOR/EURIBOR), kredietproductie door aangesloten Rabobanken, verleende bouwvergunningen en een zelf geconstrueerde krapte-indicator voor de markt van bestaande koopwoningen op basis van NVM-gegevens¹. Helaas gaat de historie van de tijdreeksen van de bouwvergunningen en de krapte-indicator niet zo ver terug. Daarom hebben we twee voorspellers gemaakt: een 'brede' voorspeller met de vijf genoemde deelvariabelen vanaf 1986 en een 'smalle' voorspeller met drie deelvariabelen (korte en lange rente en kredietproductie aangesloten Rabobanken) vanaf 1979. Het verloop van beide voorspellers toont een grote gelijkens en de omslagpunten liggen op vrijwel dezelfde tijdstippen. Figuur 1 geeft de Rabobank-huizenprijsindicator weer met de 'smalle' voorspeller. De omslagpunten van de wijzer en voorspeller komen goed overeen en de correlatie tussen beide lijnen bedraagt 0,89. De voorspeller loopt zes maanden voor op de wijzer.

¹ De krapte-indicator is gedefinieerd als het aanbod van koopwoningen in periode t ten opzichte van de verkooprijd in de voorgaande periode ($t-1$) ofwel $\text{aanbod}/\text{verkooprijd} = (\text{verkooprijd} \times \text{aantal transacties}) / \text{verkooprijd}$.



Figuur 1.
De huizenprijs-
indicator

men. In de eerste helft van de jaren negentig liep het kwantitatieve woningtekort zelfs op.

Op dit moment lijkt er geleidelijk een afkoeling te komen aan de vraagzijde. De zeer forse huizenprijsstijgingen en de inmiddels weer opgelopen rente hebben de betaalbaarheid van koopwoningen verslechterd. Ook is het de vraag of de huidige hoogconjunctuur ongelimiteerd kan voortgaan. Voor 2000 zijn de verwachtingen omtrent de conjunctuur nog gunstig gestemd. Voor de daarop volgende jaren mag evenwel een geleidelijke afzwakking worden verwacht, als het effect van de recente rentestijging zich in de reële sfeer uit, de Amerikaanse economie afkoelt en het groeitempo in Europa niet verder versnelt. Bovendien kunnen de aanpassingen in het fiscale regime per 1 januari 2001 ongunstig uitpakken voor de potentiële woningkoper: denk aan maatregelen als lagere marginale tarieven en dus een geringere hypotheekrente-af trek en geen aftrek meer op de tweede of

volgende woning. Voorts is het waarschijnlijk dat het stimulerende effect op de koopwoningvraag vanuit de tweeverdienersregeling, de demografische ontwikkeling en het financiële verschil tussen kopen en huren geleidelijk aan kracht verliezen. Al met al is op termijn een geleidelijke afzwakking van de vraag naar koopwoningen te voorzien. Onze verwachting is dan ook, dat de trendmatige toename van de gemiddelde huizenprijs de komende jaren juist weer wat lager zal liggen dan in de jaren negentig en eerder overeen kan komen met de trendgroei in de tweede helft van de jaren tachtig (circa vier procent per jaar).

Conclusie

De Rabobank-huizenprijsindicator, die een weerspiegeling geeft van de conjunctuur in de huizenprijs, toont dit jaar een duidelijke omslag. Verwacht kan worden dat de huizenprijzen minder dan trendmatig zullen gaan stijgen. Aangezien de trendmatige prijsstij-

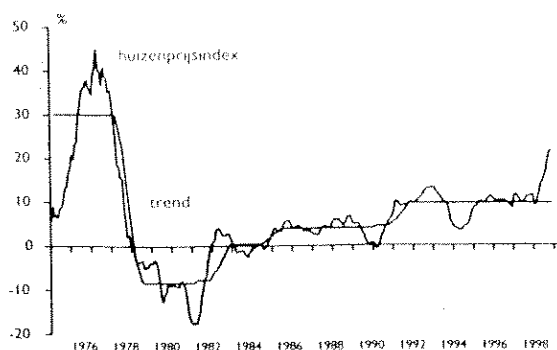
ging aanzienlijk is, worden evenwel geen prijsdalingen verwacht.

Ook de trend zelf zal de komende jaren vermoedelijk lager gaan uitvallen, doordat de belangrijkste vraagimpulsen wegvallen en de betaalbaarheid van koopwoningen in bepaalde segmenten in het gedrang is gekomen. In 2000 kan het stijgingspercentage toch nog hoog uitvallen vanwege aanzienlijke overloopeffecten uit 1999. Zelfs als de gemiddelde huizenprijs constant blijft, zou de gemiddelde prijsstijging dit jaar al 8¼ procent (jaar-op-jaar) bedragen. Een daling in de gemiddelde jaarlijkse huizenprijs wordt de komende jaren niet verwacht.

Een belangrijke beleidsimplicatie van de voorzien afkoeling van de prijsontwikkeling op de woningmarkt is dat maatregelen om de woningmarkt te temperen vanuit de politiek of De Nederlandsche Bank, zoals bijvoorbeeld het beperken van het hypotheekrecht bij tophypotheek, nu onnodig of zelfs ongewenst zijn. Daarmee zou de betaalbaarheid van koopwoningen immers verder kunnen afnemen, waardoor er vraaguitval en stagnatie op de woningmarkt kan ontstaan. De grote positieve vermogenseffecten van de huizenprijsstijgingen die de consumptie in de laatste jaren stevig hebben gestimuleerd, zouden dan juist in hun tegendeel kunnen omslaan. Dit zou de reële economie hard kunnen raken. Daarmee zouden de beleidsmakers de macro-economische instabiliteit, die zij zeggen te willen voorkomen, juist kunnen veroorzaken.

Al met al voorzien we een afvlakking in de stijging van de huizenprijzen. De komende jaren kan de koopwoningmarkt in rustiger vaarwater terechtkomen, waarbij de te verwachten stijging van de huizenprijs gekarakteriseerd kan worden als "redelijk evenwichtig".

Wim van Assenbergh



Figuur 2.
Huizenprijsindex
en trend,
jaar-op-jaar
mutatie 1976-
1998