Statistische Signaalverwerking

Statistische Signaalverwerking

R.L. Lagendijk
J. Biemond

Technische Universiteit Delft
Faculteit Informatietechnologie en Systemen

VSSD

Eerste druk 1994, verbeterd 1999

Uitgegeven door:
VSSD
Leeghwatertsraat 42, 2628 CA Delft, The Netherlands
tel. +31 15 2782124, fax +31 15 2787585, e-mail: hlf@vssd.nl
internet: http://www.vssd.nl/hlf
URL over dit boek: http://www.vssd.nl/hlf/informatie.html

Gedrukte editie
ISBN-10 90-407-1258-1
ISBN-13 978-90-407-1258-6

Elektronische versie:
ISBN-10 90-6562-142-3
ISBN-13 978-90-6562-142-9

Trefw.: statistische signaalverwerking / detectietheorie / schattingstheorie

Voorwoord

In de laatste decennia heeft de signaalverwerking sterke veranderingen doorgemaakt. Door de opkomst van geïntegreerde digitale signaalverwerkingsprocessoren en door de ontwikkeling van gespecialiseerde geïntegreerde circuits voor complexe signaalverwerkingsalgoritmen heeft de bewerking van elektrische signalen in real-time zich ontwikkeld van abstract idee tot dagelijkse praktijk. Algoritmen die lange tijd slechts conceptuele oplossingen van theoretische problemen waren, worden momenteel routinematig gebruikt in een groot aantal toepassingen binnen de Elektrotechniek (telecommunicatie: detectie en schatting van signalen, netwerktheorie: realisatie van algoritmen, regeltechniek: regelen en meten van systemen, en informatietheorie: transport en extractie van informatie), maar ook daar buiten. De statistische signaalverwerking concentreert zich op dat deel van de signaalverwerking waarin het gebruik van stochastische signaalmodellen en statistische kenmerken essentieel is. Bij alle te bespreken onderwerpen draait het telkens om het scheiden van informatie en verstoring, waarbij de verstoring meestal gemodelleerd wordt als een (additieve) ruiscomponent.

Dit boek is gebaseerd op het college *Statistische Signaalverwerking* zoals dat sinds 1983 gegeven wordt aan de faculteit der Elektrotechniek van de Technische Universiteit in Delft. Door de jaren heen is het studiemateriaal geëvolueerd van collegeaantekeningen, via collegediktaten tot het voor u liggende boek.

Aanpak en inhoud

Dit boek behandelt de elementaire onderwerpen uit de statistische signaalverwerking, zodat na bestudering de (vrij uitgebreide) literatuur op dit gebied goed toegankelijk is. Er wordt ingegaan op detectie- en schattingstheorie, lineaire kleinste-kwadraten schatters (Wiener, Kalman schatters en adaptieve filters) en methoden voor het schatten van spectra.

In dit boek staan de methodologieën voorop en niet zozeer de toepassingen, alhoewel de behandelde stof wel zoveel mogelijk geïllustreerd wordt aan de hand van voorbeelden uit de praktijk van de signaalverwerking. De methoden die besproken worden, vormen de basis voor veel gerealiseerde systemen. Voordat een gekozen methode in de praktijk gebruikt kan worden, zal eerst een geschikte implementatie gevonden moeten worden. Dit niet onbelangrijke architectuur- en hardware-aspect wordt slechts op enkele plaatsen in het boek genoemd. De nadruk van dit boek ligt dus meer op het

ontwikkelen van methoden en algoritmen dan op de realisatie hiervan.
In hoofdstuk 1 wordt een nadere toelichting gegeven op de doelstellingen van de statistische signaalverwerking aan de hand van enkele voorbeelden. Vervolgens wordt een resumé gegeven van de belangrijkste begrippen uit de kansrekening en de theorie van de stochastische processen. Op basis hiervan wordt daarna een aantal stochastische signaalmodellen geïntroduceerd (AR, MA, ARMA), omdat deze essentieel zijn voor de ontwikkeling van Kalman filters en spectrum schatters.
In hoofdstuk 2 wordt het meest elementaire probleem uit de statistische signaalverwerking besproken, namelijk het beslissen of een signaal wel of niet aanwezig is. Bij dit detectieprobleem zijn de beslissingscriteria van groot belang. Een toepassing van de detectietheorie is te vinden in het matched filter.
Nadat een signaal gedetecteerd is, kunnen hieraan bepaalde eigenschappen gemeten worden. Bij stochastische signalen spreekt men dan over het schatten van signaalparameters, wat het onderwerp is van hoofdstuk 3. Omdat bij deze schatting altijd afwijkingen optreden ten opzichte van de werkelijke parameterwaarde, is ook bij de schattingstheorie de keuze van een foutcriterium belangrijk. Vanwege implementatieoverwegingen beperkt men zich vaak tot lineaire schatters voor parameters, dat wil zeggen, schatters die een gewogen gemiddelde van waargenomen signaalwaarden zijn. Dit onderwerp vormt de overgang naar de lineaire filtertheorie (hoofdstuk 4, 5 en 6).
In de lineaire filtertheorie probeert men vanuit een waargenomen verstoord tijdsignaal een zo goed mogelijke schatting te maken van het onverstoorde (dat wil zeggen, het originele) informatiedragende tijdsignaal. We beperken ons daarbij tot lineaire schatters die een kwadratisch foutcriterium minimaliseren. Vervolgens wordt een algemeen geldig stelsel van vergelijkingen afgeleid, waaruit de weegcoëfficiënten van de lineaire schatter bepaald dienen te worden. Onder zekere voorwaarden kan dit redelijk eenvoudig gebeuren. We spreken dan van Wiener schatters, waarvan een aantal mogelijke vormen in hoofdstuk 4 aan de orde komt.
Veelal zijn Wiener schatters rekentechnisch onaantrekkelijk. In hoofdstuk 5 komt een klasse van rekentechnisch meer aantrekkelijke schatters aan de orde, namelijk de Kalman schatters. De twee belangrijkste representanten van deze categorie schatters worden afgeleid, namelijk de Kalman 1-staps predictor en het Kalman filter. Uitbreidingen naar meer complexe situaties worden kort genoemd.
Kalman en Wiener schatters zijn gebaseerd op stochastische modellen voor de signalen en de ruiscomponenten. In sommige situaties is echter meer voorkennis over verstoring of signaal beschikbaar. We kunnen dan een lineaire schatter gebruiken in de zogenaamde ‘noise cancelling’ configuratie. Tevens blijkt dan een hoge mate van adaptiviteit aan signaal en ruis mogelijk te zijn, wat schatters met betere prestaties levert dan de Kalman of Wiener schatters. Hoofdstuk 6 beschrijft de basisvorm van het ‘noise cancelling’ concept, behandelt een efficiënte implementatie hiervan en geeft tevens een aantal toepassingen.

Zoals gezegd hebben vele lineaire schatters voorkennis nodig over de te schatten signalen, namelijk het vermogensdichtheidsspectrum of de autocorrelatiefunctie. Helaas zijn beide niet zomaar beschikbaar en zullen dus uit een gemeten signaal geschat moeten worden. In hoofdstuk 7 wordt eerst een aantal klassieke spectrum-schattingsmethoden besproken. Vervolgens wordt uitgebreid ingegaan op de belangrijke klasse van (AR/MA/ARMA) model-gebaseerde spectrum schatters.

Vereiste voorkennis

Om aansluiting te kunnen vinden bij het materiaal dat in dit boek besproken wordt, is voorkennis vereist over de basisbegrippen rond kansrekening en stochastische processen: (vectoriële) stochastische variabele, kansdichtheidsfunctie, Gaussische stochastische variabele, verwachting, correlatie en covariantie, onafhankelijkheid, ongecorreleerdheid, orthogonaliteit, stationariteit en ergodiciteit. Hoofdstuk 1 van dit boek geeft een resumé van de vereiste voorkennis, en introduceert tevens diverse notaties. Concepten uit de lineaire algebra worden bekend verondersteld, zoals matrix-vector rekenen en het begrip orthogonaliteit. Er wordt tevens van uitgegaan dat de begrippen Fourier transformatie (voornamelijk op tijd-discrete signalen) en z-transformatie en de eigenschappen hiervan bekend en operationeel zijn.

Dankbetuigingen

Het samenstellen van collegeaantekeningen, diktaten of boeken vergt interactie tussen docent en studenten. Ook aan dit boek hebben op een groot aantal plaatsen studenten hun bijdrage geleverd door het stellen van gerichte vragen, het opmerken van fouten, en soms ook door het suggereren van een wijziging in aanpak of formulering. De auteurs bedanken iedereen die op een dergelijke wijze aan verbetering van het studiemateriaal heeft meegewerkt.
Alhoewel het materiaal met uiterste zorg is samengesteld, zijn onvolkomenheden nooit uitgesloten. De auteurs stellen daarom opmerkingen en suggesties met betrekking tot de inhoud zeer op prijs.

Delft, januari 1994
Inald Lagendijk
Jan Biemond

Inhoud

VOORWOORD 5

1. INTRODUCTIE IN DE STATISTISCHE SIGNAALVERWERKING 13
   1.1. Toepassingen 13
   1.2. Resumé stochastische variabelen 17
      1.2.1. Het kansbegrip 17
      1.2.2. Gezamenlijke en conditionele kans, onafhankelijkheid 19
      1.2.3. Stochastische variabelen 19
      1.2.4. Vectoriële stochastische variabelen 22
      1.2.5. Stochastische processen en signalen 26
      1.2.6. Gemiddelde, correlatie en covariantie van stochastische signalen 28
      1.2.7. Stationariteit en ergodiciteit 30
      1.2.8 Autocorrelatiefunctie en vermogensdichtheidsspectrum 31
      1.2.9. Witte ruis 34
   1.3. AR, MA en ARMA signaalmodellen 36
      1.3.1. Klassen van signaalmodellen 36
      1.3.2. Autoregressief signaalmodel 38
      1.3.3. Moving-average signaalmodel 39
      1.3.4. Autoregressief moving-average signaalmodel 40
      1.3.5. Toestandsbeschrijving van ARMA modellen 41
   1.4. Samenvatting en belangrijke termen 45
      Lijst van belangrijke termen 45

2. DETECTIETHEORIE 46
   2.1. Inleiding 46
   2.2. Het toetsen van hypothesen 47
   2.3. Beslissingscriteria 50
      2.3.1. Bayes criterium 51
      2.3.2 MAP-criterium 55
      2.3.3 Minimale foutkans criterium 55
      2.3.4 Minimax criterium 57
      2.3.5 Neyman-Pearson toets 59
      2.3.6. Maximum likelihood 63

2.4. Meervoudige waarnemingen 63
2.5. Sequentiële detectie 66
2.6. Detectie van bekende signalen in witte ruis 68
2.6.1. Tijd-discrete geval 68
2.6.2. Tijd-continue geval 69
2.6.3. Matched filter implementatie van de ontvanger 71
2.7. Samenvatting en belangrijke termen 73
Lijst van belangrijke termen 75
Vraagstukken 75

3. SCHATTINGSTHEORIE 80
3.1. Inleiding 80
3.2. Stochastische parameters: Bayes schatters 81
3.2.1. Minimum-variantie schatter 83
3.2.2. Mediaan schatter 84
3.2.3. Maximum a posteriori schatter 84
3.3. Deterministische parameters: maximum likelihood schatters 90
3.4. Eigenschappen van schatters 92
3.5. Lineaire kleinste kwadraten schatters voor stochastische parameters 98
3.6. Samenvatting en belangrijke termen 102
Lijst van belangrijke termen 102
Vraagstukken 103

4. LINEAIRE KLEINSTE KWADRATEN SCHATTERS VOOR SIGNALEN: WIENER FILTERS 105
4.1. Inleiding 105
4.2. Wiener-Hopf vergelijking 109
4.2.1. Orthogonaliteitsprincipe 109
4.2.2. Tijd-variante Wiener-Hopf vergelijking 112
4.2.3. Tijd-invariante Wiener-Hopf vergelijking 113
4.3. Niet-causale Wiener filter 115
4.4. Causale Wiener schatter met oneindig geheugen 119
4.4.1. Whitening benadering 119
4.4.2. Spectrale factorisatie 120
4.4.3. Oplossing voor de causale schatter 122
4.5. Causale Wiener schatter met eindig geheugen 125
4.6. Gegeneraliseerde Wiener filter 127
4.6.1. Matrix-vector formulering 128
4.6.2. Efficiënte implementatie met unitaire transformaties 130
4.6.3. Fourier domein filter 132
4.7. Tijd-continue Wiener schatters 134

4.8. Samenvatting en belangrijke termen 134
Lijst van belangrijke termen 136
Vraagstukken 136

5. KALMAN FILTERS 141
5.1. Inleiding 141
5.2. Wold-decompositie 145
5.3. Kalman 1-staps predictor 149
5.3.1. Afleiding 149
5.3.2. Eigenschappen van het innovatieproces 152
5.3.3. Kalmanversterking 154
5.3.4. Ricatti differentievergelijking 155
5.3.5. Begincondities en stabiliteit 156
5.3.6. Steady-state gedrag 157
5.4. Kalman filter 160
5.4.1. Afleiding 160
5.4.2. Kalmanversterking en Ricatti differentievergelijking 162
5.4.3. Steady-state gedrag 162
5.5. Uitbreidingen op de standaard Kalman schatters 164
5.6. Beperkingen en voorbeelden 165
5.6.1. Beperkingen 165
5.6.2. Filteren van beelden 166
5.6.3. Filteren van niet-stationaire ruis 167
5.7. Samenvatting en belangrijke termen 168
Lijst van belangrijke termen 169
Vraagstukken 169

6. ADAPTIEF FILTEREN IN DE ‘NOISE CANCELING’ CONFIGURATIE 172
6.1. Inleiding 172
6.2. Het ‘adaptive noise canceling’ principe 172
6.3. Het adaptieve filter 175
6.3.1. ‘Single-input’ adaptief filter 176
6.3.2. ‘Multiple-input’ adaptief filter 177
6.3.3. De optimale weegvector 178
6.4. Het LMS-algoritme 180
6.5. Toepassingen 182
6.6. Samenvatting en belangrijke termen 190
Lijst van belangrijke termen 192
Vraagstukken 192

7. SPECTRUM SCHATTEN 194
7.1. Inleiding 194
7.2. De ‘taper-and-transform’ methoden 198
7.2.1. Periodogram spectrum schatter 199
7.2.2. Oplossend vermogen en leakage 199
7.2.3. Variantie van de periodogram schatter 203
7.2.4. Middeling van het periodogram 207
7.2.5. Blackman-Tukey schatter 208
7.3. Autoregressieve spectrum schatter 210
7.3.1. AR modellering 210
7.3.2. Yule-Walker vergelijkingen 211
7.3.3. Schatting van de autocorrelatiecoëfficiënten 214
7.3.4. Extrapolatie van de autocorrelatiefunctie 216
7.3.5. Orde van het model 219
7.4. Moving-average en autoregressieve moving-average spectrum schatter 221
7.4.1. MA modellering 221
7.4.2. AR-oneindig benadering 223
7.4.3. Autoregressieve moving-average spectrum schatter 224
7.5. Samenvatting en belangrijke termen 226
Lijst van belangrijke termen 226
Vraagstukken 227

APPENDIX: TWEEZIJDIGE Z-TRANSFORMATIE 229

ANTWOORDEN 232

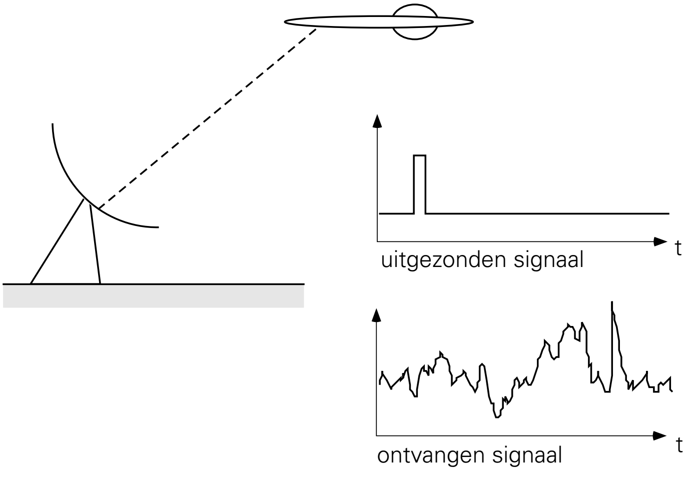
LITERATUUR 240

TREFWOORDENLIJST 242

1 Introductie in de statistische signaalverwerking

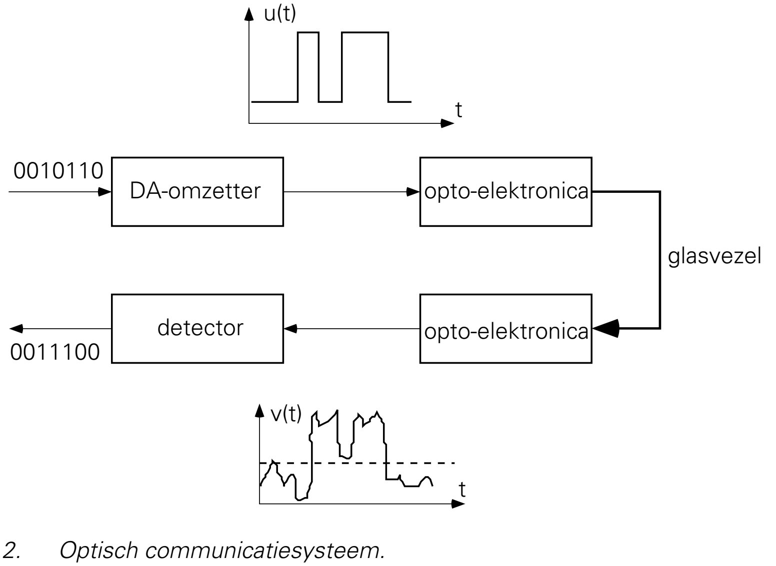
1.1. Toepassingen

Statistische signaalverwerking is dát gebied van de signaalverwerking dat zich bezighoudt met het scheiden van informatie en verstoring in een waargenomen (gemeten) signaal. Het scheiden van informatie en verstoring is relatief eenvoudig als de informatie deterministisch is of als de verstoringen bekende fenomenen zijn die nauwkeurig beschreven kunnen worden. Het wordt echter anders wanneer de signalen stochastische grootheden zijn, zoals in de praktijk meestal het geval is. De precieze beschrijving van informatie en verstoring is dan niet langer mogelijk, en zullen er statistische kenmerken zoals verwachting, variantie en correlatie gebruikt moeten worden. Tevens is het bij stochastische signalen meestal niet mogelijk de gezochte informatie exact te bepalen en moeten schattingsmethoden ingezet worden om een (gemiddeld gezien) zo goed mogelijke scheiding tussen informatie en verstoring te verkrijgen.
Een aantal voorbeelden uit de praktijk kan het nut van statistische signaalverwerking illustreren. Stel we willen met radar bepalen of zich een vliegtuig in een bepaald deel van het luchtruim bevindt (figuur 1.1). Hiertoe wordt een radarbundel op het betreffende deel van het luchtruim gericht en wordt er vervolgens bepaald of zich reflecties voordoen. Er doen zich nu twee vragen voor. In de eerste plaats moet



Figuur 1.1. Volgen van een bewegend object met radar.

beslist worden of er wel of geen reflectie optreedt. En als zich dan een reflectie voordoet, is vervolgens de vraag wat de eigenschappen van het betreffende object (vliegtuig?) zijn, zoals positie en snelheid. Als er niets is waartegen de radarpuls kan reflecteren, zal er in ideale omstandigheden ook inderdaad niets waargenomen worden. Indien er wel een object aanwezig is, dan ontvangen we een reflectie τ seconden na het uitzenden van een radarpuls. De afstand tot het object is dan $\ell = 0{,}5c\tau$, waarbij c de voortplantingssnelheid van elektromagnetische golven is.
Helaas gaan bovenstaande ideale omstandigheden in de praktijk nooit op. We hebben altijd te maken met stochastische verstoringen (ruis) van het waargenomen signaal, bijvoorbeeld ten gevolge van verstoringen in het transmissiemedium (multipad-effecten, irrelevante reflecties) en ruis in de detectieapparatuur. Hierdoor is het mogelijk dat een reflectie wordt gedetecteerd terwijl er geen object aanwezig is, of dat de reflectie op een verkeerd tijdstip wordt gedetecteerd, wat tot een verkeerde bepaling van ℓ leidt. Aan de andere kant is het ook mogelijk dat een reflectie gemist wordt, waardoor een vliegtuig niet gedetecteerd wordt. In beide gevallen leidt de aanwezigheid van ruis tot een verkeerde beslissing of verkeerde afstandsbepaling.
De statistische signaalverwerking probeert nu de kans op een foute beslissing dan wel de grootte van de fout in de afstandsbepaling te minimaliseren, zodat uit de verstoorde waarnemingen toch de gezochte informatie (wel/geen vliegtuig, welke afstand) zo goed mogelijk geschat kan worden.
Een vergelijkbare situatie doet zich voor in de digitale transmissie van signalen over glasvezels (figuur 1.2). Aan de zendzijde wordt elke T seconden ofwel een rechthoekige puls (een digitale ‘1’) ofwel geen puls (een digitale ‘0’) verzonden. In ideale omstandigheden kan aan het eind van de glasvezel een sensor bepalen of er wel of



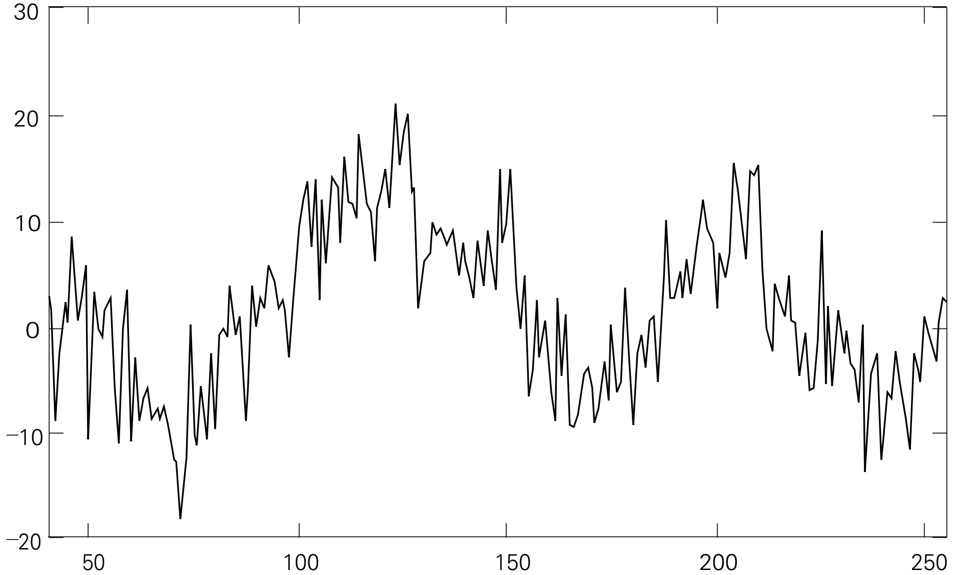
Figuur 1.2. Optisch communicatiesysteem.

geen puls was uitgezonden. Helaas zal de puls tijdens de transmissie verstoord worden door bijvoorbeeld looptijdverschillen, dempingen en imperfecties in de glasvezel. Bovendien is de sensor zelf niet ruisvrij. Hierdoor is het mogelijk dat de sensor geen puls detecteert terwijl die wel aanwezig was en vice versa. Met behulp van statistische signaalverwerkingsmethoden kan een detectiesysteem ontworpen worden dat gemiddeld gezien zo min mogelijk fouten maakt, wat de betrouwbaarheid van de communicatieverbinding vergroot.

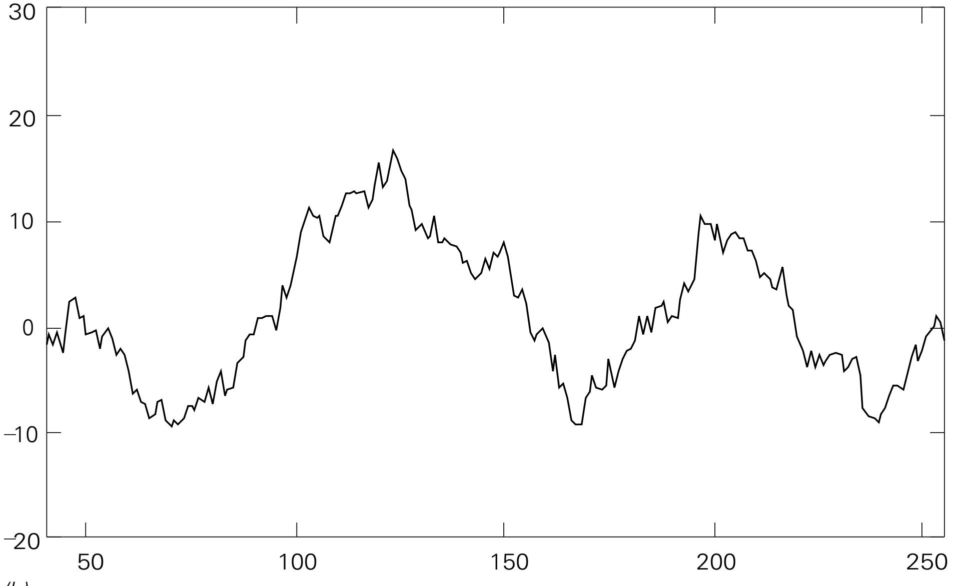
Een heel ander toepassingsgebied is het filteren van signalen. In veel situaties is een signaal dat afkomstig is van een fysisch proces imperfect door verstoringen die ontstaan tijdens het verrichten van de metingen of ten gevolge van de wijze waarop het signaal gemeten wordt. Denk hierbij bijvoorbeeld aan het meten van een EEG (elektro-encefalogram). Deze vorm van afbeelden van hersenactiviteiten vindt toepassing in een groot aantal klinische situaties. De EEG signalen zijn het resultaat van zeer kleine potentiaalverschillen in de hersenen en zijn daarom moeilijk te registreren. Allerlei verstorende effecten zoals spieractiviteiten, omgevingsinvloeden en meetruis in de sensoren zorgen ervoor dat het gemeten EEG een verruiste versie is van de feitelijke hersenactiviteiten. Door middel van statistische signaalverwerkingsmethoden zoals Wiener, Kalman en adaptieve schatters is het mogelijk de verstorende componenten te reduceren (zie figuur 1.3). Dit maakt vervolgens de interpretatie (bijvoorbeeld een medische diagnose) van de gemeten signalen eenvoudiger en betrouwbaarder.

Andere voorbeelden uit deze categorie van statistische signaalverwerking zijn de regeltechniek (meten/regelen/modelleren van industriële processen), de beeldverwerking (analyse en codering van beelden of beeldsequenties), de robotica (positie/oriëntatie/snelheidsmetingen), de navigatie (global position systems) en de biomedische techniek.

In elk van de genoemde toepassingen is er sprake van *waarnemingen*, *informatiedragende signalen* en *verstoringen*. In figuur 1.4 worden deze termen in relatie tot elkaar geplaatst. Een fysisch of andersoortig proces genereert een zeker informatiedragend signaal, dat echter niet rechtstreeks kan worden waargenomen (aangegeven met ????). De waarneming (meting) die verkregen wordt, is een verstoorde versie van het informatiedragende signaal. Door modellering van het oorspronkelijke proces en de verstoringen kunnen schattingsalgoritmen worden afgeleid die optimaal zijn volgens een bepaald criterium. Toepassen van deze algoritmen op de waarneming leidt dan tot een schatting van het gezochte signaal of informatie. In de statistische signaalverwerking wordt dus altijd uit de (verstoorde) waarnemingen informatie onttrokken via schattingsalgoritmen. Stochastische modellen voor de diverse signalen spelen hierbij een cruciale rol.

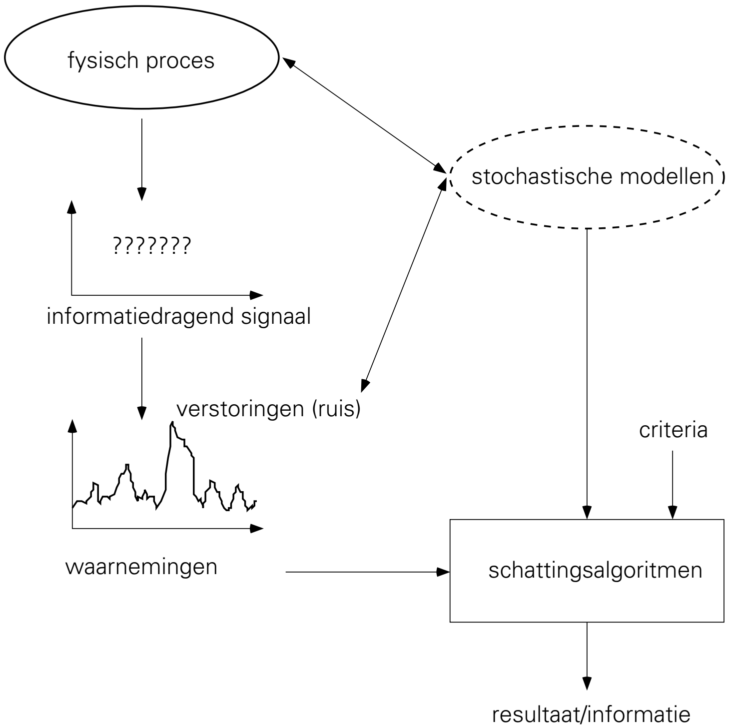


(a)



(b)

Figuur 1.3. *Filteren van signaal met ruis.*



Figuur 1.4. Ingrediënten van de statistische signaalverwerking.

1.2. Resumé stochastische variabelen

Deze paragraaf geeft een resumé van enkele van de belangrijkste begrippen uit de kansrekening en stochastische processen, waarbij de nadruk wordt gelegd op de aspecten die voor de modellering van signalen belangrijk zijn. Een meer diepgaande presentatie van het materiaal in deze paragraaf is te vinden in elk goed tekstboek over waarschijnlijkheidsrekening, bijvoorbeeld [1] en [2]. Paragraaf 1.3 bouwt verder op de hier besproken begrippen en introduceert de belangrijke klasse van autoregressieve (AR), moving-average (MA) en autoregressieve moving-average (ARMA) signaalmodellen.
Opgemerkt wordt dat deze paragraaf zich beperkt tot de behandeling van *reëel-waardige* stochastische variabelen en signalen.

1.2.1. Het kansbegrip

Het deel van de wiskunde dat zich bezighoudt met het bestuderen van verschijnselen die niet in detail voorspelbaar zijn, is de waarschijnlijkheids- of kansrekening. In de kansrekening wordt verondersteld dat er sprake is van een *experiment,* dat goed omschreven moet worden om dubbelzinnigheden en onduidelijkheden te vermijden.

Bovendien wordt verondersteld dat het experiment willekeurig vaak herhaald kan worden onder identieke omstandigheden. De laatste veronderstelling waarvan uitgegaan wordt, is dat men aan kan geven welke *uitkomsten* mogelijk zijn bij het uitvoeren van een experiment. De set van mogelijke uitkomsten noemt men de *uitkomstenruimte*.
Als voorafgaande veronderstellingen waar zijn, dan is het mogelijk om te spreken over de *kans* dat een experiment een bepaalde uitkomst aanneemt. Aan elke uitkomst wordt dan een getal verbonden tussen 0 en 1 dat aangeeft hoe groot de kans is dat deze uitkomst zich voordoet.

Men kan slechts over de kans van een uitkomst spreken als een experiment willekeurig vaak herhaald kan worden. Definiëren we de *relatieve frequentie* of *frequentiecoëfficiënt* van een uitkomst A_i als

$$f(A_i) = \frac{N(A_i)}{N} \qquad 0 \le f(A_i) \le 1 \tag{1.1}$$

waarbij N het aantal malen herhaling van het experiment is, en $N(A_i)$ het aantal malen dat uitkomst A_i voorkomt, dan kan de kans op A_i gedefinieerd worden als:

$$P(A_i) = \lim_{N\to\infty} f(A_i) = \lim_{N\to\infty} \frac{N(A_i)}{N} \tag{1.2}$$

Een op deze manier gedefinieerde kans is een hypothetische grootheid, omdat oneindig veel waarnemingen gedaan zouden moeten worden. In feite levert een eindig aantal waarnemingen slechts een schatting op van de kans. Om deze reden wordt overgegaan op een axiomatische definitie.

Eerst wordt een *gebeurtenis* gedefinieerd als een verzameling van één of meer uitkomsten. De uitkomstenruimte S komt overeen met de zogenaamde zekere gebeurtenis (Ω) waarin alle uitkomsten opgenomen zijn. Aan elke gebeurtenis wordt nu een getal, de kans of waarschijnlijkheid, toegekend dat voldoet aan een drietal axioma's:

axioma 1: Voor elke gebeurtenis A_i geldt: $P(A_i) \ge 0$,
axioma 2: Voor de zekere gebeurtenis geldt: $P(\Omega) = 1$,
axioma 3: Zijn er N disjuncte gebeurtenissen A_1, A_2, ..., A_N, dan moet gelden:

$$P(A_1 \cup A_2 \cup \ldots \cup A_N) = P(A_1) + P(A_2) + \ldots + P(A_N) \tag{1.3}$$

Uitgaande van deze axioma's kunnen verdere eigenschappen van kansen worden afgeleid, zoals

$$P(A_{iC}) = 1 - P(A_i)$$

$$P(\emptyset) = 0 \tag{1.4}$$

$0 \le P(A_i) \le 1$

$P(A_i \cup A_j) = P(A_i) + P(A_j) - P(A_i \cap A_j)$

1.2.2. Gezamenlijke en conditionele kans, onafhankelijkheid

Soms kan men bij een enkel experiment twee soorten uitkomsten onderscheiden, zodat er in feite sprake is van een combinatie van twee subexperimenten. We kunnen nu de kans op de gebeurtenis $(A_i \cap B_j)$ opvatten als de kans op het gezamenlijk (tegelijkertijd) optreden van A_i en B_j, en wordt $P(A_i \cap B_j) = P(A_i, B_j)$ een *gezamenlijke kans* genoemd. Uit de gezamenlijke kansen zijn de *marginale kansen* $P(A_i)$ en $P(B_j)$ direct te bepalen. Het omgekeerde is niet noodzakelijkerwijs waar.

Naast de marginale en gezamenlijke kans bestaat er nog een derde vorm, namelijk de *conditionele kans*. Die doet zich voor wanneer we de kans op een bepaalde gebeurtenis A_i willen weten indien een andere gebeurtenis B_j reeds is opgetreden (en dus bekend is). De conditionele kans van A_i onder de voorwaarde (of gegeven) B_j wordt gedefinieerd als:

$$P(A_i / B_j) = \frac{P(A_i \cap B_j)}{P(B_j)} \qquad \text{mits } P(B_j) > 0 \tag{1.5}$$

Een belangrijke relatie die vaak bij manipuleren van conditionele kansen gebruikt wordt, is de regel van Bayes:

$$P(A_i / B_j) = \frac{P(B_j / A_i) P(A_i)}{P(B_j)} = \frac{P(B_j / A_i) P(A_i)}{\sum_i P(A_i) P(B_j / A_i)} \tag{1.6}$$

In het algemeen is de conditionele kans $P(A_i/B_j)$ groter of kleiner dan de marginale kans $P(A_i)$. Echter in het geval dat $P(A_i/B_j) = P(A_i)$ heeft kennis over de gebeurtenis B_j blijkbaar geen invloed op de conditionele kans van A_i. We noemen de gebeurtenissen dan *stochastisch onafhankelijk*. Meestal wordt onafhankelijkheid als volgt geverifieerd:

$$P(A_i, B_j) = P(A_i) P(B_j) \quad \Leftrightarrow \quad \text{stochastisch onafhankelijk} \tag{1.7}$$

1.2.3. Stochastische variabelen

Een stochastische variabele is formeel gesproken een numerieke functie die aan elke uitkomst van een experiment een *reëel-waardig getal* koppelt. Gebeurtenissen zijn nu eenvoudig te definiëren als een deelverzameling van de reële as. Wanneer de stochastische variabele slechts waarden x_i aanneemt die individueel aanwijsbaar zijn, spreken we van een discrete stochastische variabele. Bij continue stochastische variabelen zijn de uitkomsten van het experiment niet individueel aanwijsbaar.

Het gedrag van een stochastische variabele X wordt volledig beschreven door de *cumulatieve distributiefunctie* (cdf) $F(x)$, gedefinieerd als:

$$F(x) = P(X \leq x) \qquad -\infty < x < \infty \tag{1.8}$$

Uit de definitie van een stochastische variabele en de eigenschappen van kansen volgen ondermeer de volgende eigenschappen van de cumulatieve distributiefunctie:

$$\begin{aligned}
&F(-\infty) = 0 \\
&F(\infty) = 1 \\
&0 \leq F(x) \leq 1 \\
&F(a) \leq F(b) \qquad \text{voor } a < b \\
&\lim_{\varepsilon \downarrow 0} F(x + \varepsilon) = F(x) \\
&P(X > x) = 1 - F(x)
\end{aligned} \tag{1.9}$$

Vanuit de cumulatieve distributieverdeling wordt de *kansdichtheidsfunctie* (pdf) $p(x)$ van de stochastische variabele X gedefinieerd als

$$p(x) = \frac{\mathrm{d}}{\mathrm{d}\,x} F(x) \tag{1.10}$$

waarbij discontinuïteiten in $F(x)$ delta-impulsen in $p(x)$ opleveren. Voor discrete stochastische variabelen X wordt soms ook van de zogenaamde *kansfunctie* p_i gebruik gemaakt, gedefinieerd als:

$$p_i = P(X = x_i) \tag{1.11}$$

In praktische situaties is vaak van een stochastische variabele niet de volledige kansdichtheidsfunctie bekend, maar is slechts kennis over gemiddelden beschikbaar. Dit leidt tot een meer globale en daardoor minder precieze omschrijving van het gedrag van een stochastische variabele. Veelal worden de *verwachting, verwachte waarde* of *eerste moment*

$$\mu = E[X] = \int_{-\infty}^{\infty} x p(x)\,\mathrm{d}\,x \tag{1.12}$$

en de *variantie* of *tweede centrale moment*

$$\sigma^2 = \mathrm{var}[X] = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)\,\mathrm{d}\,x \tag{1.13}$$

gebruikt.