

Introduktion til Tidsrækker

Preben Blæsild

Institut for Matematik
2013

Copyright © 2013 Preben Blæsild

Layout af Lars Madsen

Sat med Kp-Fonts, memoir-klassen og pdfL^AT_EX

Indhold

Forord	iii
1 Indledning	1
1.1 Eksempler	1
2 Stokastiske processer	6
2.1 Stationære processer	8
2.2 ARMA-processer	14
2.3 Egenskaber ved simple ARMA-processer	26
AR(1)-processen	26
MA(1)-processen	28
AR(2)-processen	30
MA(2)-processen	34
ARMA(1,1)-processen	37
3 Empiriske tidsrækker I	40
3.1 Endimensionale tidsrækker	40
Identifikation	44
Estimation	61
Modelkontrol	75
4 VARMA-processer	87
4.1 Svagt stationære K -dimensionale tidsrækker	87
4.2 VAR-processer	89
4.3 Kovarianser og korrelationer	96
4.4 VMA-processer	101
4.5 VAR- og VMA-repræsentationer af VARMA-processen	103
4.6 Kovarianser for en VARMA(p, q)-proces	105
4.7 Forecasting	106
Forecasting i VARMA-processer	110
4.8 Strukturel analyse af VARMA-processer	111
Granger-kausaltet	111

Øjeblikkelig kausalitet	115
Impuls-response analyse	118
Opsplitning af variansen på forecast error	121
5 Empiriske tidsrækker II	124
5.1 Flerdimensionale tidsrækker	124
Identifikation	125
Krydskorrelationsmatricer	125
Partielle autoregressionsmatricer	126
Partielle krydskorrelationsmatricer	126
Partielle kanoniske korrelationer	127
Estimation	132
Modelkontrol	138
Strukturel analyse og forecasting	145
Appendikser	152
A Nyttige resultater	153
B Differensligninger	156
Opgaver	161
Indeks	164
Litteratur	171

Forord

Som titlen antyder, giver disse forelæsningsnoter en introduktion til teorien for og anvendelsen af tidsrækker. Notesættet, der er beregnet til brug i kurset *Statistiske Modeller 2*, koncentrerer sig om den af del af tidsrækkeanalysen, der vedrører ARMA-modellerne, som blev introduceret af [Box og Jenkins \(1970\)](#), og generalisationen af disse modeller til flerdimensionale tidsrækker, som beskrevet i for eksempel [Lütkepohl \(2007\)](#).

En stor del af materialet til noterne for de endimensionale tidsrækker stammer fra Andersen og Blæsild (2003), og omtalen af flerdimensionale tidsrækker tager udgangspunkt i specialet af Brix (2010), som er inspireret af Lütkepohl (2007).

Emnet *Tidsrækkeanalyse* er kun ét blandt flere i *Statistiske Modeller 2*. Tiden tillader derfor ikke en så detaljeret behandling af enmet som i referencerne ovenfor og heller ikke omtale af kointegration, som er et vigtigt begreb inden for flerdimensionale tidsrækker.

Den nærværende version af noterne er den anden version men langt fra den endelige. Den indeholder således ikke gengivelser af SAS-programmer. SAS-proceduren ARIMA er en gennemprøvet ældre procedure, der benyttes til analyse af endimensionale tidsrækker, mens beregningerne i flerdimensionale tidsrækker foretages ved hjælp af proceduren VARMAX, der er forholdsvis ny og tilsyneladende ikke har fundet sin form, idet den varierer fra version til version. Jeg har derfor valgt ikke at gengive SAS-programmerne her. Programmer i SAS og i R til fire af noternes eksempler findes på henholdsvis *EksemplerSAS.zip* og *EksemplerR.zip* i *Dokumenter/Noter 2013* på kursets hjemmeside. Data til og beskrivelser af programmerne findes også på disse filer.

Jeg vil gerne takke J  r  my Auneau-Cognacq for grundigt at have genneml  st tidligere versioner af notes  ttet. Jeg vil ogs   gerne takke Lotte W  rtz Brix for at stille SAS-programmer og T  X-filer fra hendes speciale til r  dighed for mig. Endelig tak til Lars Madsen for hans store og t  lmodige hj  lp med ops  tningen af notes  ttet.

Preben Blæsild, April 2013

Kapitel 1

Indledning

En tidsrække er en række observationer af et veldefineret stokastisk system foretaget sekventielt i tid til ækvidistante tidspunkter.

Metoder til analyse af tidsrækker er et vigtigt område inden for statistik. Eksempler på tidsrækker findes inden for en lang række områder, for eksempel økonomi (daglige priser på aktier, værdien af den samlede månedlige eksport, den kvartårlige inflationsrate, årlige opgørelser af et firmas udbytte, osv.), meteorologi og geofysik (regnmængder i successive dage, målinger af lufttemperatur, nedbør og vandstand hver time eller dagligt), marketing (salgstal i successive dage, måneder, kvartaler eller år), demografi (månedlige, kvartalsvise eller årlige målinger af populationers størrelse), procesovervågning (kontrolmålinger af en fabriksproduktion) og medicin (målinger hver time af blodtryk, temperatur, puls osv.).

Formålet med at analysere tidsrækker er foruden at modellere data ofte at prediktere fremtidige observationer af systemet eller at kontrollere systemet. Hvis observationerne er flerdimensionale er det ofte også af interesse at undersøge, hvordan de forskellige komponenter påvirker hinanden.

Vi betragter nu nogle eksempler på tidsrækker.

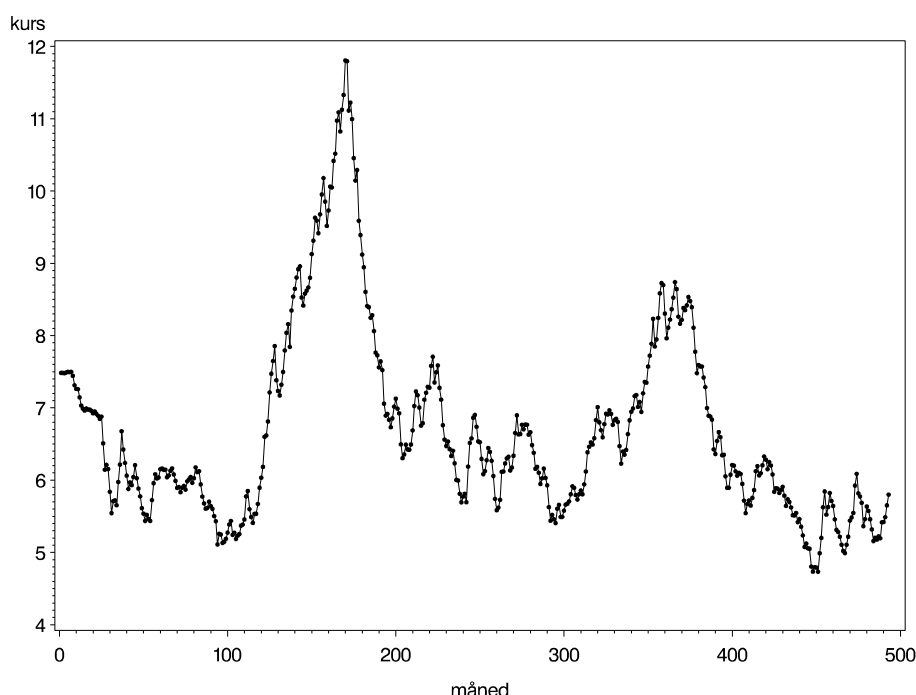
1.1 Eksempler

Eksempel 1.1. [Figur 1.1](#) viser valutakursen for US\$ fra januar 1971 til og med januar 2012, i alt 493 måneder. Kurserne er downloadet fra Federal Reserve Board i USA,¹ hvor man kan finde en lang række lignende datasæt.

Som det ses af [Figur 1.1](#) svinger kursen på US\$ meget i den betragtede periode med en største værdi på 11.81 kr. i februar 1985 og en mindste værdi på 4.73 kr. i juli 2008. Selv på baggrund af de meget svingende kurser er det af oplagt in-

¹<http://www.federalreserve.gov>

teresse at kunne prediktere kurserne fremover og udtale sig om præcisionen af disse forudsigelser.



Figur 1.1: Kursen på US\$ fra januar 1971 til og med januar 2012.

Eksempel 1.2. Klimaet har stor indflydelse på det økonomiske liv i Grønland, specielt på fiskeindustrien. Forekomsten af de forskellige fiskearter er meget afhængig af havets temperatur. Den følgende beskrivelse af et projekt, der har til opgave at undersøge mulige klimatiske forandringer i Grønland, er hentet fra [Risager \(1976\)](#):²

“U. S. Cold Region Research and Engineering Laboratory accomplished in 1966 a project aiming at producing an ice kernel by drilling through the ice cap by Camp Century 225 km east of Thule. By investigating the variation in the concentration of the O^{18} -isotope down through the ice kernel one gets an impression of the behaviour of the climate throughout the period in which the kernel was formed. [...]

On the ice kernel, which was produced by the project mentioned above, simultaneous measurements of the age of the ice and content of the O^{18} -isotope in the ice have been made. The isotope curve is determined from these measurements as the concentration ($\delta(O^{18})$) of the O^{18} -isotope as a function of the age of the ice. [...]

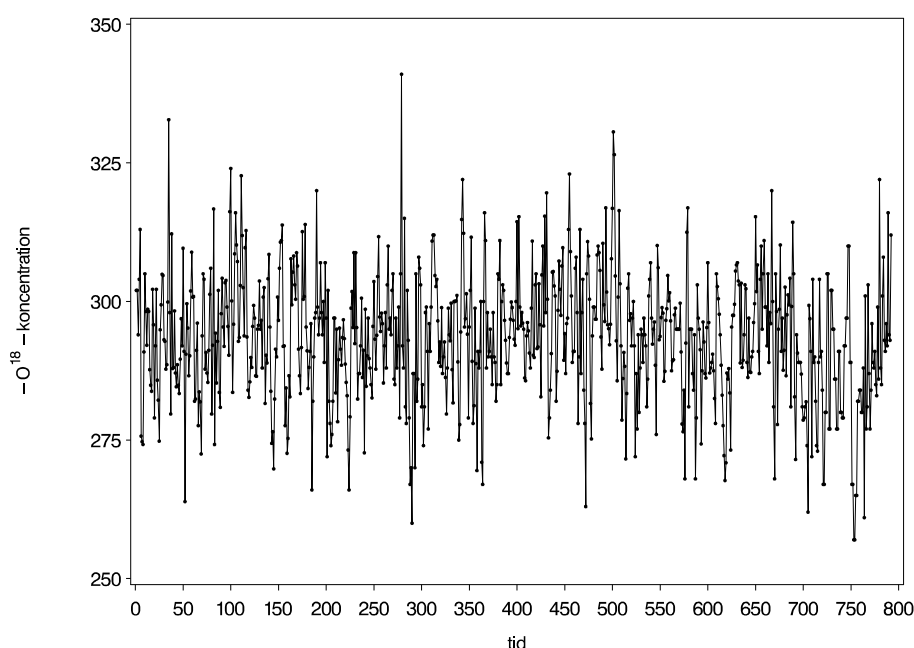
²Notationen ‘ O^{18} ’ er citeret fra artiklen, i dag ville man skrive ^{18}O om den omtalte ilt-isotop.

The reason for the interest in the isotope curve is that other investigations have shown the amount of O^{18} in polar snow depends nearly exclusively of the temperature. The lower the content of $\delta(O^{18})$ is in the ice the lower the temperature was when the ice was formed. [...]

In a period of 792 years – the concentration of O^{18} was measured. [...]

The data to be considered here, which we will refer to as the isotope curve, consists of the 792 measurements of $-\delta(O^{18})$, but for curious reasons – and somewhat confusing – the minus sign has been omitted on the observations. Thus a high value of the observation corresponds to a low value of $\delta(O^{18})$ indicating a low temperature. Similarly, a low value of the observation indicates a high temperature. The first observation corresponds to the year 1175.

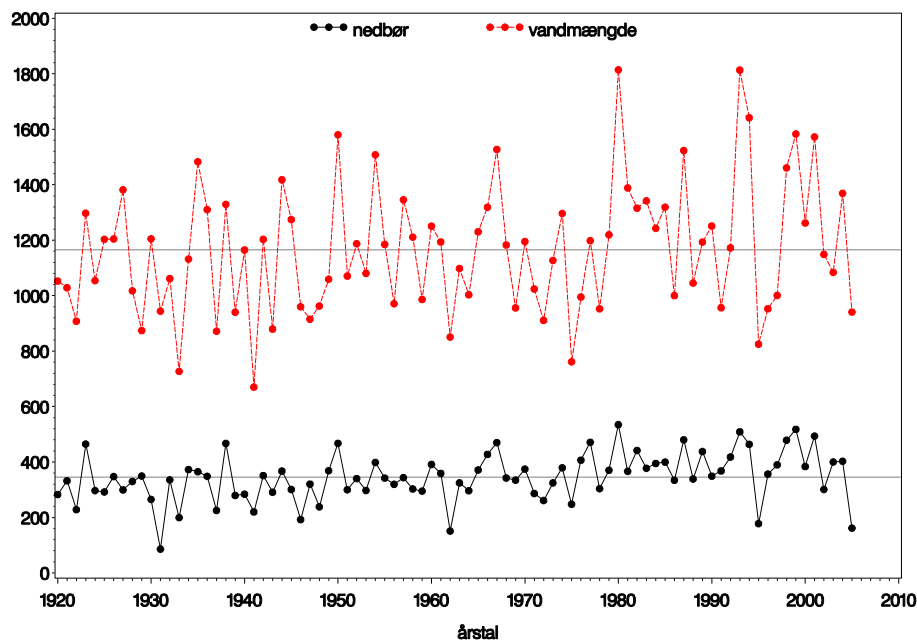
In Figure 1.2 the isotope curve corresponding to the 792 measurements is shown. [...] Unfortunately, in a period, approximately 1890–1930, only one determination of $\delta(O^{18})$ has been made for pieces of ice corresponding to two years. Consequently, the measurements from this period are pairwise alike. We must accept this as a incorrigible fault by the data; however, since the defective part of the data is only a small part of the data, we will at first allow ourself to ignore this defect.”



Figur 1.2: Camp Century isotopkurven.

Eksempel 1.3. Her betragtes nedbørs- og vandløbsmålinger i perioden 1920–2005, i alt 86 år. Vandløbsmålingerne er foretaget i Gudenåen ved Tvillumbro, mens nedbørsmålingerne stammer fra Grønbæk. Målinger er foretaget i vintermånederne, oktober, november, december, januar, februar og marts. Nedbørsmå-

lignerne består af den samlede nedbør (i mm) i de seks måneder. Vandløbsmålingerne er fremkommet ved at bestemme et døgnmiddel (i hl/sek) for hver af de seks måneder og derefter addere disse. Data er vist i [Figur 1.3](#).

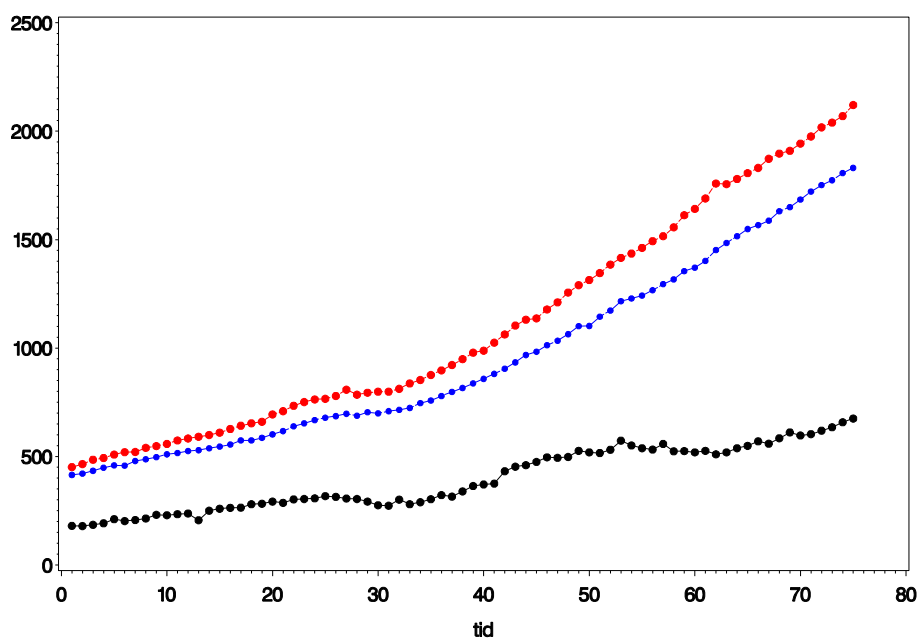


Figur 1.3: Målinger af nedbør og vandmængde i vintermånederne i årene 1920–2005.

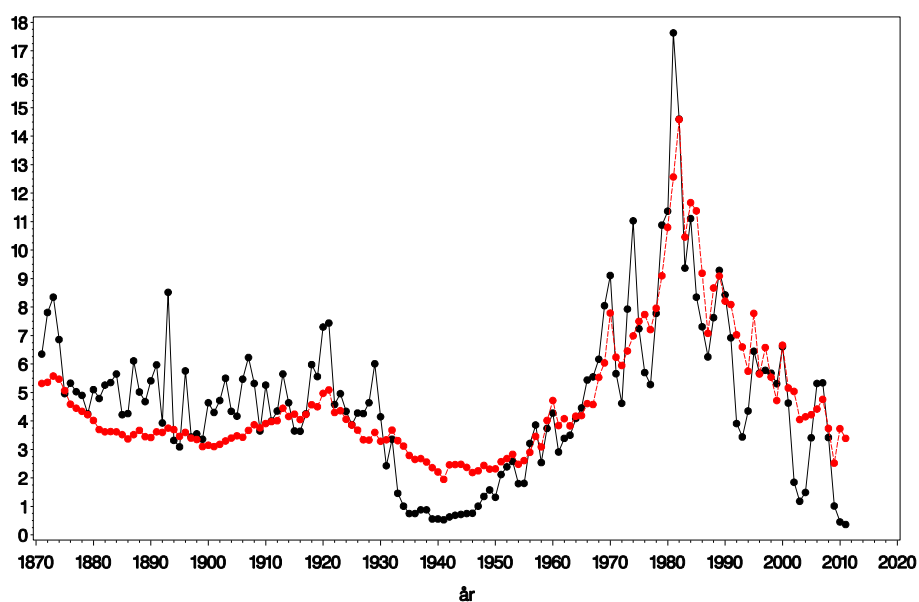
Foruden at betragte udviklingen af nedbør og vandmængde hver for sig er det af interesse at undersøge, om disse variable påvirker hinanden.

Eksempel 1.4. Data i dette eksempel er fra [Lütkepohl \(2007\)](#) og vedrører sæson-korrigerede kvartalsvise simultane observationer af faste investeringer, disponibel indkomst og udgifter til forbrug i Vesttyskland i perioden 1960–1978, i alt 76 observationer. Det er af interesse at beskrive udviklingen i disse tre variable og deres samspil.

Eksempel 1.5. [Figur 1.5](#) viser renten på 1-årige og på 10-årige amerikanske statsobligationer i perioden 1871–2011, i alt 141 år. Her er det af interesse at modellere rentens udvikling over tiden og et eventuelt samspil mellem renten på de to slags obligationer.



Figur 1.4: Faste investeringer (sort), disponibel indkomst (rød) og udgifter til forbrug (blå) i Vesttyskland kvartalsvis i perioden 1960–1978.



Figur 1.5: Renten på 1-årige (sort) og 10-årige (rød) amerikanske statsobligationer i årene 1871–2011.

Kapitel 2

Stokastiske processer

I dette afsnit introducerer vi begrebet *stokastiske processer*. Vi koncentrerer os mest om diskussion af specielle stokastiske processer, såsom strengt stationære processer, anden ordens processer og svagt stationære processer. I de teoretiske diskussioner betragter vi processer med komplekse værdier, mens processerne, der betragtes i anvendelserne, antager reelle værdier. Kovariansfunktionen for en svagt stationær proces studeres i detaljer og der omtales sætninger, der karakteriserer kovariansfunktionen for en svagt stationær proces. Begrebet et lineært filter introduceres. Endelig introduceres den partielle autokorrelationsfunktion for svagt stationære processer.

Lad (E, I, P) være et sandsynlighedsrum, (E', I') et målrum og T en vilkårlig mængde og antag, at der for ethvert $t \in T$ eksisterer en målelig afbildning X_t fra E til E' . Mængden $X = \{X_t : t \in T\}$ af stokastiske variable kaldes en *stokastisk proces* på (E, I) med *tilstandsrum* E' og med *indeksmængde* T .

Mængden T er ofte en delmængde af \mathbb{R} af formen $T = (a, b)$ eller $T = \{t_0, t_0 \pm h, t_0 \pm 2h, \dots\}$. I disse tilfælde omtales den indicerende variabel t næsten altid som *tiden* og processen siges at have henholdsvis *kontinuert* eller *diskret tid*. En proces med $T = \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$ eller $T = \mathbb{N}_0 = \{0, 1, 2, \dots\}$ omtales som en tidsrække.

Tilstandsrummet E' er typisk en delmængde af \mathbb{R}^k og klassifikationen af tilstandsrummet som værende en- eller flerdimensionalt er indlysende. Tilstandsrummet siges at være *diskret* henholdsvis *kontinuert*, hvis de stokastiske variable (vektorer) $X_t, t \in T$, er diskrete henholdsvis (absolut) kontinuerte.

Dette set-up er dog for generelt til at være brugbart, og for at opnå modeller for stokastiske processer, som vi kan håndtere matematisk, er det nødvendigt at gøre antagelser vedrørende $\{X_t\}_{t \in T}$ eller ækvivalent hermed om sandsynlighedsmålet P . Da vi i det følgende mest skal betragte stokastiske processer med endimensionalt og kontinuert udfaldsrum og med diskret tid, formuleres disse antagelser vedrørende P for en stokastisk proces med $E' = \mathbb{R}$ og med $T = \mathbb{Z}$. Det overlades til læseren at vurdere, hvorvidt antagelserne er direkte anvendelige i

andre tilfælde eller om modifikationer og/eller tilføjelser er påkrævet.

For at introducere de relevante antagelser er det nødvendigt med lidt mere notation. Lad (E, I, P) være et sandsynlighedsrum og lad $X = \{X_t : t \in \mathbb{Z}\}$ være en stokastisk proces på (E, I) med tilstandsrum \mathbb{R} . Med $\mathbb{R}^{\mathbb{Z}}$ betegner vi *mængden af afbildninger fra \mathbb{Z} ind i \mathbb{R}* , dvs.

$$\begin{aligned}\mathbb{R}^{\mathbb{Z}} &= \{\mathbf{x} = (\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots) : x_t \in \mathbb{R}, t \in \mathbb{Z}\} \\ &= \{\mathbf{x} = \{x_t\}_{t \in \mathbb{Z}} : x_t \in \mathbb{R}\}.\end{aligned}$$

Vi udstyrer nu $\mathbb{R}^{\mathbb{Z}}$ med en σ -algebra og definerer et sandsynlighedsmål på $\mathbb{R}^{\mathbb{Z}}$. Lad $\Delta_{\mathbb{Z}}$ være *mængden af alle endelige og ordnede delmængder af \mathbb{Z}* , dvs.

$$\Delta_{\mathbb{Z}} = \bigcup_{n=1}^{\infty} \{(t_1, t_2, \dots, t_n) \in \mathbb{Z}^n : t_1 < t_2 < \dots < t_n\}.$$

Lad \mathcal{B} betegne Borel delmængderne i \mathbb{R} . Hvis $\mathbf{t} = (t_1, t_2, \dots, t_n) \in \Delta_{\mathbb{Z}}$ og $B_{t_j} \in \mathcal{B}$, $j = 1, 2, \dots, n$, sætter vi

$$B_{\mathbf{t}} = B_{t_1} \times B_{t_2} \times \dots \times B_{t_n} = \{\mathbf{x} \in \mathbb{R}^{\mathbb{Z}} : x_{t_j} \in B_{t_j}, j = 1, 2, \dots, n\}.$$

Borel σ -algebraen $\mathcal{B}^{\mathbb{Z}}$ på $\mathbb{R}^{\mathbb{Z}}$ er σ -algebraen frembragt af disse mængder, dvs.

$$\mathcal{B}^{\mathbb{Z}} = \sigma\{B_{\mathbf{t}} : \mathbf{t} \in \Delta_{\mathbb{Z}}\}.$$

Da X_t er en stokastisk variabel for alle $t \in \mathbb{Z}$, ses det, at afbildningen

$$\begin{aligned}X : E &\longrightarrow \mathbb{R}^{\mathbb{Z}} \\ e &\longmapsto X(e) = \{X_t(e)\}_{t \in \mathbb{Z}}\end{aligned}$$

er $(I, \mathcal{B}^{\mathbb{Z}})$ -målelig. Værdien $X(e)$ kaldes *udfaldsfunktionen* svarende til e . Løftes sandsynlighedsmålet P på (E, I) ved hjælp af X , får vi et sandsynlighedsmål $P \cdot X$ på $(\mathbb{R}, \mathcal{B}^{\mathbb{Z}})$ givet ved

$$\begin{aligned}P \cdot X : \mathcal{B}^{\mathbb{Z}} &\longrightarrow [0, 1] \\ A &\longmapsto P \cdot X(A) = P(X^{-1}(A)) = P(\{e \in E : X(e) \in A\}).\end{aligned}$$

Specielt har vi for $\mathbf{t} \in \Delta_{\mathbb{Z}}$, at

$$P \cdot X(B_{t_1} \times \dots \times B_{t_n}) = P(\{e \in E : X_{t_1}(e) \in B_{t_1}, \dots, X_{t_n}(e) \in B_{t_n}\}).$$

Vi kan derfor for enhver endelig delmængde af ordnede tidspunkter $\mathbf{t} = (t_1, t_2, \dots, t_n) \in \Delta_{\mathbb{Z}}$ beregne den simultane fordelingsfunktion af de tilsvarende stokastiske variable $(X_{t_1}, \dots, X_{t_n})$ ved hjælp af $P \cdot X$.

For $\mathbf{t} = (t_1, \dots, t_n) \in \Delta_{\mathbb{Z}}$ lader vi $F_{\mathbf{t}}$ betegne den simultane fordelingsfunktion af $(X_{t_1}, \dots, X_{t_n})$, dvs.

$$F_{\mathbf{t}}(\mathbf{x}) = P\left(\bigcap_{i=1}^n \{e \in E : X_{t_i}(e) \leq x_i\}\right), \quad \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Familien $\{F_{\mathbf{t}}\}_{\mathbf{t} \in \Delta_{\mathbb{Z}}}$ kaldes *familien af endelig dimensionale fordelinger* for \mathbf{X} og opfylder:

- (i) $F_{\mathbf{t}}$ er en fordelingsfunktion på \mathbb{R}^n for alle $\mathbf{t} = (t_1, \dots, t_n) \in \Delta_{\mathbb{Z}}$
- (ii) $F_{(t_1, \dots, t_n)}(x_1, \dots, x_{k-1}, \infty, x_{k+1}, \dots, x_n) = F_{(t_1, \dots, t_{k-1}, t_{k+1}, \dots, t_n)}(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n)$ for alle $(t_1, \dots, t_n) \in \Delta_{\mathbb{Z}}$ og for alle $(x_1, \dots, x_n) \in \mathbb{R}^n$.

En familie af fordelingsfunktioner $\{F_{\mathbf{t}}\}_{\mathbf{t} \in \Delta_{\mathbb{Z}}}$, der opfylder betingelserne (i) og (ii), siges at være *konsistent*. Familien af endelig dimensionale fordelingsfunktioner for en stokastisk proces er således konsistent. Omvendt gælder der, som vist af [Kolmogorov \(1933\)](#):

Sætning 2.1. Hvis $\{F_{\mathbf{t}}\}_{\mathbf{t} \in \Delta_{\mathbb{Z}}}$ er en konsistent familie af fordelingsfunktioner, så eksisterer der et sandsynlighedsrum (E, I, P) og en stokastisk proces med tid \mathbb{Z} , $\mathbf{X} = \{X_t\}_{t \in \mathbb{Z}}$, således at $\{F_{\mathbf{t}}\}_{\mathbf{t} \in \Delta_{\mathbb{Z}}}$ er familien af endelig dimensionale fordelingsfunktioner for \mathbf{X} .

[Sætning 2.1](#) er bekvem, idet den tillader os at betragte $(\mathbb{R}^{\mathbb{Z}}, \mathcal{B}^{\mathbb{Z}})$ som det basale målrum i stedet for (E, I) . Desuden tillader den os at formulere antagelser vedrørende sandsynligheds målet på $(\mathbb{R}^{\mathbb{Z}}, \mathcal{B}^{\mathbb{Z}})$ ved hjælp af de endelig dimensionale fordelingsfunktioner, når blot vi gør det konsistent.

2.1 Stationære processer

En matematisk bekvem, men i praksis undertiden for restriktiv, antagelse om en stokastisk proces er, at den er strengt stationær, hvilket vil sige, at de endelig dimensionale fordelinger er invariante under translation af tiden.

Definition 2.2. En stokastisk proces $\{X_t\}_{t \in \mathbb{Z}}$ er **strengt stationær**, hvis det for alle $\mathbf{t} = (t_1, \dots, t_n) \in \Delta_{\mathbb{Z}}$ og for alle $k \in \mathbb{Z}$ gælder, at

$$F_{(t_1, \dots, t_n)} = F_{(t_1+k, \dots, t_n+k)}, \quad (2.1)$$

hvilket er ækvivalent med, at $(X_{t_1}, \dots, X_{t_n})$ har samme fordeling som $(X_{t_1+k}, \dots, X_{t_n+k})$.

Definition 2.3. En stokastisk proces $\{X_t\}_{t \in \mathbb{Z}}$ er en **anden ordens proces**, hvis

$$EX_t^2 < \infty, \quad \text{for alle } t \in \mathbb{Z}.$$

For en sådan proces kaldes funktionen

$$\begin{aligned} m : \mathbb{Z} &\longrightarrow \mathbb{R} \\ t &\longmapsto EX_t \end{aligned}$$

for **middelværdifunktionen** og funktionen

$$\begin{aligned} q : \mathbb{Z}^2 &\longrightarrow \mathbb{R} \\ (t, s) &\longmapsto \text{Cov}(X_t, X_s) = E\{(X_t - EX_t)(X_s - EX_s)\} \end{aligned}$$

omtales som **kovariansfunktionen**.

Sætning 2.4. Hvis $\{X_t\}_{t \in \mathbb{Z}}$ er en strengt stationær proces og en anden ordens proces, så er middelværdifunktionen m konstant, dvs.

$$m(t) = m, \quad t \in \mathbb{Z},$$

og kovariansfunktionen q afhænger af (t, s) kun gennem $t - s$.

Bevis. Det fås fra (2.1), at vi for alle $k \in \mathbb{Z}$ har

$$m(t) = m(t + k), \quad t \in \mathbb{Z}$$

og

$$q(t, s) = q(t + k, s + k), \quad (t, s) \in \mathbb{Z}^2,$$

hvilket viser sætningen. □

Definition 2.5. Hvis $\{X_t\}_{t \in \mathbb{Z}}$ er en anden ordens proces, så middelværdifunktionen m er konstant,

$$m(t) = m, \quad t \in \mathbb{Z},$$

og hvis kovariansfunktionen q afhænger af (t, s) kun gennem $t - s$,

$$q(t, s) = \gamma(t - s), \quad (t, s) \in \mathbb{Z}^2,$$

så kaldes $\{X_t\}_{t \in \mathbb{Z}}$ **en anden ordens stationær proces** eller **en svagt stationær proces** med **middelværdi** m og **kovariansfunktion** γ . Bemærk, at $\gamma(-k) = \gamma(k)$ for $k \in \mathbb{N}$.

Det ses, at en strengt stationær anden orden proces også er svagt stationær.

Særlig interesse knytter sig til normale eller Gaussiske processer, som defineres således:

Definition 2.6. En stokastisk proces $\{X_t\}_{t \in \mathbb{Z}}$ kaldes en **Gaussisk proces** eller en **normal proces**, hvis fordelingen af $(X_{t_1}, \dots, X_{t_n})$ er en n -dimensional normalfordeling for alle $\mathbf{t} = (t_1, \dots, t_n) \in \Delta_{\mathbb{Z}}$.

Bemærkning 2.7. En Gaussisk proces $\{X_t\}_{t \in \mathbb{Z}}$ er en anden ordens proces. Hvis tillige $\{X_t\}_{t \in \mathbb{Z}}$ er svagt stationær, ses det, at processen er strengt stationær, da den n -dimensionale normalfordeling er fuldstændigt karakteriseret af sin middelværdivektor og sin kovariansmatriks. ■

Den næste sætning giver en karakterisering af kovariansfunktionen for en svagt stationær proces.

Sætning 2.8. Kovariansfunktionen γ for en svagt stationær proces $\{X_t\}_{t \in \mathbb{Z}}$ er en ikke-negativ definit funktion, dvs.

$$\sum_{j=1}^n \sum_{k=1}^n \gamma(t_j - t_k) z_j z_k \geq 0 \quad (2.2)$$

for alle $\mathbf{t} = (t_1, \dots, t_n) \in \Delta_{\mathbb{Z}}$ og for alle $(z_1, \dots, z_n) \in \mathbb{R}^n$.

Omvendt eksisterer der for enhver ikke-negativ definit funktion γ en svagt stationær proces $\{X_t\}_{t \in \mathbb{Z}}$ med γ som kovariansfunktion.

Bevis. Lad γ være kovariansfunktionen for en svagt stationær proces med middelværdi 0. Betragt for $\mathbf{t} = (t_1, \dots, t_n) \in \Delta_{\mathbb{Z}}$ og $(z_1, \dots, z_n) \in \mathbb{R}^n$ den stokastiske variabel

$$U = \sum_{i=1}^n z_i X_{t_i}.$$

Da

$$\begin{aligned} 0 \leq \text{Var } U &= \text{Cov}(U, U) = \text{Cov}\left(\sum_{j=1}^n z_j X_{t_j}, \sum_{k=1}^n z_k X_{t_k}\right) \\ &= \sum_{j=1}^n \sum_{k=1}^n z_j z_k \text{Cov}(X_{t_j}, X_{t_k}) = \sum_{j=1}^n \sum_{k=1}^n \gamma(t_j - t_k) z_j z_k, \end{aligned}$$

ses det, at γ er en ikke-negativ definit funktion.

Hvis γ er en ikke-negativ definit funktion gælder for alle $\mathbf{t} \in \Delta_{\mathbb{Z}}$, at $n \times n$ matricen

$$\Sigma_{\mathbf{t}} = \{\gamma(t_i - t_j)\}_{i,j=1}^n$$

er en symmetrisk og ikke-negativ definit matris. Da familien af fordelingsfunktioner $\{F_{\mathbf{t}}\}_{\mathbf{t} \in \Delta_{\mathbb{Z}}}$, hvor

$$F_{\mathbf{t}} \sim N_n(\mathbf{0}, \Sigma_{\mathbf{t}}), \quad \mathbf{t} = (t_1, \dots, t_n) \in \Delta_{\mathbb{Z}},$$

er konsistent, så eksisterer der en svagt stationær Gaussisk proces med γ kovariansfunktion. \square

Bemærkning 2.9. Ifølge [Bemærkning 2.7](#) er den Gaussiske proces, der blev konstrueret i beviset for [Sætning 2.8](#), strengt stationær. \blacksquare

Kovariansfunktionen kan yderligere karakteriseres ved Herglotz sætning fra 1911, hvis bevis udelades.

Sætning 2.10 (Hergoltz). En funktion $\gamma : \mathbb{Z} \rightarrow \mathbb{R}$ er kovariansfunktion for en svagt stationær proces hvis og kun hvis der eksisterer en funktion $F : [-\pi, \pi] \rightarrow \mathbb{R}$, så

$$\gamma(n) = \int_{-\pi}^{\pi} e^{inx} dF(x), \quad n \in \mathbb{Z}, \quad (2.3)$$

og

- (a) F er ikke-aftagende
- (b) F er højrekontinuert
- (c) $F(-\pi) = 0$.

Desuden gælder der, at funktionen F er entydigt bestemt.

Definition 2.11. Hvis γ er kovariansfunktion for en svagt stationær proces $\{X_t\}_{t \in \mathbb{Z}}$, så kaldes funktionen F i [Sætning 2.10](#) processens **spektralfordeling** og målet dF processens **spektralmål**.

Hvis spektralmålet har tæthed f med hensyn til Lebesguemålet λ , dvs.

$$f(x) = \frac{dF}{d\lambda}(x),$$

så kaldes f processens **spektraltæthed**. (I så tilfælde ses det, at $\gamma(n)$ på nær konstanten 2π er den nte Fourier koefficient for f .)

Definition 2.12. Kovariansfunktionen for en svagt stationær proces $\{X_t\}_{t \in \mathbb{Z}}$ omtales ofte som **autokovariansfunktionen**. Funktionen

$$\begin{aligned} \rho : \mathbb{N} &\longrightarrow [-1, 1] \\ n &\longmapsto \gamma(n)/\gamma(0) \end{aligned} \quad (2.4)$$

kaldes processens **(auto)korrelationsfunktion** og refereres ofte (i engelsk litteratur) til blot som **processens ac**. Størrelsen $\gamma(0)$ omtales som processens **varians** og betegnes med σ_X^2 , dvs.

$$\sigma_X^2 = \text{Var } X_t, \quad t \in \mathbb{Z}.$$

Korrelationen mellem X_{t+n} og X_t er altså $\rho(n)$. Ofte skriver vi γ_n og ρ_n i stedet for henholdsvis $\gamma(n)$ and $\rho(n)$. (Undertiden defineres korrelationsfunktionen for en svagt stationær proces som funktionen i [\(2.4\)](#), men defineret på \mathbb{Z} i stedet for på \mathbb{N} .)

Bemærkning 2.13. Det ses af ovenstående, at en stationær Gaussisk proces er fuldstændigt karakteriseret ved dens middelværdi, varians og korrelationsfunktion. ■

I omtalen af svagt stationære processer behøver vi yderligere en korrelationsfunktion, nemlig den partielle autokorrelationsfunktion. For at definere denne funktion er det nødvendigt med lidt mere notation relateret til en svagt stationær proces $\{X_t\}_{t \in \mathbb{Z}}$. Med Γ_k og P_k betegnes henholdsvis kovariansmatricen og korrelationsmatricen for (X_1, \dots, X_k) , dvs.

$$\Gamma_k = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{k-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{k-1} & \gamma_{k-2} & \cdots & \gamma_0 \end{pmatrix}$$

og

$$P_k = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{pmatrix},$$

så

$$\Gamma_k = \gamma_0 P_k.$$

Endvidere, lader vi

$$\rho_k = (\rho_1, \dots, \rho_k) \quad \text{og} \quad \tilde{\rho}_k = (\rho_k, \dots, \rho_1)$$

og endelig, betegner ρ_k^* den transponerede af ρ_k . Med denne notation har vi:

Definition 2.14. Lad $\{X_t\}_{t \in \mathbb{Z}}$ være en svagt stationær proces, således at Γ_k eller, ækvivalent hermed P_k , er positiv definit for alle $k \in \mathbb{N}$ og lad ϕ_k betegne vektoren

$$\phi_k = (\phi_{k,1}, \dots, \phi_{k,k})$$

bestemt ved

$$\phi_k = \rho_k P_k^{-1}. \quad (2.5)$$

Funktionen

$$\begin{aligned} \phi_{-, -} : \mathbb{N} &\longrightarrow [-1, 1] \\ k &\longmapsto \phi_{k,k} \end{aligned} \quad (2.6)$$

kaldes **processens partielle autokorrelationsfunktion**, kort **processens pac**, og $\phi_{k,k}$ omtales som den **k 'te partielle autokorrelation**.

Som det ses af [Definition 2.14](#) er den partielle autokorrelationsfunktion en funktion af autokorrelationsfunktionen og den følgende sætning giver en rekursiv formel til beregning af $\phi_{k,k}$ ud fra $\{\rho_j\}_{j \leq k}$.

Sætning 2.15. *De partielle autokorrelationer kan beregnes ud fra autokorrelationerne ved hjælp af følgende rekursionsformel:*

$$\phi_{1,1} = \rho_1,$$

og for $k = 1, 2, \dots$ er

$$\phi_{k+1,k+1} = \frac{\rho_{k+1} - \sum_{j=1}^k \phi_{k,j} \rho_{k+1-j}}{1 - \sum_{j=1}^k \phi_{k,j} \rho_j},$$

hvor

$$\phi_{k+1,j} = \phi_{k,j} - \phi_{k+1,k+1} \phi_{k,k+1-j}, \quad j = 1, \dots, k.$$

Bevis. For $k = 1, 2, \dots$ lader vi

$$\begin{aligned} \phi_k &= (\phi_{k,1}, \dots, \phi_{k,k}), & \tilde{\phi}_k &= (\phi_{k,k}, \dots, \phi_{k,1}), \\ \phi_{k,k-1} &= (\phi_{k,1}, \dots, \phi_{k,k-1}), & \tilde{\phi}_{k,k-1} &= (\phi_{k,k-1}, \dots, \phi_{k,1}). \end{aligned}$$

Formel (2.5) er ækvivalent med

$$P_k \phi_k^* = \rho_k^*.$$

Erstatter vi k med $k+1$ og bruger den netop introducerede notation bliver denne formel til

$$\begin{pmatrix} P_k & \tilde{\rho}_k^* \\ \tilde{\rho}_k & 1 \end{pmatrix} \begin{pmatrix} \phi_{k+1,k}^* \\ \phi_{k+1,k+1}^* \end{pmatrix} = \begin{pmatrix} \rho_k^* \\ \rho_{k+1}^* \end{pmatrix}$$

eller, ækvivalent hermed,

$$\begin{aligned} P_k \phi_{k+1,k}^* + \tilde{\rho}_k^* \phi_{k+1,k+1}^* &= \rho_k^* \\ \tilde{\rho}_k \phi_{k+1,k}^* + \phi_{k+1,k+1}^* &= \rho_{k+1}^*. \end{aligned} \tag{2.7}$$

Multipliserer vi den første ligning i (2.7) med P_k^{-1} fås

$$\phi_{k+1,k}^* = P_k^{-1}(\rho_k^* - \tilde{\rho}_k^* \phi_{k+1,k+1}^*), \tag{2.8}$$

og indsættes (2.8) i den anden ligning i (2.7), finder vi

$$\phi_{k+1,k+1}^* = \rho_{k+1}^* - \tilde{\rho}_k P_k^{-1}(\rho_k^* - \tilde{\rho}_k^* \phi_{k+1,k+1}^*). \tag{2.9}$$

Formel (2.5) medfører, at $P_k^{-1} \rho_k^* = \phi_k^*$ og $P_k^{-1} \tilde{\rho}_k^* = \tilde{\phi}_k^*$, og formlerne (2.8) og (2.9) bliver til

$$\phi_{k+1,k}^* = \phi_k^* - \tilde{\phi}_k^* \phi_{k+1,k+1}^*$$

og

$$\phi_{k+1,k+1}^* (1 - \tilde{\rho}_k \tilde{\phi}_k^*) = \rho_{k+1}^* - \tilde{\rho}_k \phi_k^*,$$

og det er nu let at gøre beviset færdigt. \square

Hvis processen $\{X_t\}_{t \in \mathbb{Z}}$ er Gaussisk, kan det vises, at den k 'te partielle autocorrelation $\phi_{k,k}$ er den partielle korrelation mellem X_1 og X_{k+1} inden for mængden af variable $X_1, X_2, \dots, X_k, X_{k+1}$, dvs. korrelationskoefficienten

$$\text{Cor}((X_1, X_{k+1})|(X_2, \dots, X_k))$$

mellem X_1 og X_{k+1} beregnet i den betingede fordeling af disse to variable givet de mellemliggende (X_2, \dots, X_k) .

2.2 ARMA-processer

I dette afsnit defineres en klasse af svagt stationære processer kaldet ARMA-processerne. I [Box og Jenkins \(1970\)](#) studeres disse processer meget detaljeret, og i denne bog samt i en mængde senere artikler af samme og andre forfattere findes en lang række eksempler på anvendelser af ARMA-processer som modeller for fænomener, der udvikler sig i tiden. Pålægges ARMA-processerne visse betingelser, kan de karakteriseres fuldstændigt ved middelværdien og kovariansfunktionen. Disse betingelser er relevante fra et statistisk synspunkt, og desuden bekvemme på grund af det netop nævnte. Da sætningen vedrørende eksistensen af ARMA-processer, [Sætning 2.19](#) på [side 19](#), ikke bevises her, illustreres de betingelser, der pålægges ARMA-processerne. Endvidere studeres kovariansfunktionen og den partielle kovariansfunktion for en generel ARMA-proces og for visse specialtilfælde.

Klassen af ARMA-processer kan på simpel vis udvides til at omfatte visse stokastiske processer, der ikke er svagt stationære. Disse processer vil vi kort omtale senere og her blot bemærke, at den udvidede klasse af processer kaldes ARIMA-processerne.

Definition 2.16. Lad $\{W_t : t \in \mathbb{Z}\}$ være en følge af ukorrelerede og identisk fordelte stokastiske variable med middelværdi 0 og varians σ_W^2 . Den stokastiske proces $\mathbf{W} = \{W_t\}_{t \in \mathbb{Z}}$ kaldes en **hvid støj med varians σ_W^2** .

Da

$$\begin{aligned} EW_t^2 &= \sigma_W^2 < \infty, & t \in \mathbb{Z}, \\ EW_t &= 0, & t \in \mathbb{Z}, \end{aligned}$$

og

$$\text{Cov}(W_t, W_s) = E(W_t W_s) = \begin{cases} 0 & \text{hvis } t \neq s, (t, s) \in \mathbb{Z}^2 \\ \sigma_W^2 & \text{hvis } t = s, (t, s) \in \mathbb{Z}^2, \end{cases} \quad (2.10)$$

er en hvid støj \mathbf{W} en svagt stationær proces. Det ses let, at spektraltætheden for \mathbf{W} er

$$f_W(x) = \frac{\sigma_W^2}{2\pi}, \quad x \in [-\pi, \pi]. \quad (2.11)$$

De stokastiske variable X_t i de processer $\mathbf{X} = \{X_t\}_{t \in \mathbb{Z}}$, vi kommer til at betragte, kan ofte udtrykkes som en linear kombination af de stokastiske variable i en hvid støj $\mathbf{W} = \{W_t\}_{t \in \mathbb{Z}}$; linear kombinationer hvori der kan indgå uendelig mange af komponenterne af \mathbf{W} .

Mere præcist, antag at $\mathbf{W} = \{W_t\}_{t \in \mathbb{Z}}$ er en hvid støj på et sandsynlighedsrum (E, I, P) , og lad $\{\psi_k\}_{k \in \mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}$. For alle $t \in \mathbb{Z}$ og for ethvert $K \in \mathbb{N}$, har vi ved hjælp af (2.10), at

$$\begin{aligned} E\left(\sum_{k=-K}^K \psi_k W_{t-k}\right)^2 &= E\left(\sum_{k=-K}^K \psi_k W_{t-k} \sum_{j=-K}^K \psi_j W_{t-j}\right) \\ &= \sum_{k=-K}^K \sum_{j=-K}^K \psi_k \psi_j E(W_{t-k} W_{t-j}) \\ &= \sigma_W^2 \sum_{k=-K}^K \psi_k^2. \end{aligned}$$

Det vil sige, at $\sum_{k=-K}^K \psi_k W_{t-k}$ er konvergent i $\mathcal{L}^2(E, I, P)$ for $K \rightarrow \infty$ hvis og kun hvis $\sum_{k=-\infty}^{\infty} \psi_k^2 < \infty$. Antag, at dette er tilfældet og lad X_t betegne grænseværdien, dvs.

$$X_t = \sum_{k=-\infty}^{\infty} \psi_k W_{t-k} \quad \left(= \lim_{K \rightarrow \infty} \sum_{k=-K}^K \psi_k W_{t-k} \right).$$

Da

$$EX_t = \sum_{k=-\infty}^{\infty} \psi_k EW_{t-k} = 0$$

og

$$\text{Cov}(X_t, X_s) = \sigma_W^2 \sum_{k=-\infty}^{\infty} \psi_k \psi_{k+t-s},$$

er processen $\mathbf{X} = \{X_t\}_{t \in \mathbb{Z}}$ en svagt stationær proces.

Disse overvejelser begrundes:

Definition 2.17. Lad $\mathbf{W} = \{W_t\}_{t \in \mathbb{Z}}$ være en hvid støj på (E, I, P) og lad $\{\psi_k\}_{k \in \mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}$ således at $\sum_{k=-\infty}^{\infty} |\psi_k|^2 < \infty$. Den svagt stationære proces $\mathbf{X} = \{X_t + \mu\}_{t \in \mathbb{Z}}$ givet ved

$$X_t = \sum_{k=-\infty}^{\infty} \psi_k W_{t-k}, \quad t \in \mathbb{Z},$$

hvor summen er konvergent i $\mathcal{L}^2(E, I, P)$, kaldes **et glidende gennemsnit** af W .

Hvis $\psi_k = 0$ for $k < 0$ og $\psi_0 \neq 0$ kaldes processen X **et baglæns glidende gennemsnit**.

Hvis $\mu \in \mathbb{R}$, $\psi_0 = 1$, $\psi_j = -\theta_j$, $j = 1, \dots, q$, $\theta_q \neq 0$ og $\psi_k = 0$ for $k < 0$ og $k > q$, kaldes processen $\{X_t + \mu\}_{t \in \mathbb{Z}}$ **et baglæns glidende gennemsnit af orden q** eller en **MA(q)-proces** med parametre $\theta_1, \dots, \theta_q, \sigma_W^2$ og μ . (MA \approx moving average) I dette tilfælde er

$$X_t = W_t - \theta_1 W_{t-1} - \dots - \theta_q W_{t-q}, \quad t \in \mathbb{Z}.$$

Hvis X er et glidende gennemsnit af den hvide støj W , kan X_t altså udtrykkes som en (dobbelt uendelig, uendelig eller endelig) linear kombination af $\{W_s\}_{s \in \mathbb{Z}}$. Hvis indekset t betragtes som tiden, ses det, at hvis X er et glidende gennemsnit af W , så afhænger værdien af X til tid t , X_t , af såvel fortiden, $\{W_s : s < t\}$, nutiden, W_t , som fremtiden, $\{W_s : s > t\}$. Fra et statistisk synspunkt er modeller, hvor en observation til tid t kun afhænger af fortiden og nutiden, de eneste relevante, og følgelig er et glidende gennemsnit kun relevant som en statistisk model, hvis det er et baglæns glidende gennemsnit.

Også processer X , hvor elementer W_t af en hvid støj $W = \{W_t\}_{t \in \mathbb{Z}}$ kan udtrykkes som en linear kombination af elementerne i X op til og med t , $\{X_s : s \leq t\}$, er af interesse.

Definition 2.18. Lad $\{W_t\}_{t \in \mathbb{Z}}$ være en hvid støj på (E, I, P) . En svagt stationær proces $X = \{X_t + \mu\}_{t \in \mathbb{Z}}$, hvor $\{X_t\}_{t \in \mathbb{Z}}$ er en løsning til ligningerne

$$W_t = \sum_{k=0}^{\infty} \pi_k X_{t-k}, \quad t \in \mathbb{Z}, \quad (2.12)$$

hvor $\pi_k \in \mathbb{R}$, $k \geq 0$, $\pi_0 \neq 0$ og hvor summen er konvergent i $\mathcal{L}^2(E, I, P)$, kaldes en (baglæns) **autoregression** med W som hvid støj.

Hvis $\mu \in \mathbb{R}$, $\pi_0 = 1$, $\pi_j = -\varphi_j$, $j = 1, \dots, p$, $\varphi_p \neq 0$ og $\pi_k = 0$ for $k > p$, kaldes processen $\{X_t + \mu\}_{t \in \mathbb{Z}}$ **en autoregression af orden p** eller en **AR(p)-proces** med parametre $\varphi_1, \dots, \varphi_p$, σ_W^2 and μ . (AR \approx autoregression) I dette tilfælde er

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + W_t. \quad (2.13)$$

Hvis elementerne $\{W_t\}_{t \in \mathbb{Z}}$ i den hvide støj W er uafhængige, fås det af (2.13), at

$$E(X_t | (X_{t-1}, \dots, X_{t-p}) = (x_{t-1}, \dots, x_{t-p})) = \varphi_1 x_{t-1} + \dots + \varphi_p x_{t-p}$$

og

$$\text{Var}(X_t | (X_{t-1}, \dots, X_{t-p}) = (x_{t-1}, \dots, x_{t-p})) = \sigma_W^2,$$

dvs. den betingede model for X_t givet de foregående p værdier X_{t-1}, \dots, X_{t-p} af processen er en regressionsmodel. De kovariate i denne regressionsmodel ændres fra tidspunkt til tidspunkt, hvilket retfærdiggør navnet “autoregression”.

Kravet, at løsningen X til ligningerne i (2.12) i Definition 2.18 er en svagt stationær proces, medfører, at koefficienterne $\{\varphi_k\}$ må opfylde visse betingelser. Disse er formuleret i Sætning 2.19 nedenfor.

For at illustrere dette betragter vi en autoregression af order 1, givet ved

$$W_t = X_t - \varphi X_{t-1}, \quad t \in \mathbb{Z}. \quad (2.14)$$

Lad dF_X og γ_X betegne henholdsvis spektralmålet og autokovariansfunktion for X , dvs.

$$\gamma_X(k) = \int_{-\pi}^{\pi} e^{ikx} dF_X(x), \quad k \in \mathbb{Z},$$

og lad tilsvarende dF_W og γ_W betegne henholdsvis spektralmålet og autokovariansfunktionen for W . For $k \in \mathbb{Z}$ fås – ved hjælp af (2.14) – at

$$\begin{aligned} \int_{-\pi}^{\pi} e^{ikx} \frac{\sigma_W^2}{2\pi} dx &= \gamma_W(k) \\ &= \text{Cov}(W_{t+k}, W_t) \\ &= \text{Cov}(X_{t+k} - \varphi X_{t+k-1}, X_t - \varphi X_{t-1}) \\ &= \gamma_X(k) - \varphi \gamma_X(k-1) - \varphi \gamma_X(k+1) + \varphi^2 \gamma_X(k) \\ &= \int_{-\pi}^{\pi} (1 - \varphi e^{-ix} - \varphi e^{ix} + \varphi^2) e^{ikx} dF_X(x) \\ &= \int_{-\pi}^{\pi} |1 - \varphi e^{-ix}|^2 e^{ikx} dF_X(x). \end{aligned}$$

Da spektralmålet er entydigt bestemt, følger det, at

$$|1 - \varphi e^{-ix}|^2 dF_X(x) = \frac{\sigma_W^2}{2\pi} dx$$

eller, ækvivalent hermed, at spektraltætheden for X er

$$f_X(x) = \frac{dF_X}{d\lambda}(x) = \frac{\sigma_W^2}{2\pi} |1 - \varphi e^{-ix}|^{-2}, \quad \text{hvis } e^{-ix} \neq \varphi^{-1},$$

hvor λ betegner Lebesguemålet. Da

$$\infty > \gamma_X(0) = \int_{-\pi}^{\pi} f_X(x) dx \geq \frac{\sigma_W^2}{2\pi} \int_{-\pi}^{\pi} |1 - \varphi e^{-ix}|^{-2} dx,$$

følger det, at kravet om stationaritet af X medfører, at $|\varphi| \neq 1$.

Om de stationære processer, vi vil betragte, vil vi kræve, at de kan repræsenteres som såvel autoregressioner som baglæns glidende gennemsnit. Dette begrundes umiddelbart efter Definition 2.20. Lad os for at illustrere dette krav igen

betragte AR(1)-processen givet ved (2.14), hvor vi nu antager, at $|\varphi| \neq 1$ (for at sikre svag stationaritet).

Betegner B den operator, der anvendt på X_t giver X_{t-1} , dvs.

$$BX_t = X_{t-1},$$

kan (2.14) skrives

$$W_t = (1 - \varphi B)X_t, \quad t \in \mathbb{Z}. \quad (2.15)$$

Af Definition 2.17 ses at et glidende gennemsnit kan skrives som

$$X_t = \psi(B)W_t, \quad t \in \mathbb{Z}, \quad (2.16)$$

hvor $\psi(B)$ er operatoren

$$\psi(B) = \sum_{k=-\infty}^{\infty} \psi_k B^k.$$

Regner vi formelt ses, at autoregressionen i (2.15) kan repræsenteres som et glidende gennemsnit hvis og kun hvis

$$\psi(B) = (1 - \varphi B)^{-1}.$$

Hvis $|\varphi| < 1$, kan vi formelt rækkeudvikle $(1 - \varphi B)^{-1}$, idet $|\varphi B| = |\varphi||B| = |\varphi| < 1$. Vi får

$$\psi(B) = (1 - \varphi B)^{-1} = \sum_{k=0}^{\infty} (\varphi B)^k$$

og (2.16) medfører, at

$$X_t = \sum_{k=0}^{\infty} \varphi^k W_{t-k}, \quad t \in \mathbb{Z},$$

dvs. X er et baglæns glidende gennemsnit.

På samme måde får vi for $|\varphi| > 1$, at

$$\begin{aligned} \psi(B) &= (1 - \varphi B)^{-1} \\ &= -\varphi^{-1} B^{-1} (1 - \varphi^{-1} B^{-1})^{-1} \\ &= -\varphi^{-1} B^{-1} \sum_{k=0}^{\infty} (\varphi^{-1} B^{-1})^k \\ &= -\sum_{k=1}^{\infty} (\varphi^{-1} B^{-1})^k \end{aligned}$$

og (2.16) medfører, at

$$X_t = -\sum_{k=1}^{\infty} \varphi^{-k} W_{t+k}, \quad t \in \mathbb{Z},$$

dvs. X er et glidende gennemsnit, sådan at observationen X_t til tid t udtrykkes ved hjælp af de fremtidige værdier $\{W_s\}_{s>t}$ af W .

Disse beregninger viser, at autoregression (2.15) kan repræsenteres som et baglæns glidende gennemsnit hvis og kun hvis $|\varphi| < 1$, hvilket er ækvivalent med at polynomiet $\varphi(B) = 1 - \varphi B$, $B \in \mathbb{C}$, har rødder uden for enhedscirklen. **Sætning 2.19** retfærdiggør de formelle beregninger og giver blandt andet også betingelser for, hvornår en $MA(q)$ -proces – som altid er svagt stationær – kan repræsenteres som en autoregression.

Sætning 2.19. *Lad $(p, q) \in \mathbb{N}_0^2$ og lad $\varphi_1, \dots, \varphi_p$ og $\theta_1, \dots, \theta_q$ være reelle tal, så $\varphi_p \neq 0$ og $\theta_q \neq 0$. Sæt*

$$\varphi(z) = 1 - \varphi_1 z - \dots - \varphi_p z^p, \quad z \in \mathbb{C},$$

og

$$\theta(z) = 1 - \theta_1 z - \dots - \theta_q z^q, \quad z \in \mathbb{C}.$$

Lad $W = \{W_t\}_{t \in \mathbb{Z}}$ være en hvid støj med varians $\sigma_W^2 > 0$ og betragt systemet af ligninger:

$$X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p} = W_t - \theta_1 W_{t-1} - \dots - \theta_q W_{t-q}, \quad t \in \mathbb{Z}. \quad (2.17)$$

Da gælder:

- (a) Ligningssystemet (2.17) har en svagt stationær proces $X = \{X_t\}_{t \in \mathbb{Z}}$ (med midelværdi 0) som løsning hvis og kun hvis funktionen

$$\mathbb{C} \ni z \mapsto \frac{\theta(z)}{\varphi(z)} \in \mathbb{C}$$

er begrænset på enhedscirklen $\{z \in \mathbb{C} : |z| = 1\}$, dvs. at rødder for $\varphi(z)$ på enhedscirklen også er rødder for $\theta(z)$ med mindst samme multiplicitet. Desuden er løsningen X entydigt bestemt hvis og kun hvis $\varphi(z) \neq 0$ for $|z| = 1$.

Vi antager derfor, at $\varphi(z) \neq 0$ for $|z| = 1$.

- (b) Hvis $\varphi(z)$ har alle rødder uden for enhedscirklen, dvs. hvis $\varphi(z) \neq 0$ for $|z| \leq 1$, kan løsningen X repræsenteres som et baglæns glidende gennemsnit

$$X_t = \sum_{k=0}^{\infty} \psi_k W_{t-k}, \quad t \in \mathbb{Z}, \quad (2.18)$$

hvor $\{\psi_k\}_{k \in \mathbb{N}_0}$ er koefficienterne i potensrækkeudviklingen af $\theta(z)/\varphi(z)$ for $|z| \leq 1$, dvs.

$$\frac{\theta(z)}{\varphi(z)} = \sum_{k=0}^{\infty} \psi_k z^k, \quad |z| \leq 1.$$

Vi antager derfor, at $\varphi(z) \neq 0$ for $|z| \leq 1$.

- (c) Hvis $\theta(z)$ har alle rødder uden for enhedscirklen, dvs. hvis $\theta(z) \neq 0$ for $|z| \leq 1$, kan løsningen \mathbf{X} repræsenteres som en (baglæns) autoregression

$$W_t = \sum_{k=0}^{\infty} \pi_k X_{t-k}, \quad t \in \mathbb{Z}, \quad (2.19)$$

hvor $\{\pi_k\}_{k \in \mathbb{N}_0}$ er koefficienterne i potensrækkeudviklingen af $\varphi(z)/\theta(z)$ for $|z| \leq 1$, dvs.

$$\frac{\varphi(z)}{\theta(z)} = \sum_{k=0}^{\infty} \pi_k z^k, \quad |z| \leq 1.$$

Bevis. Udelades her. Sætningen er vist i [Andersen og Blæsild \(2003\)](#). \square

Definition 2.20. Lad $\mathbf{X} = \{X_t\}_{t \in \mathbb{Z}}$ være en svagt stationær løsning til ligningerne i formel (2.17). For $\mu \in \mathbb{R}$ kaldes processen $\{X_t + \mu\}_{t \in \mathbb{Z}}$ en **ARMA-proces af orden (p, q)** med hvid støj \mathbf{W} og med parametre $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q, \sigma_W^2$ og μ og den betegnes med $\text{ARMA}(p, q)$. Undertiden skriver vi $\mathbf{X} \sim \text{ARMA}(p, q)$ for at antyde, at \mathbf{X} er en $\text{ARMA}(p, q)$ -proces.

Bemærk, at $\text{ARMA}(p, 0) = \text{AR}(p)$, $\text{ARMA}(0, q) = \text{MA}(q)$ og at en $\text{ARMA}(0, 0)$ -proces er en hvid støj.

Lad \mathbf{X} betegne $\text{ARMA}(p, q)$ -processen

$$\varphi(B)X_t = \theta(B)W_t, \quad t \in \mathbb{Z}, \quad (2.20)$$

og lad $\tau \in \mathbb{R}$ og antag, at $|\tau| \neq 1$. Så løser \mathbf{X} også ligningerne

$$(1 - \tau B)\varphi(B)X_t = (1 - \tau B)\theta(B)W_t, \quad t \in \mathbb{Z},$$

dvs. at \mathbf{X} også er en $\text{ARMA}(p+1, q+1)$ -proces. Det ses, at hvis \mathbf{X} er en $\text{ARMA}(p, q)$ -proces og polynomierne $\varphi(z)$ og $\theta(z)$ har en fælles rod, så er \mathbf{X} også en $\text{ARMA}(p-1, q-1)$ -proces.

I det følgende antager vi, at $\varphi(z)$ og $\theta(z)$ ikke har fælles rødder.

Det ses fra [Sætning 2.19](#), at hvis $\varphi(z) \neq 0$ og $\theta(z) \neq 0$ for $|z| \leq 1$, så kan en løsning \mathbf{X} til ligningerne (2.17) repræsenteres både som et baglæns glidende gennemsnit og som en (baglæns) autoregression. Umiddelbart forekommer kravet om, at for eksempel en $\text{AR}(p)$ -proces skal kunne repræsenteres som et baglæns glidende gennemsnit ikke at være en nødvendighed fra et statistisk synspunkt. Det kan imidlertid vises, at en $\text{AR}(p)$ -proces altid kan repræsenteres som et glidende gennemsnit, samt at dette er et baglæns glidende gennemsnit hvis og kun

hvis $\varphi(z) \neq 0$ for $|z| \leq 1$. Kravet er derfor nødvendigt, hvis vi stiller den rimelige betingelse, at observationen X_t til tid t kun må afhænge af fortid og nutid, uanset hvilken repræsentation vi betragter.

Lad \mathcal{A} betegne mængden af ARMA-processer, dvs.

$$\mathcal{A} = \{X : X \sim \text{ARMA}(p, q), (p, q) \in \mathbb{N}_0^2\},$$

og lad \mathcal{A}_0 betegne mængden af ARMA-processer, der kan repræsenteres som både et baglæns glidende gennemsnit og en (baglæns) autoregression, dvs.

$$\mathcal{A}_0 = \{X \in \mathcal{A} : \varphi(z) \neq 0 \text{ og } \theta(z) \neq 0 \text{ for } |z| \leq 1\}.$$

Hvis χ betegner afbildningen

$$\chi(X) = (\mu_X, \gamma_X), \quad X \in \mathcal{A},$$

hvor μ_X og γ_X betegner henholdsvis middelværdien og autokovariansfunktionen for X , så kan det vises, at restriktionen af χ til \mathcal{A}_0 er en bijektion af \mathcal{A}_0 på $\chi(\mathcal{A})$, dvs. der er en 1-1 sammenhæng mellem mængderne \mathcal{A}_0 og $\chi(\mathcal{A})$. Dette resultat viser altså, at processerne i mængden af ARMA-processer \mathcal{A}_0 kan karakteriseres ved deres middelværdi og autokovariansfunktion, og, desuden, at der til enhver proces i \mathcal{A} findes en proces i \mathcal{A}_0 med den samme middelværdi og autokovariansfunktion.

Som en sidste kommentar til [Sætning 2.19](#) bemærkes, at hvis en proces X givet ved (2.20) tilhører delmængden \mathcal{A}_0 og hvis $p \geq 1$ og $q \geq 1$, så er repræsentationerne af X som henholdsvis et baglæns glidende gennemsnit og en autoregression begge uendelige. Derfor foretrækkes repræsentation af X i (2.20), da den kun indeholder endeligt mange parametre.

Den følgende sætning indeholder resultater vedrørende autokovariansfunktionen, autokorrelationsfunktionen og spektraltætheden for en ARMA-proces. Uden tab af generalitet antager vi, at middelværdien for processen er 0.

Sætning 2.21. *Lad X være en ARMA(p, q)-proces*

$$\varphi(B)X_t = \theta(B)W_t, \quad t \in \mathbb{Z},$$

med $\mu = 0$ og således at $\varphi(z) \neq 0$ og $\theta(z) \neq 0$ for $|z| \leq 1$.

Da gælder der:

(a) Autokovariansfunktionen er af formen

$$\gamma_k = \sigma_W^2 \sum_{i=0}^{\infty} \psi_{i+k} \psi_i, \quad k \in \mathbb{N}_0,$$

hvor $\{\psi_k\}_{k \in \mathbb{N}_0}$ er koefficienterne i potensrækkeudviklingen af $\theta(z)/\varphi(z)$.

$$(b) \quad \gamma_k - \varphi_1 \gamma_{k-1} - \cdots - \varphi_p \gamma_{k-p} = 0, \quad k > q, \quad (2.21)$$

eller ækvivalent hermed

$$\varphi(B)\gamma_k = 0, \quad k > q,$$

hvor $B\gamma_k = \gamma_{k-1}$. Autokovarianserne $\{\gamma_k\}_{k \in \mathbb{Z}}$ er altså for $k > q$ løsningen til en differensligning af orden p med differensoperator $\varphi(B)$.

(c) Autokorrelationerne $\rho_k = \gamma_k/\gamma_0$, $k \in \mathbb{Z}$, opfylder ligningerne

$$\rho_k - \varphi_1 \rho_{k-1} - \cdots - \varphi_p \rho_{k-p} = 0, \quad k > q, \quad (2.22)$$

eller ækvivalent hermed

$$\varphi(B)\rho_k = 0, \quad k > q,$$

hvor $B\rho_k = \rho_{k-1}$. Autokorrelationerne $\{\rho_k\}_{k \in \mathbb{Z}}$ er altså for $k > q$ løsningen til en differensligning af orden p med differensoperator $\varphi(B)$.

(d) Lad $\{\psi_k\}_{k \in \mathbb{N}_0}$ være koefficienterne i potensrækkeudviklingen af $\theta(z)/\varphi(z)$ og lad $\psi_k = 0$ hvis $k < 0$. Så er

$$\psi_k - \varphi_1 \psi_{k-1} - \cdots - \varphi_p \psi_{k-p} = 0, \quad k > q, \quad (2.23)$$

eller ækvivalent hermed

$$\varphi(B)\psi_k = 0, \quad k > q.$$

(e) Lad $\{\pi_k\}_{k \in \mathbb{N}_0}$ være koefficienterne i potensrækkeudviklingen af $\varphi(z)/\theta(z)$ og lad $\pi_k = 0$ hvis $k < 0$. Så er

$$\pi_k - \theta_1 \pi_{k-1} - \cdots - \theta_q \pi_{k-q} = 0, \quad k > p, \quad (2.24)$$

eller ækvivalent hermed

$$\theta(B)\pi_k = 0, \quad k > p.$$

(f) Hvis $p = 0$, dvs. hvis \mathbf{X} er en MA(q)-proces, så er

$$\gamma_k = \rho_k = 0, \quad k > q. \quad (2.25)$$

(g) Hvis $q = 0$, dvs. hvis \mathbf{X} er en AR(p)-proces, så er

$$\varphi(B)\rho_k = 0, \quad k > 0 \quad (2.26)$$

eller ækvivalent hermed

$$\rho_k = \varphi_1 \rho_{k-1} + \cdots + \varphi_p \rho_{k-p}, \quad k > 0.$$

Disse ligninger for $k = 1, \dots, p$, der kaldes **Yule-Walker ligningerne**, kan skrives som

$$(\rho_1, \dots, \rho_p) = (\varphi_1, \dots, \varphi_p) \mathbf{P}_p. \quad (2.27)$$

Desuden gælder der, at den p 'te partielle autokorrelation er

$$\phi_{p,p} = \varphi_p \quad (2.28)$$

og at

$$\phi_{k,k} = 0, \quad k > p. \quad (2.29)$$

(h) Spektraltætheden for \mathbf{X} er

$$f_X(x) = \frac{\sigma_W^2}{2\pi} \frac{|1 - \sum_{k=1}^q \theta_k e^{-ikx}|^2}{|1 - \sum_{k=1}^p \varphi_k e^{-ikx}|^2}, \quad x \in [-\pi, \pi]. \quad (2.30)$$

Bemærkning 2.22. I [Appendiks B](#) findes der en sætning, der giver den fuldstændige løsning til en homogen differensligning. Af denne sætning, (2.21) og (2.22) ses det, at for en $\text{ARMA}(p, q)$ -proces i \mathcal{A}_0 vil såvel autokovarianserne γ_k som autokorrelationerne ρ_k konvergere "eksponentielt" mod 0 når $k \rightarrow \infty$. Desuden ses det, at hvis ρ_k konvergerer langsomt mod 0, skyldes det, at polynomiet $\varphi(z)$ har en rod z_0 så $|z_0| \approx 1$. Ifølge (a) i [Sætning 2.19](#) betyder dette, at processen er næsten ikke-stationær. ■

Bevis for Sætning 2.21. Fra [Sætning 2.19](#) har vi

$$X_t = \sum_{k=0}^{\infty} \psi_k W_{t-k}, \quad t \in \mathbb{Z}$$

og det ses, at

$$\begin{aligned} \gamma_k = \text{Cov}(X_{t+k}, X_t) &= \text{Cov}\left(\sum_{i=0}^{\infty} \psi_i W_{t+k-i}, \sum_{j=0}^{\infty} \psi_j W_{t-j}\right) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j \text{Cov}(W_{t+k-i}, W_{t-j}) \\ &= \sigma_W^2 \sum_{j=0}^{\infty} \psi_{j+k} \psi_j, \end{aligned}$$

hvilket viser (a).

For at vise (b) bemærker vi indledningsvis, at

$$E(W_s X_t) = E\left(W_s \sum_{k=0}^{\infty} \psi_k W_{t-k}\right) = \sum_{k=0}^{\infty} \psi_k E(W_s W_{t-k}),$$

hvilket medfører, at

$$E(W_s X_t) = 0 \quad \text{hvis } t < s. \quad (2.31)$$

Multipliseres begge sider i (2.17) med X_{t-k} , får vi

$$X_t X_{t-k} - \varphi_1 X_{t-1} X_{t-k} - \cdots - \varphi_p X_{t-p} X_{t-k} = W_t X_{t-k} - \theta_1 W_{t-1} X_{t-k} - \cdots - \theta_q W_{t-q} X_{t-k}.$$

Tager vi middelværdier på begge sider i denne ligning og bruger (2.31), er formel (2.21) vist.

(c) følger umiddelbart af (b), idet $\rho_k = \gamma_k / \gamma_0$.

Ad (d). Sætter vi $\varphi_0 = -1$, har vi for alle $z \in \mathbb{C}$

$$\begin{aligned} 1 - \theta_1 z - \cdots - \theta_q z^q &= \varphi(z) \sum_{m=0}^{\infty} \psi_m z^m = \sum_{m=0}^{\infty} \left(- \sum_{j=0}^p \varphi_j z^j \right) \psi_m z^m \\ &= - \sum_{m=0}^{\infty} \sum_{j=0}^p \varphi_j \psi_m z^{j+m} = - \sum_{k=0}^{\infty} \sum_{j=0}^{\min(p,k)} \varphi_j \psi_{k-j} z^k \\ &= \sum_{k=0}^{\infty} \left(- \sum_{j=0}^p \varphi_j \psi_{k-j} \right) z^k = \sum_{k=0}^{\infty} (\varphi(B) \psi_k) z^k. \end{aligned}$$

Da dette gælder for alle $z \in \mathbb{C}$, følger formel (2.23) ved at sætte koefficienterne til z^k på de to sider af denne ligning lig med hinanden.

Ad (e). Et lignende argument som i (d).

Ad (f). For en MA(q)-proces har vi

$$X_t = W_t - \theta_1 W_{t-1} - \cdots - \theta_q W_{t-q}$$

og det fås nemt, at $\gamma_k = 0$ for $k > q$.

Ad (g). Formel (2.26) er et specielt tilfælde af (2.22), og at kombinere disse ligninger for at få Yule-Walker ligningerne er trivielt. Det følger direkte fra Yule-Walker ligningerne og definitionen af den partielle autokorrelationsfunktion, at $\phi_{p,p} = \varphi_p$. For $k > p$ kan en AR(p)-proces betragtes som en AR(k)-proces med $\varphi_{p+1} = \cdots = \varphi_k = 0$. Af det netop viste fås, at $\phi_{k,k} = 0$.

Ad (h). Sætter vi $\theta_0 = \varphi_0 = -1$ kan ligningen (2.20) skrives som

$$- \sum_{j=0}^p \varphi_j X_{t-j} = - \sum_{m=0}^q \theta_m W_{t-m}.$$

Multipliseres denne ligning med den tilsvarende med t erstattet med $t+k$, hvor $k \in \mathbb{N}$, og tager vi middelværdier, finder vi følgende ækvivalente ligheder:

$$\begin{aligned} \sum_{j=0}^p \sum_{l=0}^p \varphi_j \varphi_l E(X_{t-j} X_{t+k-l}) &= \sum_{m=0}^q \sum_{n=0}^q \theta_m \theta_n E(W_{t-m} W_{t+k-n}), \\ \sum_{j=0}^p \sum_{l=0}^p \varphi_j \varphi_l \gamma_X(k-l+j) &= \sum_{m=0}^q \sum_{n=0}^q \theta_m \theta_n \gamma_W(k-n+m), \\ \sum_{j=0}^p \sum_{l=0}^p \varphi_j \varphi_l \int_{-\pi}^{\pi} e^{i(k-l+j)x} dF_X(x) &= \sum_{m=0}^q \sum_{n=0}^q \theta_m \theta_n \int_{-\pi}^{\pi} e^{i(k-n+m)x} dF_W(x), \\ \int_{-\pi}^{\pi} \left(\sum_{j=0}^p \varphi_j e^{ijx} \right) \left(\sum_{l=0}^p \varphi_l e^{-ilx} \right) e^{ikx} dF_X(x) &= \int_{-\pi}^{\pi} \left(\sum_{m=0}^q \theta_m e^{imx} \right) \left(\sum_{n=0}^q \theta_n e^{-inx} \right) e^{ikx} dF_W(x), \\ \int_{-\pi}^{\pi} \left| \sum_{l=0}^p \varphi_l e^{-ilx} \right|^2 e^{ikx} dF_X(x) &= \int_{-\pi}^{\pi} \left| \sum_{n=0}^q \theta_n e^{-inx} \right|^2 e^{ikx} dF_W(x). \end{aligned}$$

Da spektralmålet er entydigt bestemt, er (h) vist. □

Lad

$$\Phi_p = \{ \boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_p) \in R^p : 1 - \varphi_1 z - \dots - \varphi_p z^p \neq 0 \text{ hvis } |z| \leq 1 \}$$

og

$$\Theta_q = \{ \boldsymbol{\theta} = (\theta_1, \dots, \theta_q) \in R^q : 1 - \theta_1 z - \dots - \theta_q z^q \neq 0 \text{ hvis } |z| \leq 1 \}.$$

Med denne notation har vi, at ARMA(p, q)-processen \mathbf{X} givet ved

$$\varphi(B)X_t = \theta(B)W_t, \quad t \in \mathbb{Z}$$

kan repræsenteres såvel som et baglæns glidende gennemsnit som en (baglæns) autoregression hvis og kun hvis $(\boldsymbol{\varphi}, \boldsymbol{\theta}) \in \Phi_p \times \Theta_q$, dvs.

$$\mathbf{X} \in \mathcal{A}_0 \iff (\boldsymbol{\varphi}, \boldsymbol{\theta}) \in \Phi_p \times \Theta_q. \quad (2.32)$$

2.3 Egenskaber ved simple ARMA-processer

I dette afsnit listes nogle af de vigtigste egenskaber ved svagt stationære ARMA(p, q)-processer med $p + q < 3$.

AR(1)-processen

Af (2.32) fås, at AR(1)-processen (med middelværdi 0)

$$X_t = \varphi_1 X_{t-1} + W_t, \quad t \in \mathbb{Z},$$

kan repræsenteres som et baglæns glidende gennemsnit, hvis

$$\varphi_1 \in \Phi_1 =]-1, 1[.$$

Formel (2.31) medfører, at $E(X_{t-1} W_t) = 0$ og vi finder

$$\begin{aligned} EX_t^2 &= E(\varphi_1 X_{t-1} + W_t)^2 \\ &= \varphi_1^2 EX_{t-1}^2 + EW_t^2, \end{aligned}$$

Heraf ses, at

$$\sigma_X^2 = \frac{\sigma_W^2}{1 - \varphi_1^2}.$$

Ved hjælp af (2.26) bliver autokorrelationsfunktionen

$$\rho_k = \varphi_1^k, \quad k \in \mathbb{N}_0,$$

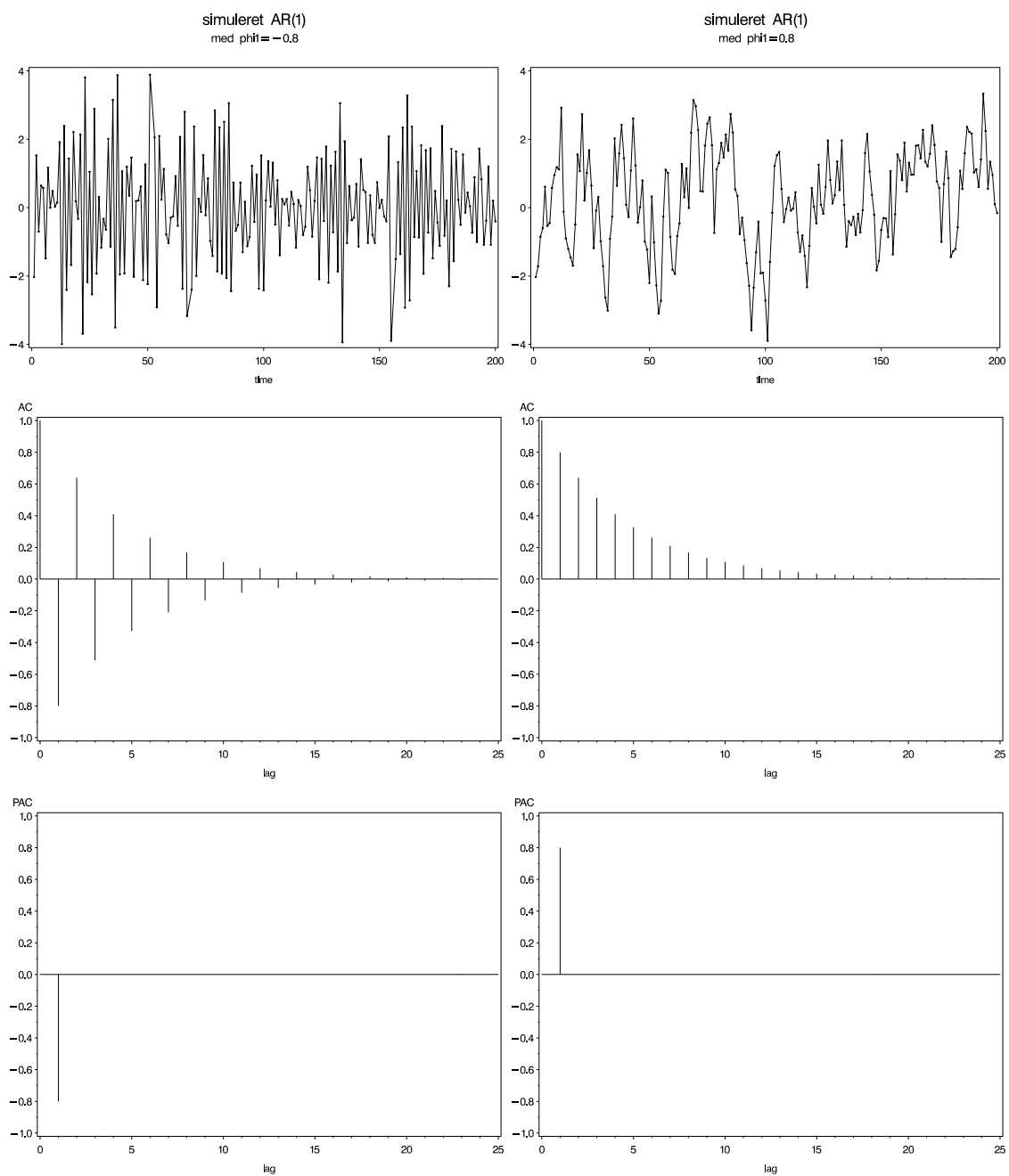
og ifølge (g) i Sætning 2.21 er den partielle autokorrelationsfunktion

$$\phi_{k,k} = \begin{cases} \varphi_1 & \text{hvis } k = 1 \\ 0 & \text{hvis } k > 1. \end{cases}$$

Fra (2.30) fås det, at spektraltætheden er

$$f_X(x) = \frac{\sigma_W^2}{2\pi} |1 - \varphi_1 e^{-ix}|^{-2}, \quad x \in [-\pi, \pi].$$

Figur 2.1 på næste side viser simulerede udfaldsfunktioner og de teoretiske korrelationsfunktioner for to AR(1)-processer.



Figur 2.1: Simuleret udfaldsfunktion og de teoretiske korrelationsfunktioner for en AR(1)-proces med $\phi_1 = -0.8$ til venstre og med $\phi_1 = 0.8$ til højre.

MA(1)-processen

MA(1)-processen (med middelværdi 0)

$$X_t = W_t - \theta_1 W_{t-1}, \quad t \in \mathbb{Z},$$

kan repræsenteres som en baglæns autoregression, hvis

$$\theta_1 \in \Theta_1 =]-1, 1[.$$

Da W er en hvid støj, har vi

$$\begin{aligned} EX_t^2 &= E(W_t - \theta_1 W_{t-1})^2 \\ &= EW_t^2 + \theta_1^2 EW_{t-1}^2, \end{aligned}$$

dvs.

$$\sigma_X^2 = (1 + \theta_1^2)\sigma_W^2,$$

og

$$\begin{aligned} \gamma_k &= E\{(W_t - \theta_1 W_{t-1})(W_{t+k} - \theta_1 W_{t+k-1})\} \\ &= E(W_t W_{t+k}) - \theta_1 E(W_{t-1} W_{t+k}) - \theta_1 E(W_t W_{t+k-1}) + \theta_1^2 E(W_{t-1} W_{t+k-1}) \\ &= \begin{cases} -\theta_1 \sigma_W^2 & \text{hvis } |k| = 1 \\ 0 & \text{hvis } |k| \neq 1, \end{cases} \end{aligned}$$

så autokorrelationsfunktionen er

$$\rho_k = \begin{cases} \frac{-\theta_1}{1+\theta_1^2} & \text{hvis } k = 1 \\ 0 & \text{hvis } k \geq 2. \end{cases}$$

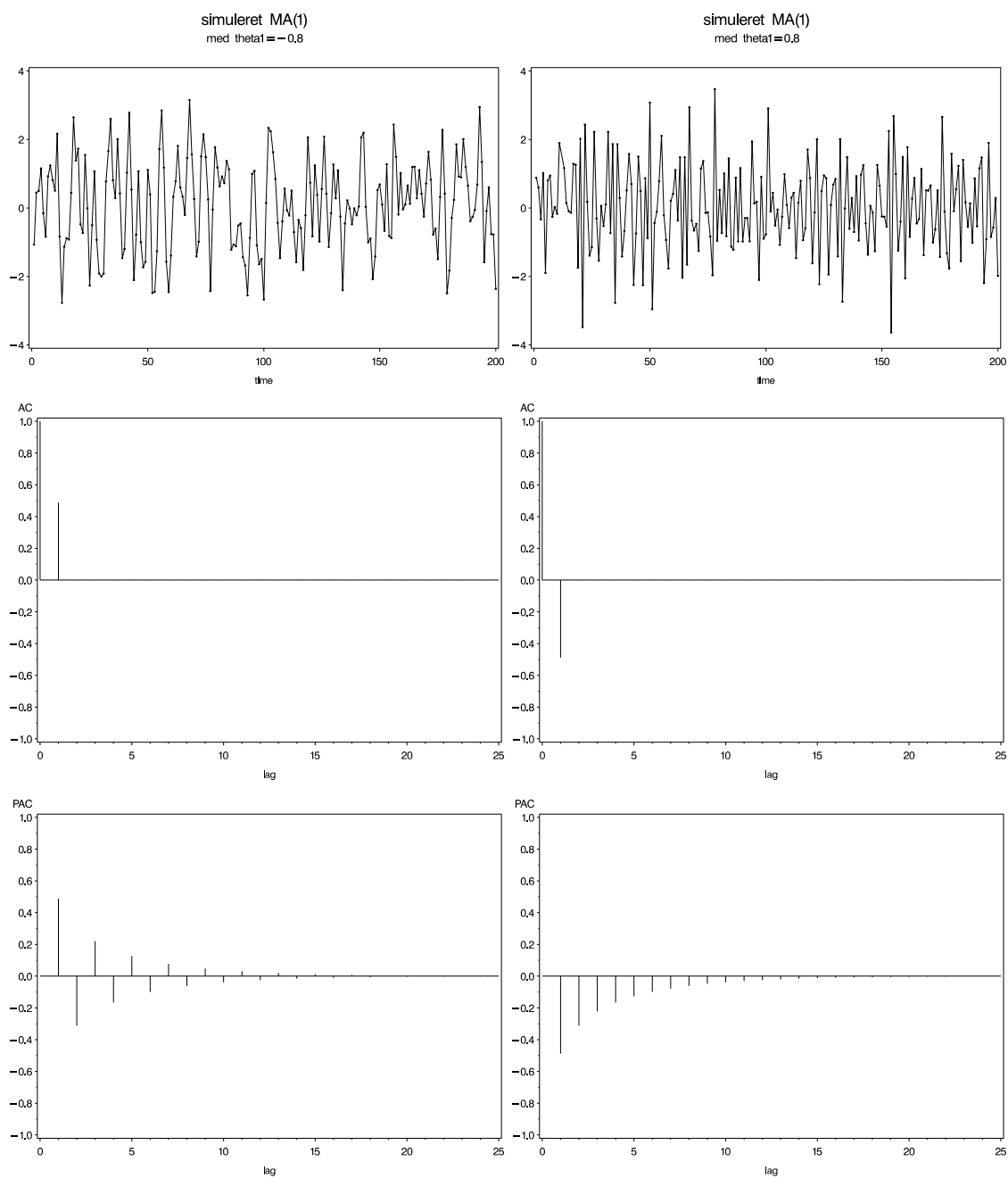
Benyttes formlerne, der udtrykker de partielle autokorrelationer $\{\phi_{k,k}\}$ ved hjælp af autokorrelationerne $\{\rho_k\}$ ses det, at den partielle autokorrelationsfunktion er

$$\phi_{k,k} = -\theta_1^k \frac{1 - \theta_1^2}{1 - \theta_1^{2(k+1)}}, \quad k \geq 1.$$

Endelig medfører formel (2.30), at spektraltætheden er

$$f_X(x) = \frac{\sigma_W^2}{2\pi} |1 - \theta_1 e^{-ix}|^2, \quad x \in [-\pi, \pi].$$

Figur 2.2 på modstående side viser simulerede udfaldsfunktioner for to MA(1)-processer sammen med de tilsvarende teoretiske korrelationsfunktioner.



Figur 2.2: Simuleret udfaldsfunktion og de teoretiske korrelationsfunktioner for en MA(1)-proces med $\theta_1 = -0.8$ til venstre og med $\theta_1 = 0.8$ til højre.

AR(2)-processen

AR(2)-processen (med middelværdi 0)

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + W_t, \quad t \in \mathbb{Z}, \quad (2.33)$$

kan repræsenteres som et baglæns glidende gennemsnit, hvis

$$(\varphi_1, \varphi_2) \in \Phi_2 = \{\varphi = (\varphi_1, \varphi_2) \in \mathbb{R}^2 : \varphi(z) = 1 - \varphi_1 z - \varphi_2 z^2 \neq 0 \text{ hvis } |z| \leq 1\}.$$

Mængden Φ_2 er det indre af mængden bestemt af linjerne

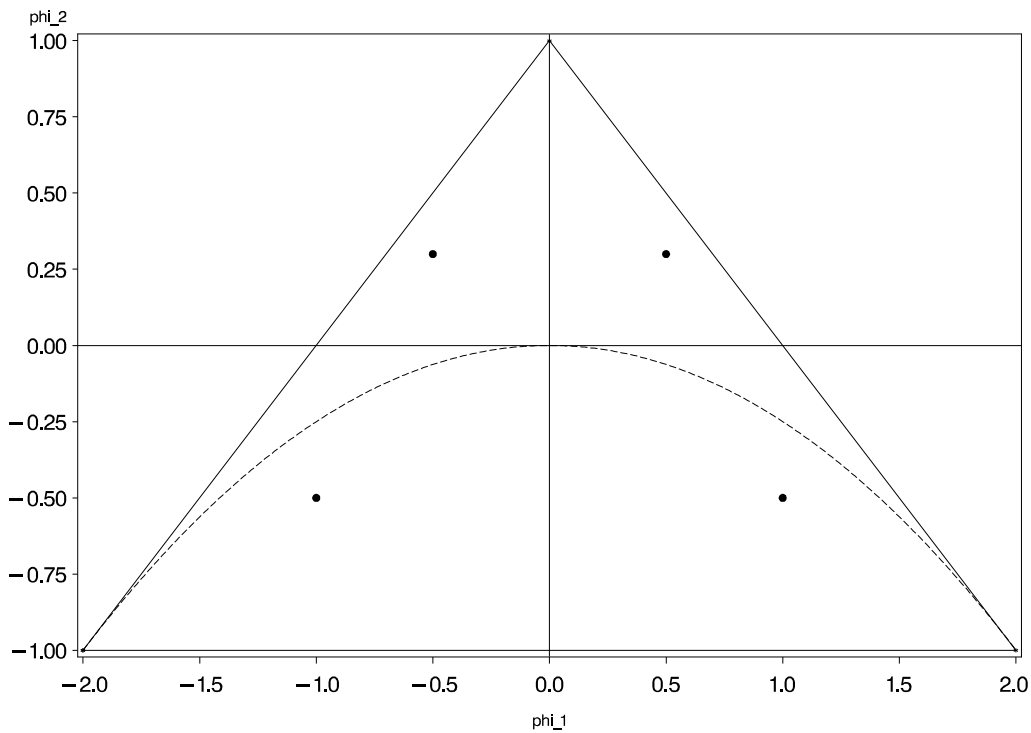
$$\varphi_2 + \varphi_1 = 1,$$

$$\varphi_2 - \varphi_1 = 1$$

og

$$\varphi_2 = -1,$$

eller, ækvivalent hermed, det indre af trekanten T bestemt af punkterne $(\varphi_1, \varphi_2) = (0, 1)$, $(\varphi_1, \varphi_2) = (2, -1)$ and $(\varphi_1, \varphi_2) = (-2, -1)$, se [Figur 2.3](#).



Figur 2.3: Parametermængden Φ_2 for en AR(2)-proces. For parametre over og under den stiplede linje har polynomiet $\varphi(z)$ henholdsvis reelle og komplekse rødder. Prikkerne antyder parametrene for de fire AR(2)-processer, der simuleres nedenfor.

Yule-Walker ligningerne for en AR(2)-proces er, se [\(2.27\)](#),

$$\rho_1 = \varphi_1 + \varphi_2 \rho_1 \quad (2.34)$$

$$\rho_2 = \varphi_1 \rho_1 + \varphi_2 \quad (2.35)$$

og løsningen til disse ligninger er

$$\varphi_1 = \rho_1(1 - \rho_2)/(1 - \rho_1^2) \quad (2.36)$$

$$\varphi_2 = (\rho_2 - \rho_1^2)/(1 - \rho_1^2). \quad (2.37)$$

Fra (2.34) finder vi, at

$$\rho_1 = \varphi_1/(1 - \varphi_2) \quad (2.38)$$

$$\rho_2 = \varphi_2 + \varphi_1^2/(1 - \varphi_2). \quad (2.39)$$

For at finde σ_X^2 multiplicerer vi med X_t på begge sider af (2.33), hvilket giver

$$X_t X_t = \varphi_1 X_{t-1} X_t + \varphi_2 X_{t-2} X_t + W_t(\varphi_1 X_{t-1} + \varphi_2 X_{t-2} + W_t).$$

Beregner vi middelværdien på begge sider og bruger (2.31), får vi at

$$\gamma_0 - \varphi_1 \gamma_1 - \varphi_2 \gamma_2 = \sigma_W^2,$$

eller

$$\sigma_X^2 = \frac{\sigma_W^2}{1 - \varphi_1 \rho_1 - \varphi_2 \rho_2}$$

hvilket udtrykt ved φ 'er bliver

$$\sigma_X^2 = \frac{1 - \varphi_2}{1 + \varphi_2} \frac{\sigma_W^2}{(1 - \varphi_2)^2 - \varphi_1^2}.$$

Autokorrelationsfunktionen opfylder – ifølge (g) i Sætning 2.21 – differensligningen $\varphi(B)\rho_k = 0$, $k > 0$. Hvis G_1^{-1} og G_2^{-1} betegner rødderne i $\varphi(z) = 0$, fås det af Sætning B.1 i Appendiks B, at

$$\rho_k = A_1 G_1^k + A_2 G_2^k, \quad k \geq -1,$$

hvis $G_1 \neq G_2$ og

$$\rho_k = (A_1 + kA_2)G_1^k, \quad k \geq -1,$$

hvis $G_1 = G_2$. I begge tilfælde kan konstanterne A_1 og A_2 findes fra (2.39).

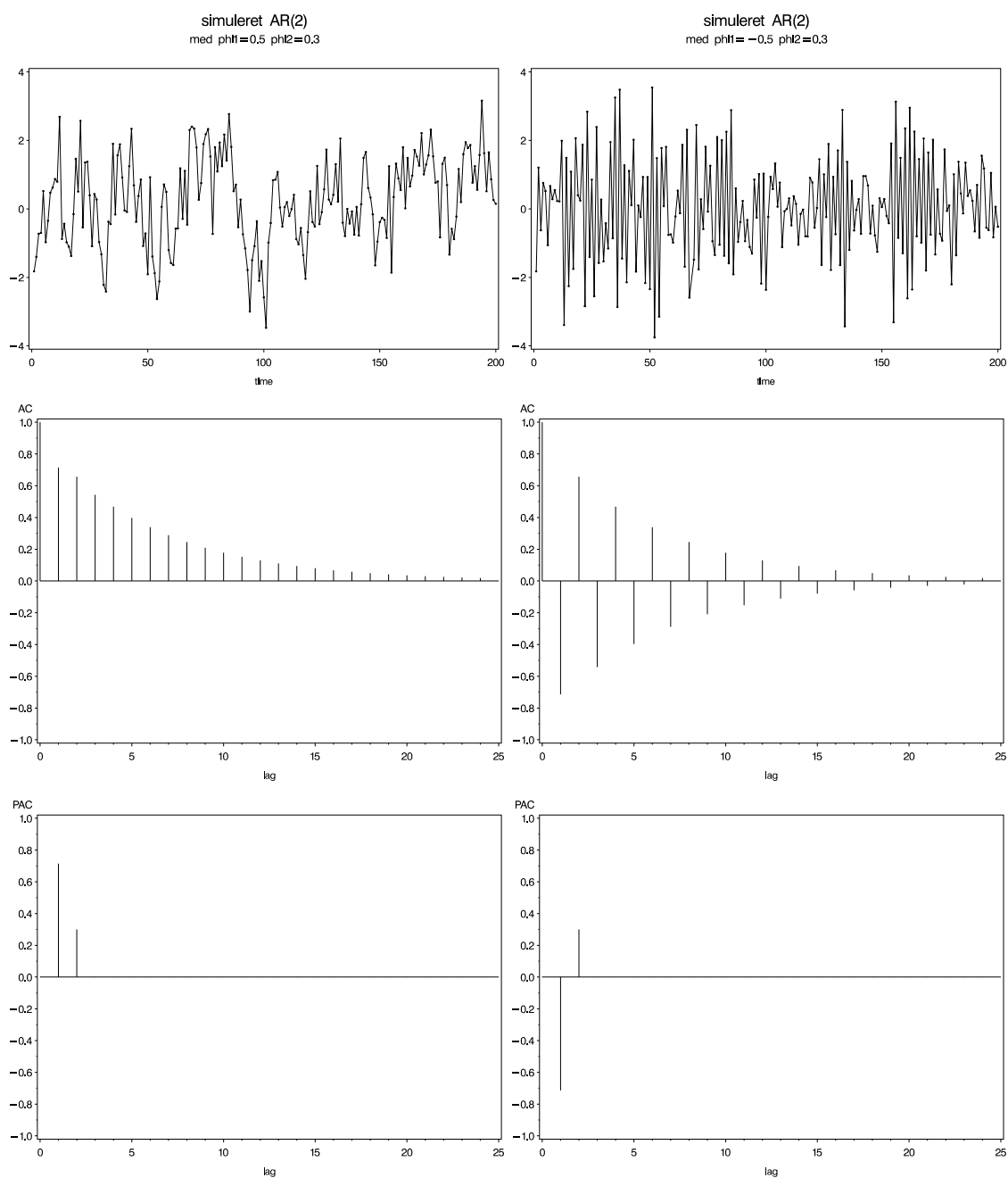
Fra (g) i Sætning 2.21 får vi, at den partielle autokorrelationsfunktion er

$$\phi_{k,k} = \begin{cases} \frac{\varphi_1}{1 - \varphi_2} & \text{hvis } k = 1 \\ \varphi_2 & \text{hvis } k = 2 \\ 0 & \text{hvis } k > 2, \end{cases} \quad (2.40)$$

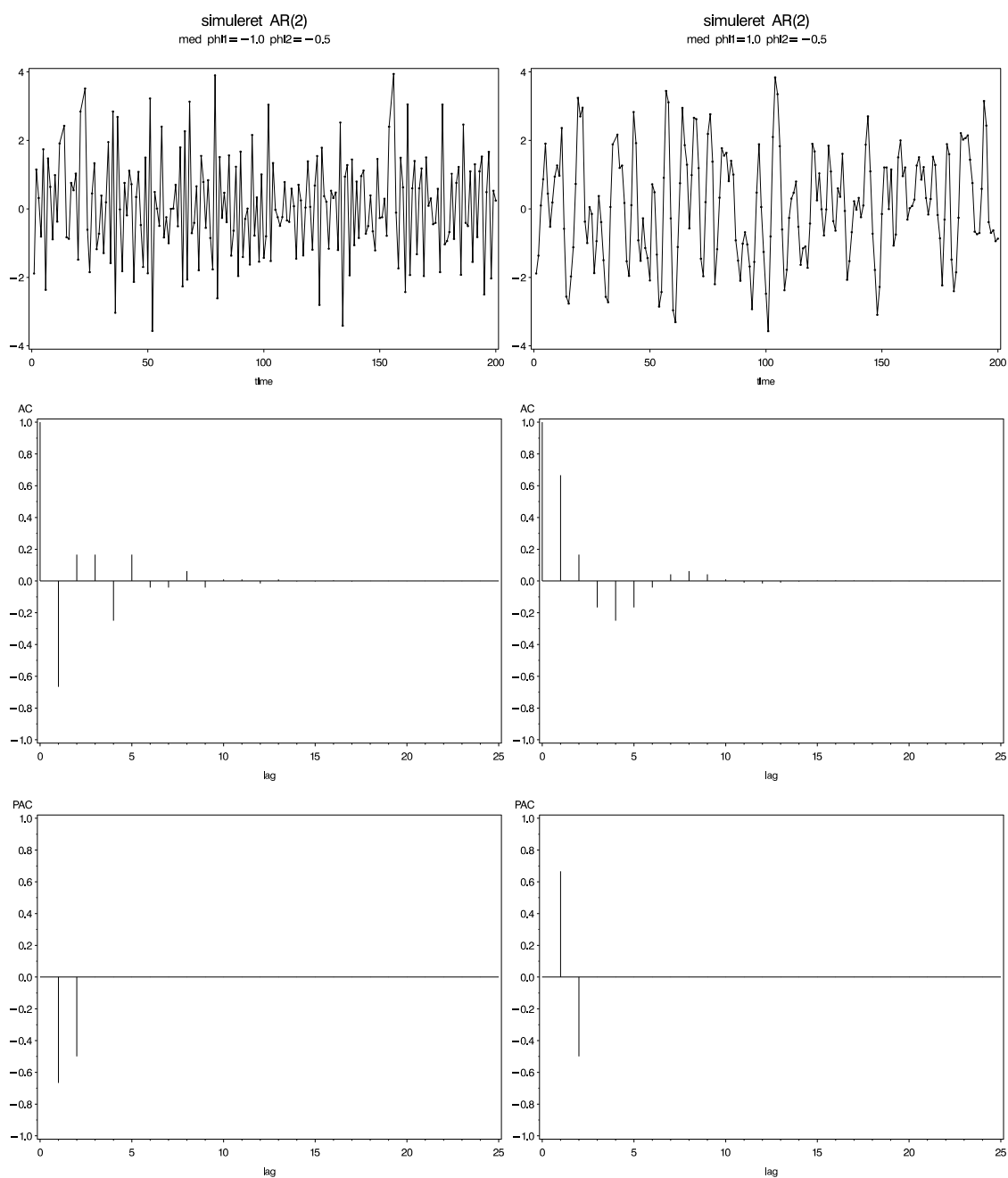
og fra (h) i Sætning 2.21 fås, at spektraltætheden er

$$f_X(x) = \frac{\sigma_W^2}{2\pi} |1 - \varphi_1 e^{-ix} - \varphi_2 e^{-i2x}|^{-2}, \quad x \in [-\pi, \pi].$$

Figur 2.4 og Figur 2.5 nedenfor viser simulerede udfaldsfunktioner for fire AR(2)-processer sammen med de teoretiske korrelationsfunktioner. Parametrene for processerne er antydnet i Figur 2.3.



Figur 2.4: Simuleret udfaldsfunktion og de teoretiske korrelationsfunktioner for en AR(2)-proces med $\varphi_1 = 0.5$ og $\varphi_2 = 0.3$ til venstre og med $\varphi_1 = -0.5$ og $\varphi_2 = 0.3$ til højre.



Figur 2.5: Simuleret udfaldsfunktion og de teoretiske korrelationsfunktioner for en AR(2)-proces med $\varphi_1 = -1.0$ og $\varphi_2 = -0.5$ til venstre og med $\varphi_1 = 1.0$ og $\varphi_2 = -0.5$ til højre.

MA(2)-processen

MA(2)-processen (med middelværdi 0)

$$X_t = W_t - \theta_1 W_{t-1} - \theta_2 W_{t-2}, \quad t \in \mathbb{Z}, \quad (2.41)$$

kan repræsenteres som en (baglæns) autoregression, hvis

$$(\theta_1, \theta_2) \in \Theta_2 = \{\theta = (\theta_1, \theta_2) \in \mathbb{R}^2 : \theta(z) = 1 - \theta_1 z - \theta_2 z^2 \neq 0 \text{ hvis } |z| \leq 1\}.$$

Mængden Θ_2 er det indre af mængden bestemt ved ligningerne

$$\theta_2 + \theta_1 = 1,$$

$$\theta_2 - \theta_1 = 1$$

og

$$\theta_2 = -1,$$

eller, ækvivalent hermed, det indre af trekanten bestemt af punkterne $(\theta_1, \theta_2) = (0, 1)$, $(\theta_1, \theta_2) = (2, -1)$ og $(\theta_1, \theta_2) = (-2, -1)$.

Variansen er

$$\sigma_X^2 = \sigma_W^2(1 + \theta_1^2 + \theta_2^2)$$

og autokorrelationsfunktionen er givet ved

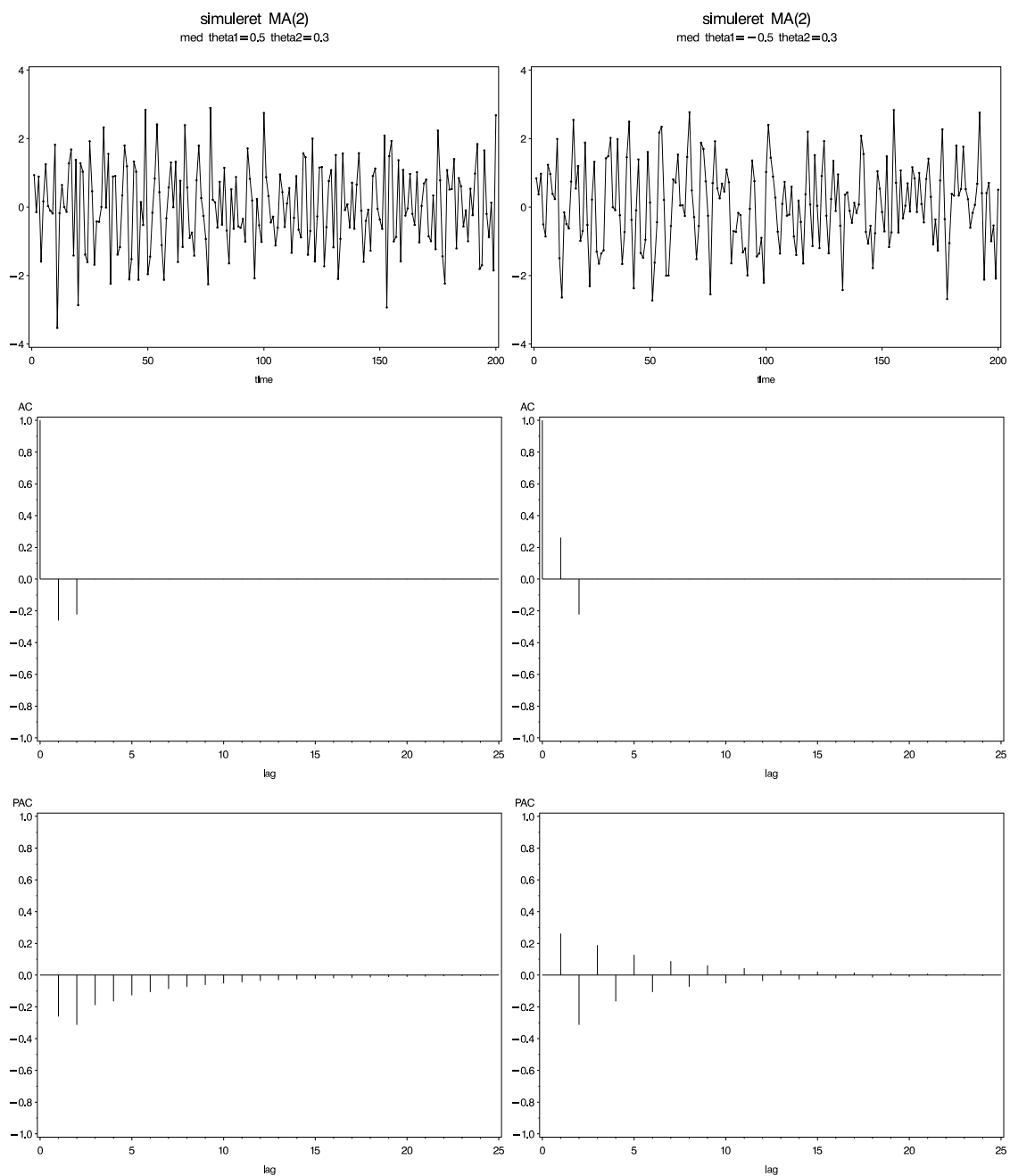
$$\rho_k = \begin{cases} \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} & \text{hvis } k = 1 \\ \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} & \text{hvis } k = 2 \\ 0 & \text{hvis } k \geq 3. \end{cases} \quad (2.42)$$

Det er muligt at udtrykke de partielle autokorrelationer $\phi_{k,k}$ som en funktion af (θ_1, θ_2) ved hjælp af rekursionsformlen for de partielle autokorrelationer og (2.42). Udtrykkene er forholdsvis komplicerede, men det kan vises, at den partielle autokorrelationsfunktion for en MA(2)-proces ligner autokorrelationsfunktionen for en AR(2)-proces. Sammenlign Figur 2.4 og Figur 2.5 med Figur 2.6 og Figur 2.7.

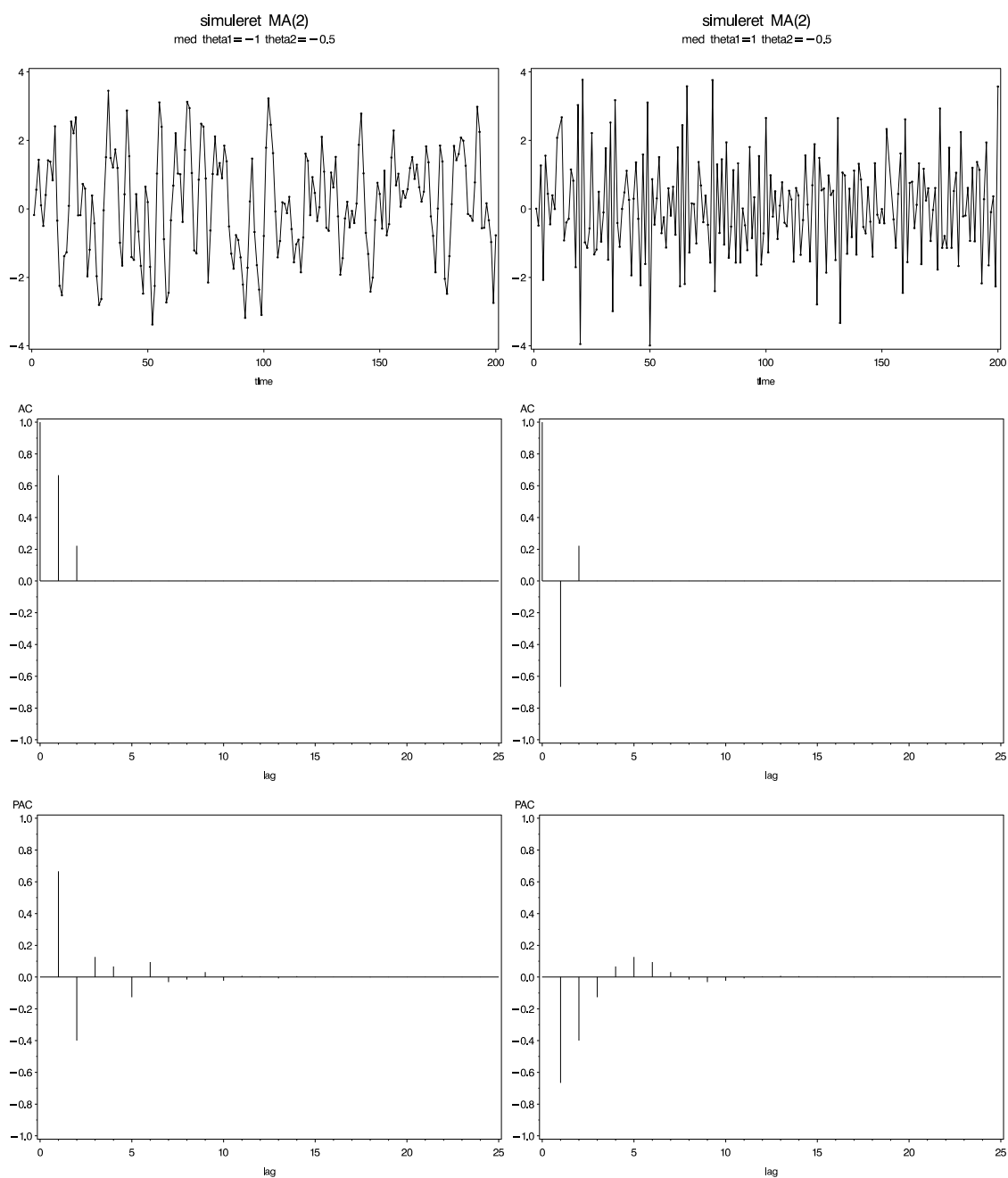
Spektraltætheden er

$$f_X(x) = \frac{\sigma_W^2}{2\pi} |1 - \theta_1 e^{-ix} - \theta_2 e^{-i2x}|^2, \quad x \in [-\pi, \pi].$$

Figur 2.6 og Figur 2.7 nedenfor viser simulerede udfaldsfunktioner for fire MA(2)-processer sammen med de teoretiske korrelationsfunktioner.



Figur 2.6: Simuleret udfaldsfunktion og de teoretiske korrelationsfunktioner for en MA(2)-proces med $\theta_1 = 0.5$ og $\theta_2 = 0.3$ til venstre og med $\theta_1 = -0.5$ og $\theta_2 = 0.3$ til højre.



Figur 2.7: Simuleret udfaldsfunktion og de teoretiske korrelationsfunktioner for en AR(2)-proces med $\theta_1 = -1.0$ og $\theta_2 = -0.5$ til venstre og med $\theta_1 = 1.0$ og $\theta_2 = -0.5$ til højre.

ARMA(1, 1)-processen

Processen

$$X_t = \varphi_1 X_{t-1} + W_t - \theta_1 W_{t-1}, \quad t \in \mathbb{Z}, \quad (2.43)$$

kan repræsenteres såvel som et baglæns glidende gennemsnit som en baglæns autoregression, hvis

$$(\varphi_1, \theta_1) \in (-1, 1)^2.$$

Det vises let, at

$$\sigma_X^2 = \sigma_W^2 \frac{1 + \theta_1^2 - 2\varphi_1\theta_1}{1 - \varphi_1^2},$$

$$\rho_k = \begin{cases} \frac{(1 - \varphi_1\theta_1)(\varphi_1 - \theta_1)}{1 + \theta_1^2 - 2\varphi_1\theta_1} & \text{hvis } k = 1 \\ \varphi_1^{k-1} \rho_1 & \text{hvis } k \geq 2, \end{cases} \quad (2.44)$$

og at

$$f_X(x) = \frac{\sigma_W^2}{2\pi} \frac{|1 - \theta_1 e^{-ix}|^2}{|1 - \varphi_1 e^{-ix}|^2}, \quad x \in [-\pi, \pi].$$

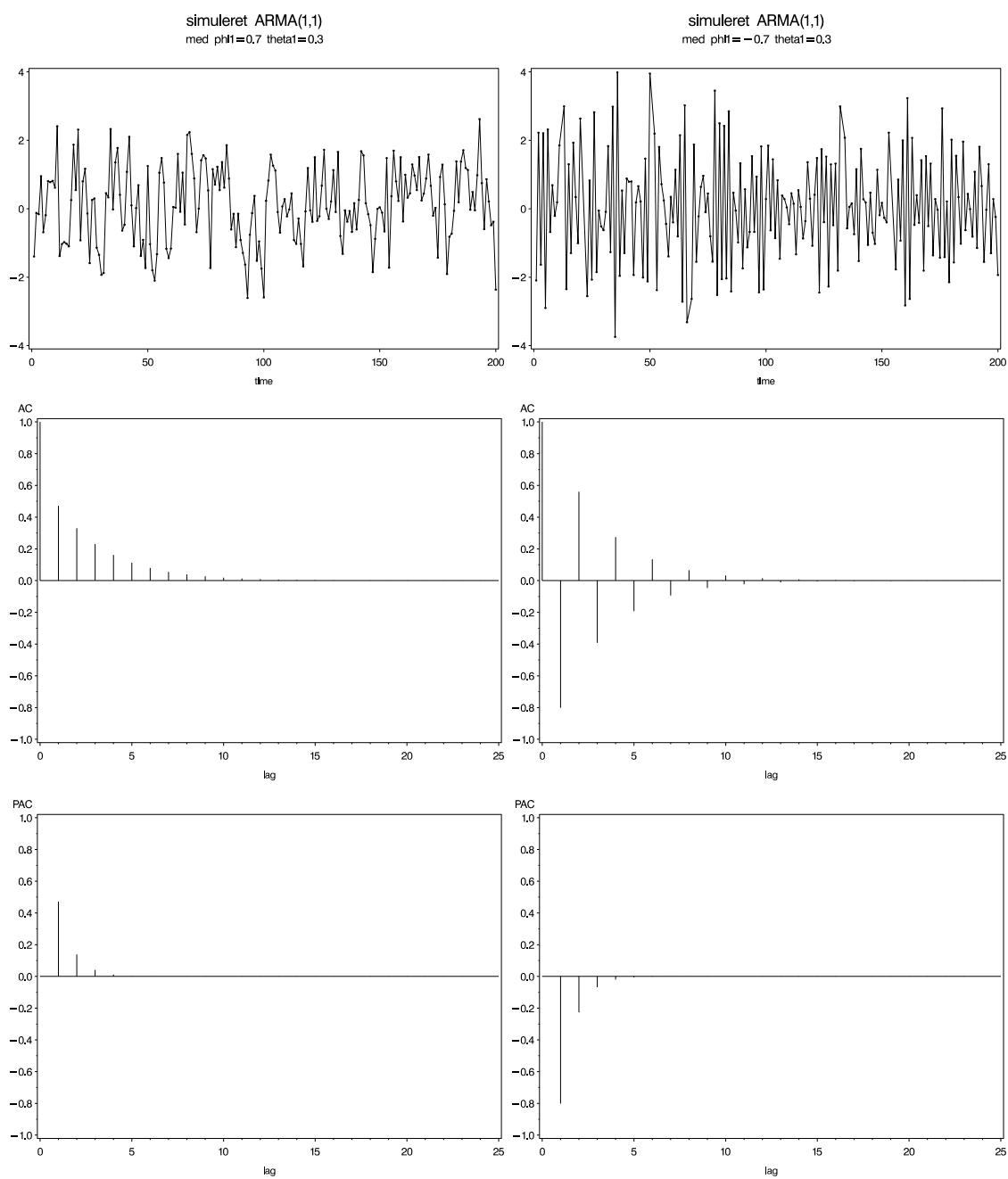
Det kan desuden vises, at den numeriske værdi af den partielle autokorrelationsfunktion for en ARMA(1, 1)-proces aftager eksponentielt mod 0. Hvis $\theta_1 > 0$, er følgen $\{\rho_{k,k}\}$ monoton, voksende hvis $\theta_1 > \varphi_1$, og aftagende hvis $\theta_1 < \varphi_1$. For $\theta_1 < 0$ er den partielle autokorrelationsfunktion oscillerende.

Bemærk, at en ARMA(1, 1)-model med $\varphi_1 = \theta_1 = \tau$, dvs.

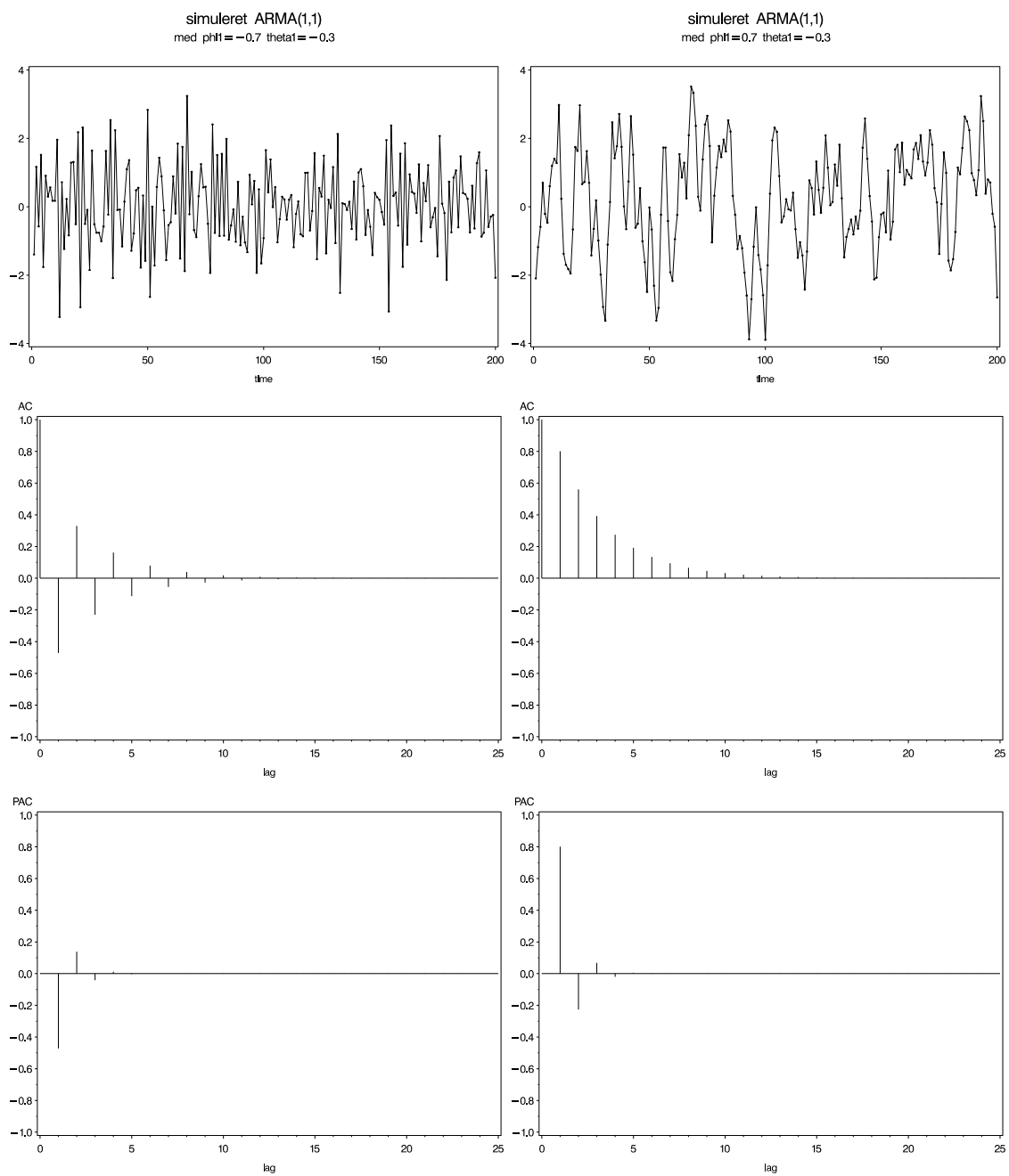
$$(1 - \tau B)X_t = (1 - \tau B)W_t, \quad t \in \mathbb{Z},$$

er en hvid støj, hvilket ses ved at forkorte med $(1 - \tau B)$ på begge sider i ovenstående ligning.

[Figur 2.8](#) og [Figur 2.9](#) nedenfor viser simulerede udfaldsfunktioner for fire ARMA(1, 1)-processer samt de tilsvarende teoretiske korrelationsfunktioner.



Figur 2.8: Simuleret udfaldsfunktion og de teoretiske korrelationsfunktioner for en ARMA(1,1)-proces med $\varphi_1 = 0.7$ og $\theta_1 = 0.3$ til venstre og med $\varphi_1 = -0.7$ og $\theta_1 = 0.3$ til højre.



Figur 2.9: Simuleret udfaldsfunktion og de teoretiske korrelationsfunktioner for en $ARMA(1,1)$ -proces med $\varphi_1 = -0.7$ og $\theta_1 = -0.3$ til venstre og med $\varphi_1 = 0.7$ og $\theta_1 = -0.3$ til højre.

Kapitel 3

Empiriske tidsrækker I

I disse noter bruges SAS til at analysere empiriske tidsrækker. SAS har to procedurer til dette formål, nemlig PROC ARIMA og PROC VARMAX. PROC ARIMA er en forholdsvis gammel procedure, der i adskillige år har været brugt til blandt andet at analysere endimensionale tidsrækker ved hjælp af ARIMA-modeller, sæsonmodeller, transfermodeller m.m. En begrænsning ved proceduren er, at den ikke er i stand til at analysere flerdimensionale tidsrækker. Det kan den forholdsvis nye procedure PROC VARMAX, som også er i stand til at lave beregninger i endimensionale tidsrækker. Proceduren har dog stadig nogle skavanker, der bevirker, at PROC ARIMA foretrækkes ved analysen af endimensionale tidsrækker. Det forekommer derfor naturligt, at opdele omtalen af empiriske tidsrækker i to dele, hvor vi betragter henholdsvis en- og flerdimensionale tidsrækker.

3.1 Endimensionale tidsrækker

Ved en endimensional empirisk tidsrække forstår vi målinger, x_1, \dots, x_N , foretaget til ækvidistante tidspunkter af en relevant størrelse ved et system, der udvikler sig i tid. Idet vi betragter differensen mellem to på hinanden følgende tidspunkter som tidsenhed, vil vi omtale x_t som observationen til tid t . Empiriske tidsrækker kan stamme fra et system med kontinuert tid ved at måle på systemet til tidspunkter $1, \dots, N$, for eksempel målinger af temperatur, vandstand osv., eller ved at akkumulere en størrelse over et tidsinterval, for eksempel nedbør per dag. Endimensionale tidsrækker kan også være komponenterne i en flerdimensional tidsrække.

Formålet med at analysere en tidsrække er typisk at opnå indsigt i det system, der betragtes, for at kunne kontrollere systemet eller for at kunne forudsige systemets fremtidige udvikling.

Den statistiske analyse af en empirisk tidsrække, x_1, \dots, x_N , skal tage i betragtning, at det betragtede system udvikler sig i tiden, og at observationen x_t til tid

t er mere eller mindre afhængig af den tidligere udvikling af systemet og derfor også af værdierne af x_1, \dots, x_{t-1} . De sædvanlige statistiske modeller, hvor observationerne betragtes som udfald af uafhængige stokastiske variable, kan derfor ikke anvendes i denne sammenhæng. Det, der er brug for i tidsrækkeanalyse, er modeller, der er i stand til at beskrive systemets udvikling, som det kommer til udtryk ved den målte størrelse, og derfor også sammenhængen mellem observationer svarende til forskellige tidspunkter.

En statistisk model for en empirisk tidsrække kan opnås ved at betragte x_1, \dots, x_N som *en del af* udfaldsfunktionen for en stokastisk proces $X = \{X_t\}_{t \in \mathbb{Z}}$. Antagelserne vedrørende sammenhængen mellem observationer til forskellige tidspunkter kan så formuleres ved hjælp af familien af endelig dimensionale fordelinger for X .

I de eksempler, vi betragter i dette afsnit, måles der som nævnt kun en enkelt reel størrelse, så derfor antager vi, at *processen X er reel*.

Visse tidsrækker, som for eksempel aktiekurser, er ikke reproducerbare i den forstand, at vi kan ikke fastholde eller rekonstruere de betingelser, som observationerne x_1, \dots, x_N blev foretaget under. Dette tager vi højde for ved, som beskrevet ovenfor, at betragte tidsrækken som en del af en udfaldsfunktion for den stokastiske proces X . Den sædvanlige antagelse, at observationerne er udfald af uafhængige og identisk fordelte stokastiske variable, grunder næsten altid i, at betingelserne, under hvilke målingerne er foretaget, er reproducerbare. Derfor er det betragtede setup for tidsrækker forskelligt fra det, vi normalt betragter.

Da en empirisk tidsrække kun er en del af udfaldsfunktionen for den stokastiske proces X , er det indlysende, at det er nødvendigt at lave relativt restriktive antagelser om familien af endelig dimensionale fordelinger for X , såsom stationaritet, svag stationaritet og/eller en simpel kovariansstruktur.

Rimeligheden af antagelsen om svag stationaritet kan vurderes ved at indtegne punkterne (t, x_t) , $t = 1, \dots, N$, i et koordinatsystem som i [Figur 1.1–1.4](#). Hvis middelværdien μ for processen X estimeres ved

$$\hat{\mu} = \bar{x} = \frac{1}{N} \sum_{t=1}^N x_t,$$

skal punkterne i tegningen ligge omkring niveauet \bar{x} uden systematiske afvigelser, såsom, for eksempel, et niveauskift. For en svagt stationær proces afhænger den teoretiske kovarians mellem X_t og X_{t+k} kun af afstanden k mellem tidspunkterne, og også med hensyn til dette aspekt må punkterne i tegningen ikke afvige systematisk.

Lad $\{\gamma_k\}_{k \in \mathbb{N}_0}$ betegne kovariansfunktionen for den svagt stationære proces X . Som et estimat for γ_k baseret på den empiriske tidsrække x_1, \dots, x_N betragter [Box](#)

og Jenkins (1970) størrelsen

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (x_t - \bar{x})(x_{t+k} - \bar{x}), \quad k = 0, 1, \dots, K, \quad (3.1)$$

for $K \leq N/4$ og $N \geq 50$. Hvis vi dividerer med antallet af led i summen, $N - k$, i stedet for med N , får vi et estimat \tilde{c}_k for γ_k , der sommetider bruges i stedet for c_k . Forskellen mellem c_k og \tilde{c}_k er naturligvis størst for store værdier af k og små værdier af N .

Korrelationen ρ_k mellem X_t og X_{t+k} estimeres ved

$$r_k = \frac{c_k}{c_0}, \quad k = 0, 1, \dots, K. \quad (3.2)$$

For visse modeller kan de *empiriske autokorrelationer* r_k , $k = 0, 1, \dots, K$, bruges til at vurdere antagelsen om svag stationaritet. Som det vises senere, er dette tilfældet for ARMA-modeller.

I praksis er antagelsen om stationaritet af X generelt umulig at checke. Hvis det yderligere antages, at X er en anden ordens proces, så er X også en svagt stationær proces, og vi kan benytte metoderne nævnt ovenfor. For Gaussiske processer er antagelserne om stationaritet og svag stationaritet ækvivalente.

For en svagt stationær Gaussisk proces giver Bartlett (1955) følgende approksimationer af fordelingen for r_k . For $N \rightarrow \infty$ er r_k approksimativt normalfordelt og

$$Er_k \approx \rho_k. \quad (3.3)$$

Yderligere gælder der for $N \rightarrow \infty$, at

$$\begin{aligned} & \text{Cov}(r_k, r_{k+l}) \\ & \approx \frac{1}{N} \sum_{i=-\infty}^{\infty} (\rho_i \rho_{i+l} + \rho_{i+l+k} \rho_{i-k} + 2\rho_k \rho_{k+l} \rho_i^2 - 2\rho_k \rho_i \rho_{k+l-i} - 2\rho_{k+l} \rho_i \rho_{i-k}) \end{aligned} \quad (3.4)$$

og derfor, at

$$\text{Var } r_k \approx \frac{1}{N} \sum_{i=-\infty}^{\infty} (\rho_i^2 + \rho_{i+k} \rho_{i-k} + 2\rho_k^2 \rho_i^2 - 4\rho_k \rho_i \rho_{i-k}). \quad (3.5)$$

Hvis det ikke er muligt at beskrive den empiriske tidsrække x_1, \dots, x_N ved hjælp af en svagt stationær proces, er den mest anvendte "udvej" at definere en ny empirisk tidsrække $\tilde{x}_{v+1}, \dots, \tilde{x}_{N-u}$ ud fra x_1, \dots, x_N ved

$$\tilde{x}_t = a_{-v} x_{t-v} + \dots + a_0 x_t + \dots + a_u x_{t+u}, \quad t = v+1, \dots, N-u, \quad (3.6)$$

hvor a -erne er konstanter, således at den nye tidsrække kan beskrives ved hjælp af en svagt stationær proces \tilde{X} .

I [Sætning 2.19](#) så vi, at der eksisterer en svagt stationær løsning X til ligningerne

$$\varphi_1(B)X_t = \theta(B)W_t, \quad t \in \mathbb{Z}, \quad (3.7)$$

hvor $\varphi_1(z)$ og $\theta(z)$ ikke har fælles rødder, hvis og kun hvis $\varphi_1(z) \neq 0$ for $|z| = 1$. En simpel måde at udvide klassen af ARMA-processer til at omfatte processer, som ikke er svagt stationære, er at tillade polynomiet $\varphi_1(z)$ at have rødder på enhedscirklen. Tillader vi, at $z = 1$ er en rod i $\varphi_1(z)$ af multiplicitet d , bliver formel (3.7) til

$$\varphi(B)\nabla^d X_t = \theta(B)W_t, \quad t \in \mathbb{Z}, \quad (3.8)$$

hvor $\varphi(z) \neq 0$ for $|z| = 1$ og hvor $\nabla = 1 - B$. En process med denne egenskab, det vil sige, at 1 er rod i AR-polynomiet af multiplicitet d , kaldes en $I(d)$ -proces (integreret af orden d).

Da $\nabla X_t = X_t - X_{t-1}$, $\nabla^2 X_t = X_t - 2X_{t-1} + X_{t-2}$ osv. er tidsrækken $\tilde{x}_t = \nabla^d x_t$ et eksempel på konstruktionen i (3.6). Dette motiverer følgende definition.

Definition 3.1. Lad $X = \{X_t\}_{t \in \mathbb{Z}}$ være en proces så $\{\nabla^d X_t + \mu\}_{t \in \mathbb{Z}}$ er en ARMA(p, q)-proces. Processen X kaldes da en **ARIMA-proces** af orden (p, d, q) . Parametrene for ARIMA-processen er parametrene for processen $\{\nabla^d X_t + \mu\}_{t \in \mathbb{Z}}$, dvs. $\mu, \sigma_W^2, \varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q$.

(ARIMA = *autoregressive integrated moving average*.)

Bemærk, at en ARIMA-proces X per definition opfylder (3.8).

Analysen af en empirisk tidsrække ved hjælp af ARIMA-modellen

$$\varphi(B)\nabla^d X_t = \theta(B)W_t, \quad t \in \mathbb{Z},$$

hvor $p, d, q \in \mathbb{N}_0$, $\varphi(z) = 1 - \varphi_1 z - \dots - \varphi_p z^p \neq 0$ hvis $|z| \leq 1$, og $\theta(z) = 1 - \theta_1 z - \dots - \theta_q z^q \neq 0$ hvis $|z| \leq 1$, opdeles på en naturlig måde i tre trin:

- (a) identifikation
- (b) estimation
- (c) modelkontrol.

Ved identifikation forstås bestemmelse af d efterfulgt af bestemmelse af en relevant model for $\nabla^d x_{d+1}, \dots, \nabla^d x_N$ fra klassen $\mathcal{A}_0(\mathbb{R})$ af ARMA-modeller med reelle koefficienter, herunder bestemmelse af ordenen (p, q) . Herefter estimeres parametrene $\mu, \sigma_W^2, \varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q$ for den identificerede model. Til sidst undersøges det, om den estimerede model giver en passende god beskrivelse af data.

Statistiske modeller for uafhængige stokastiske variable baseret på normalfordelingen betragtes som den vigtigste klasse af modeller dels på grund af, at

modellerne matematisk er rimeligt lette at håndtere, og dels på grund af deres store anvendelighed. Da en lignende bemærkning også gælder for modeller for empiriske tidsrækker, *vil vi i det følgende altid antage at den hvide støj W er Gaussisk.*

Identifikation

ARMA-modeller Identifikationen af den relevante ARMA(p, q)-proces fra klassen $\mathcal{A}_0(\mathbb{R})$ for en empirisk tidsrække x_1, \dots, x_N udføres ved hjælp af den empiriske autokorrelationsfunktion $\{r_k\}_{k=0}^K$ og den empiriske partielle autokorrelationsfunktion $\{\hat{\phi}_{k,k}\}_{k=1}^K$, som bestemmes rekursivt ud fra den førstnævnte på samme måde som de teoretiske partielle autokorrelationer $\{\phi_{k,k}\}$ beregnes ud fra de teoretiske autokorrelationer $\{\rho_k\}$. Ved hjælp af resultaterne for de teoretiske korrelationsfunktioner for en ARMA(p, q)-proces i [Sætning 2.21](#) kan vi bestemme p og q ved hjælp af de empiriske autokorrelationsfunktioner, så derfor en kort oversigt over disse resultater for ARMA(p, q)-processen X ,

$$\varphi(B)X_t = \theta(B)W_t, \quad t \in \mathbb{Z}. \quad (3.9)$$

(a) Autokorrelationsfunktionen $\{\rho_k\}_{k \in \mathbb{Z}}$ for X opfylder differensligningen

$$\varphi(B)\rho_k = 0, \quad k > q.$$

Hvis $p \neq 0$, er autokorrelationsfunktionen en sum af eksponentielt aftagende led og eksponentielt dæmpede sinussvingninger.

Hvis $p = 0$, gælder der

$$X \sim MA(q) \Rightarrow \rho_k = 0, \quad k > q.$$

(b) Den partielle autokorrelationsfunktion $\{\phi_{k,k}\}_{k \in \mathbb{Z}}$ for X opfylder “næsten” differensligningen

$$\theta(B)\phi_{k,k} = 0, \quad k > p + q.$$

Hvis $q \neq 0$, er den partielle autokorrelationsfunktion “næsten” en sum af eksponentielt aftagende led og eksponentielt dæmpede sinussvingninger.

Hvis $q = 0$, gælder der

$$X \sim AR(p) \Rightarrow \phi_{k,k} = 0, \quad k > p.$$

Det er derfor kun rimeligt at modellere den empiriske tidsrække x_1, \dots, x_N ved hjælp af (3.9), hvis de empiriske autokorrelationsfunktioner begge går mod 0 når

$k \rightarrow \infty$. Hvis dette ikke er tilfældet, beregnes de empiriske autokorrelationsfunktioner for rækken $\nabla x_2, \dots, \nabla x_N$, der fremkommer af x_1, \dots, x_N ved at tage differenser. Hvis $\nabla x_2, \dots, \nabla x_N$ kan beskrives ved hjælp af en model i klassen $\mathcal{A}_0(\mathbb{R})$, er $d = 1$, så processen er $I(1)$. Ellers tages differenser af $\nabla x_2, \dots, \nabla x_N$ og så videre. Til hjælp ved bestemmelsen af d er Dickey-Fuller testet.

Bemærkning 3.2. *Dickey-Fuller testet* er et test for stationaritet af en tidsrække. Testet vedrører AR(p)-processen $\varphi(B)x_t = w_t$ eller

$$x_t = \varphi_1 x_{t-1} + \dots + \varphi_p x_{t-p} + w_t, \quad (3.10)$$

og tester hypotesen, at $z = 1$ er rod i polynomiet $\varphi(z)$, dvs. at processen *ikke* er stationær.

(Hvis $p > 1$ omtales testet ofte som det udvidede (augmented) Dickey-Fuller test.)

Testet fremkommer ved hjælp af følgende omskrivning af (3.10)

$$\nabla x_t = -\varphi(1)x_{t-1} + \gamma_1 \nabla x_{t-1} + \dots + \gamma_{p-1} \nabla x_{t-(p-1)} + w_t, \quad (3.11)$$

hvor

$$\gamma_i = -\sum_{j=i+1}^p \varphi_j, \quad j = 1, \dots, p-1,$$

ved at betragte den multiple regression af ∇x_t på $x_{t-1}, \nabla x_{t-1}, \dots, \nabla x_{t-(p-1)}$ og ved i denne at teste om koefficienten til x_{t-1} er 0, dvs. om $\varphi(1) = 0$.

Testet afhænger naturligvis af ordenen p af AR(p)-processen. Varianter af det nævnte test fremkommer ved at tilføje en konstant α eller et lineært led $\alpha + \beta t$ på højresiden i formel (3.10).

Testet bruges ikke blot i forbindelse med AR(p)-processer men også for generelle processer, idet det kan vises, at disse kan approksimeres ved hjælp af en AR-proces af høj orden. ■

I det følgende antages, at $d = 0$. (Ellers, betragtes $\nabla^d x_{d+1}, \dots, \nabla^d x_N$ som den observerede tidsrække.)

For en MA(q)-proces gælder der, at $\rho_j = 0$ hvis $j > q$, og Bartlett's formler (3.3) og (3.5) bliver derfor

$$Er_k \approx 0, \quad k > q,$$

og

$$\text{Var } r_k = \frac{1}{N} \left(1 + 2 \sum_{j=1}^q \rho_j^2 \right), \quad k > q. \quad (3.12)$$

For store værdier af N kan fordelingen af r_k for $k > q$ derfor approksimeres med en normalfordeling med middelværdi 0 og en varians givet ved (3.12). Ved at erstatte ρ_k i formel (3.12) med estimatet r_k , fås at

$$\pm \frac{1.96}{\sqrt{N}} \left(1 + 2 \sum_{j=1}^q r_j^2 \right)^{\frac{1}{2}}$$

er grænserne for et approksimativt 95 %-konfidensinterval for ρ_k for $k > q$. Det vil sige, at ved hjælp af den empiriske autokorrelationsfunktion kan antagelsen om en MA(q)-model vurderes.

I tegningerne af den empiriske autokorrelationsfunktion nedenfor er der vist nogle approksimative 95 %-konfidenskurver. Værdien af kurverne ved lag k antyder det approksimative 95 %-konfidensinterval for ρ_k beregnet under antagelse af, at modellen er en MA($k-1$)-model, dvs.

$$\pm \text{conf}_{95}(k) = \pm \frac{1.96}{\sqrt{N}} \left(1 + 2 \sum_{j=1}^{k-1} r_j^2 \right)^{\frac{1}{2}}. \quad (3.13)$$

Ved hjælp af disse kurver kan man få approksimative 95 %-konfidensintervaller for autokorrelationerne for en vilkårlig MA(q)-model. Konfidensintervallerne for en sådan model er de parallelle linjer

$$(j, \pm \text{conf}_{95}(q+1)), \quad j = q+1, q+2, \dots \quad (3.14)$$

For en AR(p)-proces gælder der, at $\phi_{k,k} = 0$ hvis $k \geq p+1$, og [Quenouille \(1949\)](#) har vist, at for sådanne processer er de empiriske partielle autokorrelationer for $k \geq p+1$ approksimativt uafhængige med middelværdi 0 og varians $1/N$,

$$\text{Var } \hat{\phi}_{k,k} \approx \frac{1}{N}, \quad k \geq p+1,$$

for $N \rightarrow \infty$. Fra dette kan vi få approksimative 95 %-konfidensintervaller for $\phi_{k,k}$, så rimeligheden af en AR(p)-model kan vurderes ved hjælp af den empiriske partielle autokorrelationsfunktion.

I tegninger af den empiriske autokorrelationsfunktion er de approksimative 95 %-konfidensintervaller antydnet af linjerne

$$(j, \pm 1.96/\sqrt{N}), \quad j = 1, 2, \dots$$

Ved vurderingen af de empiriske autokorrelationsfunktioner bør man tage Bartlett's formel (3.4) i betragtning, idet den viser, at de estimerede autokorrelationer selv kan være stærkt korrelerede. De empiriske autokorrelationer afspejler kun i grove træk den teoretiske autokorrelationsfunktion. En lignende bemærkning gælder for den empiriske partielle autokorrelationsfunktion. Bemærk dog Quenouille's resultat for AR(p)-processer.

SAS tilbyder forskellige metoder til en foreløbig bestemmelse af ordenen af en ARMA(p, q)-proces.

Bemærkning 3.3. Vi omtaler nu *MINIC metoden* (**m**inimum **i**nformation **c**riterion), som ud fra observationerne x_1, \dots, x_N giver en foreløbig bestemmelse af ordenen af ARMA-processen

$$\varphi(B)(X_t - \mu) = \theta(B)W_t, \quad (3.15)$$

hvor $\varphi(z) \neq 0$ og $\theta(z) \neq 0$ for $|z| \leq 1$.

Vi antager, at $\mu = 0$, ellers betragtes tidsrækken $\{x_t - \bar{x}\}_{t=1}^N$.

Fra [Sætning 2.19](#) på [side 19](#) ved vi, at processen kan repræsenteres som en (baglæns) autoregression, og den hvide støj kan derfor approksimeres ved hjælp af en autoregression af høj orden \tilde{p} , dvs.

$$W_t = \varphi_{\tilde{p}}(B)X_t, \quad (3.16)$$

hvor indekset angiver ordenen.

Første trin i metoden består i at finde en passende værdi p_W af \tilde{p} . For $\tilde{p}_{\min} \leq \tilde{p} \leq \tilde{p}_{\max}$ betragtes residualerne

$$\vec{w}_{\tilde{p}t} = \vec{\varphi}_{\tilde{p}}(B)x_t$$

fra Yule-Walker estimation i (3.16), se formel (3.35) på [side 69](#), og

$$s_{\tilde{p}}^2 = \frac{1}{N} \sum_{t=\tilde{p}+1}^N \vec{w}_{\tilde{p}t}^2,$$

som er et estimat for variansen på den hvide støj baseret på Yule-Walker residualerne.

Værdien p_W vælges som den værdi af \tilde{p} , som minimerer Akaike information criterion (AIC)

$$\text{AIC}(\tilde{p}, 0) = \ln(s_{\tilde{p}}^2) + 2(\tilde{p} + 0)/N.$$

For at lette notationen betegnes $\vec{w}_{p_W t}$ i det følgende blot som \vec{w}_t .

I det næste trin i metoden ser vi på test værdier $m = p_{\min}, \dots, p_{\max}$ for graden af $\varphi(z)$ og test værdier $j = q_{\min}, \dots, q_{\max}$ for graden af $\theta(z)$. For fast værdi af (m, j) betragtes

$$x_t = \varphi_1^{(m)} x_{t-1} + \dots + \varphi_m^{(m)} x_{t-m} + \vec{w}_t - \theta_1^{(j)} \vec{w}_{t-1} - \dots - \theta_j^{(j)} \vec{w}_{t-j} + u_t, \quad (3.17)$$

som bortset fra et fejld u_t med middelværdi 0 og varians σ_u^2 er ligningen i (3.15) for en proces af grad (m, j) (med $\mu = 0$) og med w_t erstattet med \vec{w}_t .

Ligningen (3.17) kan betragtes som en multipel regression med x_t som respons og med $x_{t-1}, \dots, x_{t-m}, \vec{w}_t, \dots, \vec{w}_{t-j}$ som forklarende variable. Betragt OLS estimation i denne regression og lad \hat{u}_t være OLS residualet, dvs.

$$\hat{u}_t = x_t - \hat{\phi}_1^{(m)} x_{t-1} - \dots - \hat{\phi}_m^{(m)} x_{t-m} - \vec{w}_t + \hat{\theta}_1^{(j)} \vec{w}_{t-1} + \dots + \hat{\theta}_j^{(j)} \vec{w}_{t-j},$$

som kan beregnes for $t \geq t_0 = p_W + \max(m, j)$. Et estimat for σ_u^2 er

$$s_{(m,j)}^2 = \frac{1}{N} \sum_{t=t_0}^N \hat{u}_t^2$$

Som en foreløbig værdi af ordenen af ARMA-processen (3.15) vælger SAS proceduren PROC ARIMA som default den værdi af (m, j) , som minimerer Bayesian information criterion (BIC)

$$\text{BIC}(m, j) = \ln(s_{(m,j)}^2) + 2(m + j) \ln(N)/N \quad (3.18)$$

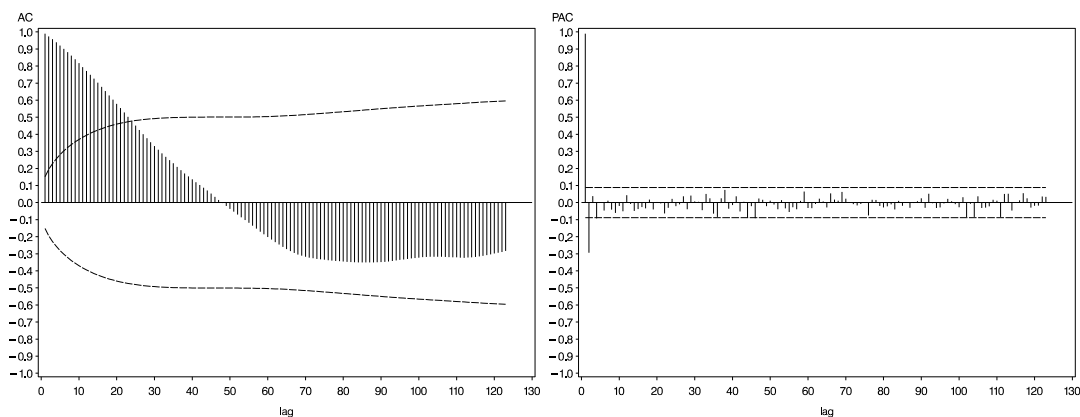
for $m = p_{\min}, \dots, p_{\max}$ og $j = q_{\min}, \dots, q_{\max}$, mens proceduren PROC VARMAX bruger AIC-kriteriet

$$\text{AIC}(m, j) = \ln(s_{(m,j)}^2) + 2(m + j)/N \quad (3.19)$$

som default. ■

Eksempel 3.4. Af Figur 1.1 på side 2 fremgår det tydeligt, at kursen ikke kan beskrives ved en stationær tidsrække, idet der for eksempel er flere niveauskift i forløbet af prisen på US\$ i den betragtede periode.

At rækken ikke er stationær, bekræftes ved hjælp af de empiriske autokorrelationer (AC) i Figur 3.1 som aftager meget langsomt.



Figur 3.1: De empiriske autokorrelationer (AC) og partielle autokorrelationer (PAC) for kursen på US\$

Disse indtryk forstærkes ved Dickey-Fuller testene i SAS-udskrift 3.1.

SAS-udskrift 3.1

Augmented Dickey-Fuller Unit Root Tests							
Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	1	-0.5527	0.5586	-0.69	0.4183		
	2	-0.5166	0.5663	-0.69	0.4196		
	3	-0.5729	0.5543	-0.69	0.4178		
	4	-0.5648	0.5560	-0.69	0.4174		
	5	-0.5812	0.5525	-0.69	0.4184		
Single Mean	1	-8.6264	0.1859	-2.07	0.2569	2.19	0.5113
	2	-7.6785	0.2332	-1.94	0.3131	1.93	0.5757
	3	-9.5302	0.1495	-2.15	0.2246	2.35	0.4681
	4	-9.3814	0.1550	-2.12	0.2357	2.30	0.4832
	5	-10.2157	0.1267	-2.20	0.2069	2.46	0.4422
Trend	1	-9.0114	0.5043	-2.11	0.5369	2.23	0.7300
	2	-8.0284	0.5796	-1.98	0.6091	1.97	0.7836
	3	-9.9409	0.4380	-2.19	0.4919	2.41	0.6953
	4	-9.7814	0.4489	-2.16	0.5090	2.34	0.7086
	5	-10.6416	0.3911	-2.24	0.4664	2.51	0.6750

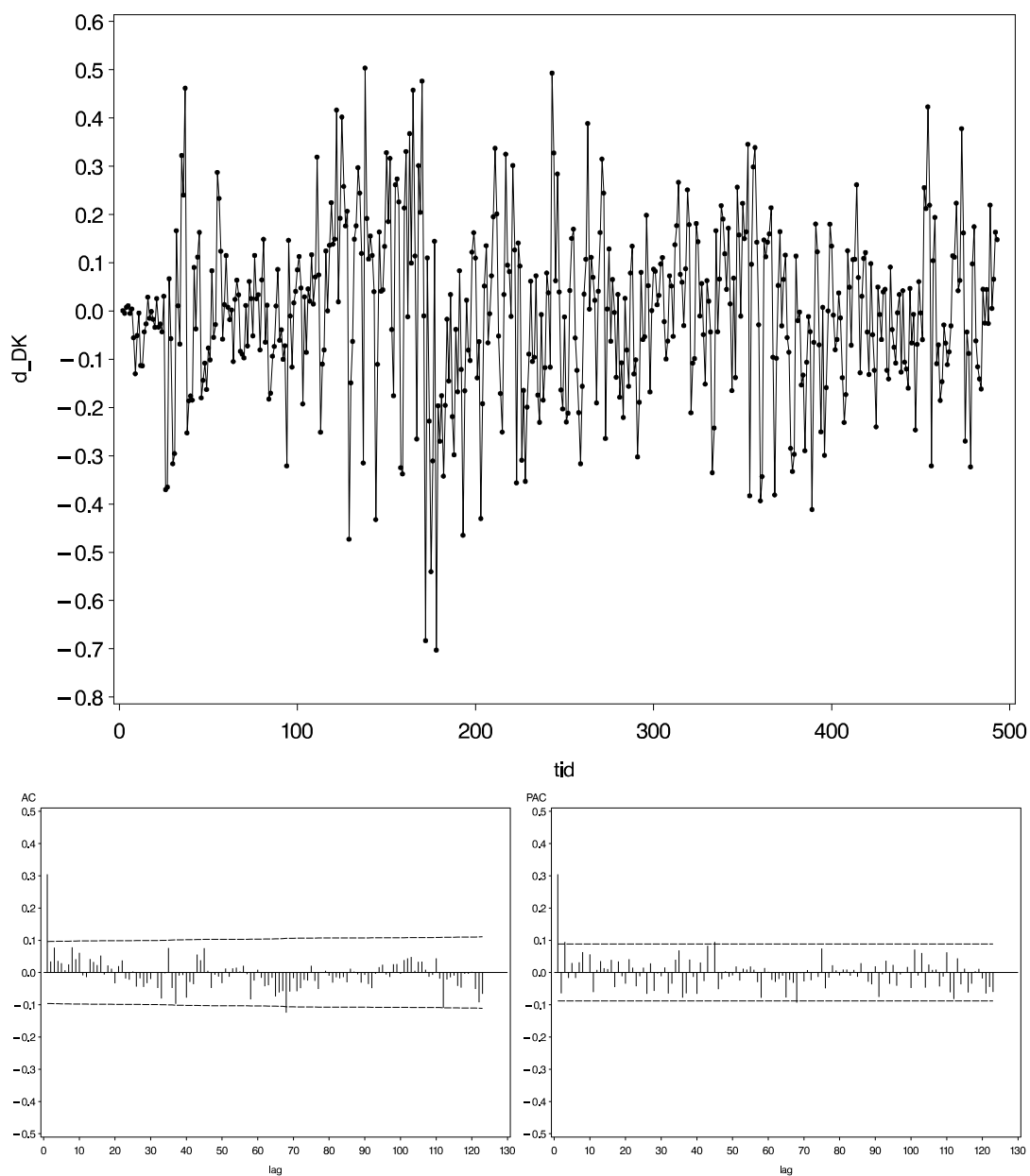
SAS beregner tre teststørrelser, Rho, Tau og F, for hypotesen om en enhedsrod, altså om at 1 er rod i AR-polynomiet, det vil sige, at tidsrækken **ikke** er stationær. De tre typer Zero Mean, Single Mean og Trend svarer til de tre varianter af testet omtalt i [Bemærkning 3.2](#). For illustrationens skyld er testet vist for $p(\text{Lags}) = 1, \dots, 5$. Ingen af testene forkaster hypotesen, altså er rækken ikke stationær.

Vi betragter derfor differenserne af kurserne. Disse er vist i [Figur 3.2](#) sammen med de tilsvarende empiriske korrelationsfunktioner. Differenser varierer omkring niveauet 0, måske bortset fra, at variansen i nogle perioder er lille i forhold til andre perioder. Dette vælger vi her at se bort fra. Hverken de empiriske autokorrelationer i [Figur 3.2](#) eller Dickey-Fuller testene i [SAS-udskrift 3.2](#) nedenfor betvivler stationariteten af differenserne, idet hypotesen om en enhedsrod forkastes.

SAS-udskrift 3.2

Augmented Dickey-Fuller Unit Root Tests							
Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	1	-386.924	0.0001	-13.86	<.0001		
	2	-298.230	0.0001	-10.63	<.0001		
Single Mean	1	-387.269	0.0001	-13.85	<.0001	95.90	0.0010
	2	-298.661	0.0001	-10.63	<.0001	56.46	0.0010
Trend	1	-387.274	0.0001	-13.83	<.0001	95.70	0.0010
	2	-298.657	0.0001	-10.61	<.0001	56.35	0.0010

Endvidere antyder AC i [Figur 3.2](#), at en MA(1)-model kan betragtes, idet AC-funktionen har et cutoff ved lag 1. I [SAS-udskrift 3.3](#) ses, at option MINIC anbefaler samme model.

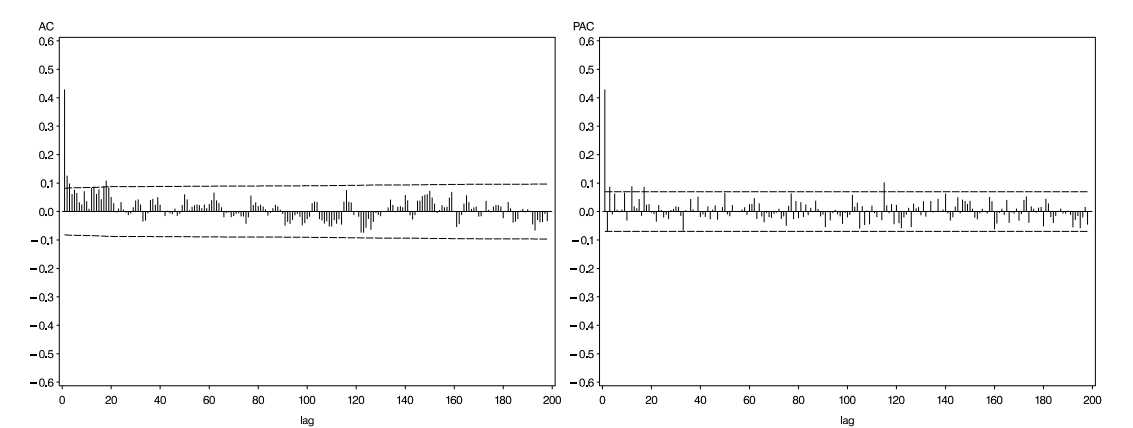


Figur 3.2: Differenserne d_DK af kursen på US\$ i [Figur 1.1](#) på [side 2](#) er vist øverst og de tilsvarende empiriske korrelationsfunktioner nederst.

SAS-udskrift 3.3

Minimum Information Criterion						
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	-3.43946	-3.53170	-3.51941	-3.51312	-3.50131	-3.48935
AR 1	-3.52466	-3.51961	-3.50943	-3.50235	-3.49001	-3.47742
AR 2	-3.51642	-3.50920	-3.50219	-3.49064	-3.47880	-3.47002
AR 3	-3.51375	-3.50182	-3.49048	-3.48026	-3.46776	-3.45772
AR 4	-3.50144	-3.48975	-3.47796	-3.46774	-3.45519	-3.44558
AR 5	-3.48976	-3.47747	-3.46593	-3.45737	-3.44477	-3.43533

Eksempel 3.5. Figur 3.3 antyder, at isotopkurven i Figur 1.2 på side 3 kan modelleres ved hjælp af en stationær tidsrække, idet de empiriske korrelationer dør hurtigt ud. Dickey-Fuller testet i SAS-udskrift 3.4 bekræfter dette, idet det i rækken Single Mean ses, at hypotesen om en enhedsrod forkastes.



Figur 3.3: De empiriske autokorrelationer (AC) og partielle autokorrelationer (PAC) for isotopkurven i Figur 1.2.

SAS-udskrift 3.4

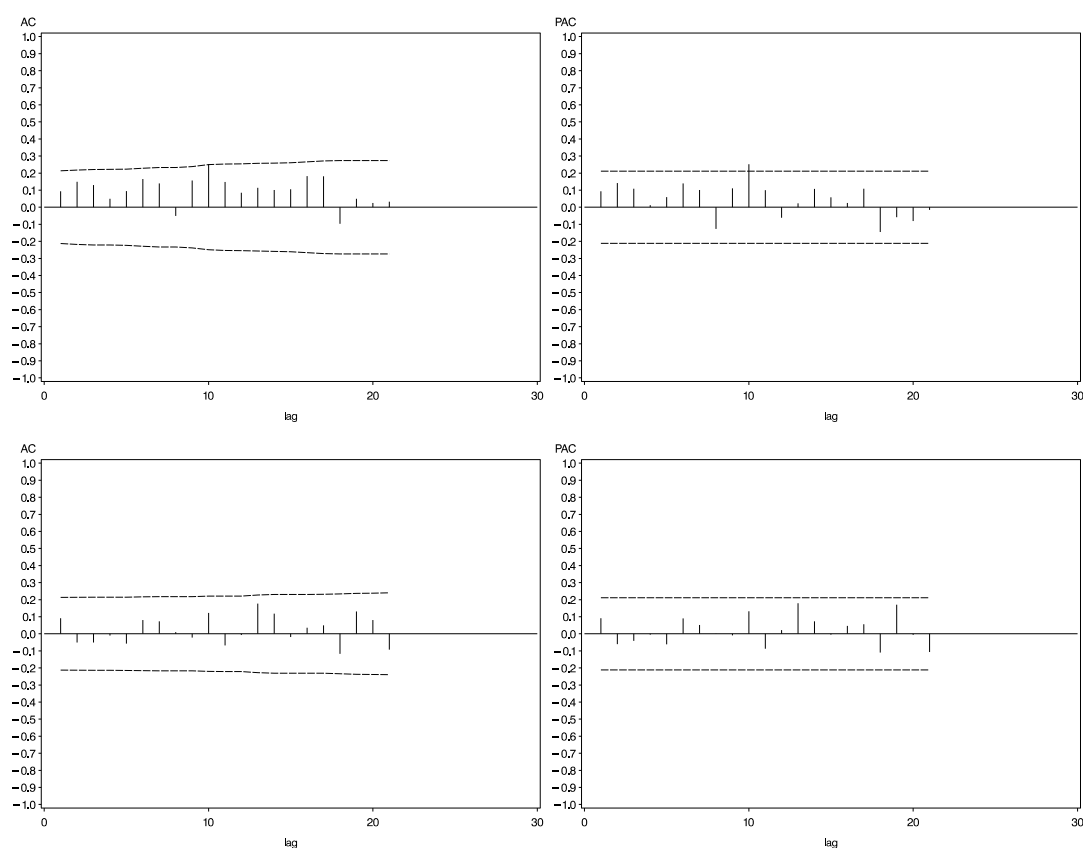
Augmented Dickey-Fuller Unit Root Tests						
Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F Pr > F
Zero Mean	1	-0.4161	0.5885	-0.44	0.5218	
Single Mean	1	-519.788	0.0001	-16.07	<.0001	129.13 0.0010
Trend	1	-528.168	0.0001	-16.18	<.0001	130.84 0.0010

I SAS-udskrift 3.5 anbefaler MINIC en AR(1)-model, mens den empiriske partielle autokorrelationsfunktion i Figur 3.3 synes at foreslå en AR(3)-model, da $\hat{\phi}_{3,3}$ falder uden for 95 %-konfidensgrænsen.

SAS-udskrift 3.5

Minimum Information Criterion						
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	4.888199	4.702496	4.700553	4.701404	4.70822	4.713134
AR 1	4.689700	4.691222	4.693788	4.701559	4.709234	4.716129
AR 2	4.693626	4.698396	4.700650	4.707253	4.714623	4.722926
AR 3	4.693508	4.701592	4.708985	4.715636	4.714063	4.722304
AR 4	4.701560	4.709639	4.716633	4.722471	4.722342	4.730720
AR 5	4.706006	4.714294	4.722697	4.726017	4.729959	4.738190
AR 5	4.706006	4.714294	4.722697	4.726017	4.729959	4.738190

Eksempel 3.6. Figur 3.4 viser de empiriske korrelationsfunktioner for 86 målinger af nedbør og vandmængde i Figur 1.3 på side 4. Tegningerne i Figur 3.4 antyder, at begge rækker er stationære og kan beskrives ved en hvid støj, ARMA(0,0), idet de empiriske korrelationer med en enkelt undtagelse alle ligger indenfor 95 %-konfidensintervallet.



Figur 3.4: De empiriske korrelationsfunktioner nedbøren i Figur 1.3 på side 4 ses øverst, mens de tilsvarende størrelser for vandmængden er vist nederst.

Antagelsen om stationaritet bekræftes af [SAS-udskrift 3.6](#), hvor vi i linjerne Single Mean kan aflæse, at for begge de variable forkastes hypotesen om en enhedsrod, dvs. tidsrækkerne er stationære.

SAS-udskrift 3.6

Dickey-Fuller Unit Root Tests					
Variable	Type	Rho	Pr < Rho	Tau	Pr < Tau
nedboer	Zero Mean	-1.55	0.3854	-0.92	0.3156
	Single Mean	-56.34	0.0008	-5.12	0.0001
	Trend	-95.73	0.0003	-6.36	<.0001
vandmaengde	Zero Mean	-1.21	0.4364	-0.76	0.3851
	Single Mean	-86.61	0.0008	-6.47	<.0001
	Trend	-105.26	0.0001	-7.01	<.0001

Option MINIC giver lidt overraskende, at et første forslag til en model for nedbøren er en AR(2)-model, som det ses i [SAS-udskrift 3.7](#), mens [SAS-udskrift 3.8](#) viser, at forslaget for vandmængden er en hvid støj, som antydnet nederst i [Figur 3.4](#).

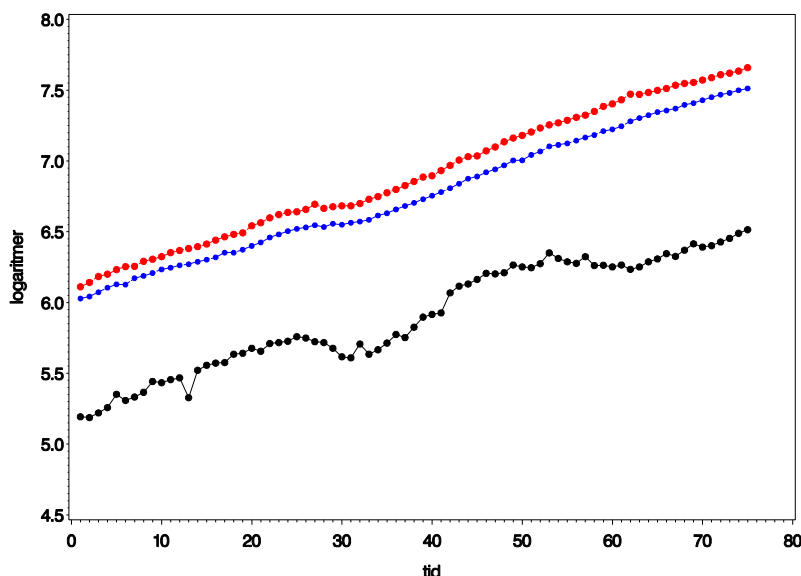
SAS-udskrift 3.7

Minimum Information Criterion						
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	8.845124	8.884180	8.788823	8.835820	8.856114	8.905490
AR 1	8.869513	8.885117	8.791813	8.841657	8.851935	8.902482
AR 2	8.784624	8.801710	8.836792	8.885736	8.899902	8.949653
AR 3	8.825128	8.850924	8.883067	8.892008	8.894871	8.942297
AR 4	8.852786	8.876393	8.911148	8.904123	8.924705	8.970015
AR 5	8.892258	8.921249	8.949653	8.951348	8.974067	8.987386

SAS-udskrift 3.8

Minimum Information Criterion						
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	10.86599	10.90670	10.95463	10.99408	11.03952	11.07292
AR 1	10.90843	10.95484	11.00575	11.04197	11.08460	11.11553
AR 2	10.95459	11.00536	11.05583	11.09331	11.13640	11.16653
AR 3	10.99476	11.04116	11.09235	11.14354	11.18754	11.21831
AR 4	11.04000	11.08390	11.13555	11.18557	11.23492	11.26974
AR 5	11.07324	11.11437	11.16499	11.21675	11.26848	11.29926

Eksempel 3.7. Figur 1.4 på side 5 viser, at hver af de tre tidsrækker for de kvartalsvise opgørelser af henholdsvis faste investeringer, disponibel indkomst og udgifter til forbrug i Vesttyskland i perioden 1960–1978 er eksponentiel voksende. Vi betragter derfor de logaritmetransformerede tidsrækker i Figur 3.5.



Figur 3.5: Logaritmerne til de variable faste investeringer (sort), disponibel indkomst (rød) og udgifter til forbrug (blå) i Vesttyskland kvartalsvis i perioden 1960–1978.

Logaritmerne varierer nærmest lineært, og for at opnå stationære tidsrækker betragter vi differenserne af logaritmerne. Disse er vist i Figur 3.6, og de forekommer at være stationære.

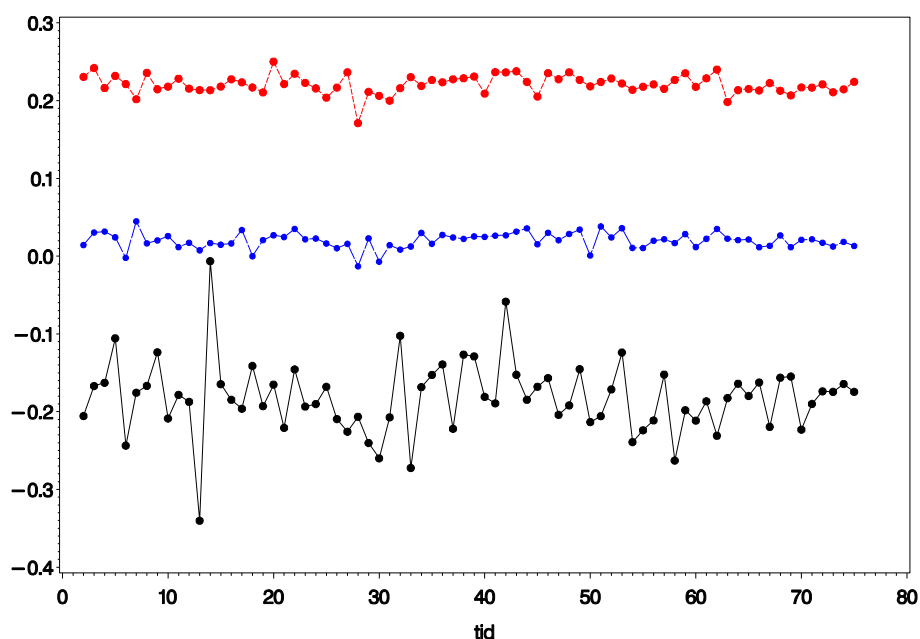
Stationariteten af de tre tidsrækker i Figur 3.6 bekræftes stort set af Dickey-Fuller testene i SAS-udskrift 3.9. Her betegner `d1l_inv` førstedifferensen for logaritmen til investeringer osv. Da det forekommer rimeligt, at førstedifferenserne har middelværdi 0, skal vi se i rækkerne `Zero Mean` i SAS-udskrift 3.9. For den variable `d1l_cons` forkastes hypotesen om en enhedsrod ikke. Bedømt ud fra de empiriske korrelationsfunktioner i Figur 3.7 på side 57 ser rækkerne dog stationære ud.

SAS-udskrift 3.9

Dickey-Fuller Unit Root Tests					
Variable	Type	Rho	Pr < Rho	Tau	Pr < Tau
d1l_inv	Zero Mean	-63.18	<.0001	-5.52	<.0001
	Single Mean	-102.89	0.0001	-7.02	0.0001
	Trend	-104.32	0.0001	-7.01	<.0001
d1l_inc	Zero Mean	-7.66	0.0523	-2.14	0.0319
	Single Mean	-59.64	0.0007	-5.49	0.0001
	Trend	-60.12	0.0002	-5.46	0.0001

d11_cons	Zero Mean	-4.40	0.1460	-1.57	0.1089
	Single Mean	-44.15	0.0007	-4.50	0.0005
	Trend	-44.12	0.0002	-4.47	0.0033

SAS-udskrifterne 3.10–3.12 antyder, at de tre rækker af førstedifferenser af logaritmerne kan analyseres med henholdsvis en en hvid støj, en hvid støj og en AR(2)-model. Figur 3.7 peger dog på en AR(4)-model for den første tidsrække.



Figur 3.6: Differenser af logaritmerne til de variable faste investeringer (sort, -0.2), disponibel indkomst (rød, $+0.2$) og udgifter til forbrug (blå) i Vesttyskland kvartalsvis i perioden 1960–1978.

SAS-udskrift 3.10

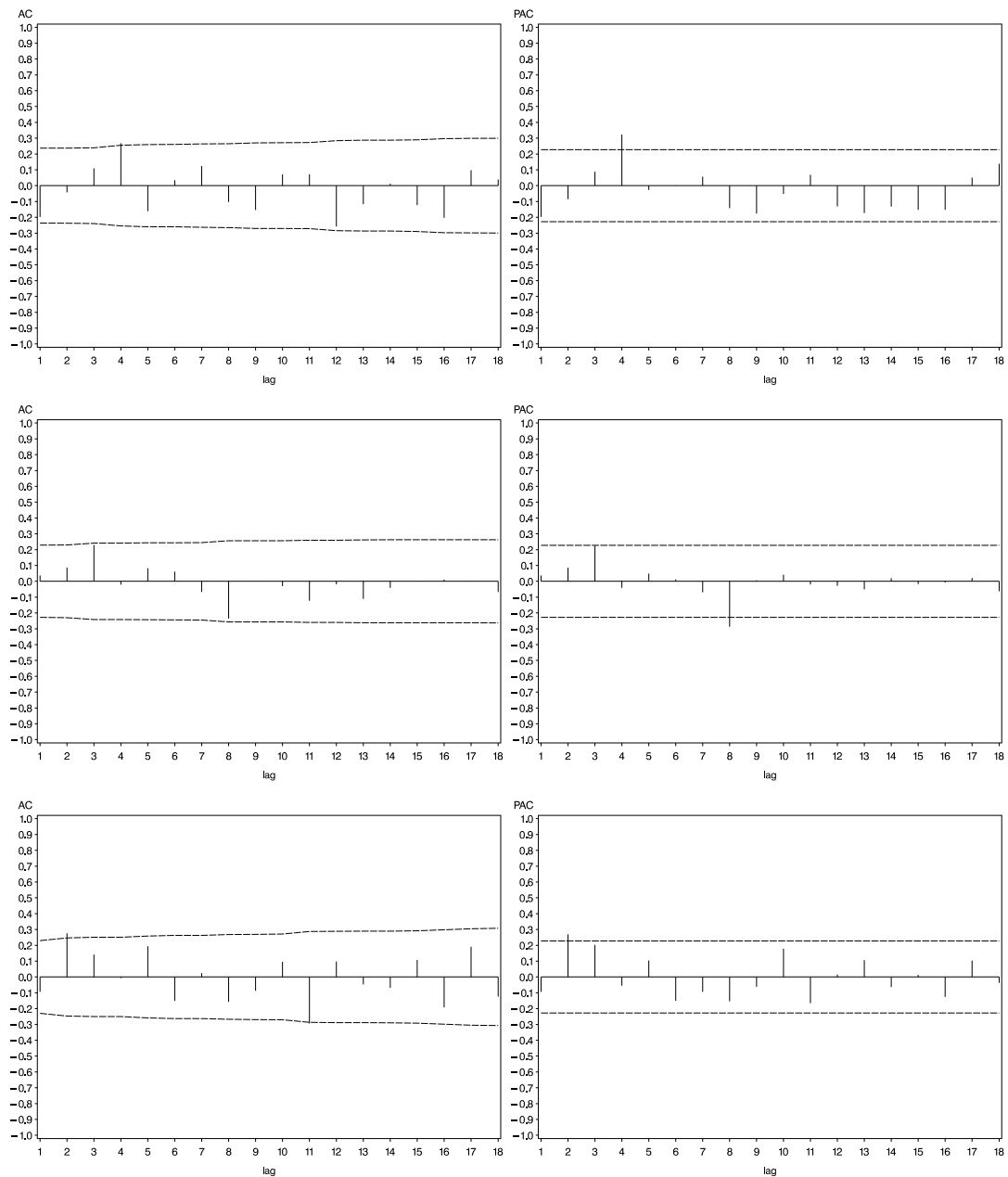
Minimum Information Criterion						
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	-6.20592	-6.18680	-6.13095	-6.12559	-6.12889	-6.08141
AR 1	-6.18370	-6.12967	-6.07358	-6.06943	-6.09779	-6.04168
AR 2	-6.13355	-6.07654	-6.03999	-6.02365	-6.05624	-6.04000
AR 3	-6.12164	-6.06513	-6.02122	-5.96623	-6.01434	-5.98323
AR 4	-6.16798	-6.11228	-6.05605	-6.00269	-5.95755	-5.93831
AR 5	-6.11312	-6.05620	-6.00204	-5.96182	-5.94028	-5.88506

SAS-udskrift 3.11

Minimum Information Criterion						
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	-8.97570	-8.91958	-8.87762	-8.86484	-8.81720	-8.79005
AR 1	-8.92332	-8.87390	-8.83414	-8.81300	-8.76257	-8.73605
AR 2	-8.88256	-8.83402	-8.77686	-8.75549	-8.70505	-8.69306
AR 3	-8.87093	-8.81518	-8.75920	-8.70185	-8.65012	-8.63585
AR 4	-8.81933	-8.76264	-8.70706	-8.65019	-8.59278	-8.57845
AR 5	-8.77099	-8.71459	-8.66823	-8.61376	-8.55622	-8.53063

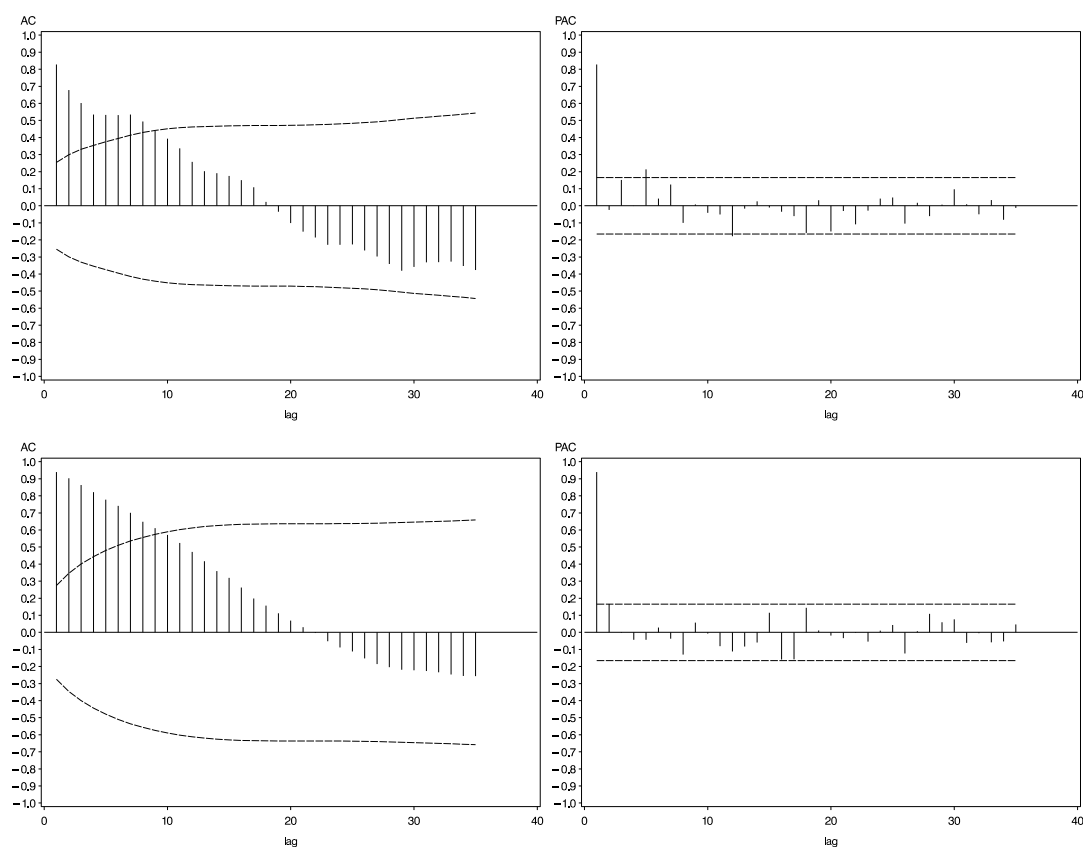
SAS-udskrift 3.12

Minimum Information Criterion						
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	-9.33907	-9.28327	-9.33949	-9.30532	-9.26389	-9.25362
AR 1	-9.28318	-9.22667	-9.28226	-9.25884	-9.22962	-9.23891
AR 2	-9.35637	-9.29977	-9.24276	-9.23173	-9.18638	-9.18343
AR 3	-9.34558	-9.30722	-9.25214	-9.19461	-9.15022	-9.12660
AR 4	-9.31717	-9.26118	-9.23760	-9.18901	-9.19012	-9.16918
AR 5	-9.28431	-9.22680	-9.18992	-9.15090	-9.15186	-9.15685



Figur 3.7: Figuren viser de empiriske korrelationsfunktioner svarende til førstedifferenserne af logaritmerne i [Figur 3.6](#) på side 55. Øverst for faste investeringer, i midten for disponibel indkomst og nederst for udgifter til forbrug.

Eksempel 3.8. Tidsrækkerne i [Figur 1.5](#) på [side 5](#) er tydeligvis ikke stationære, hvilket også fremgår de empiriske korrelationsfunktioner i [Figur 3.8](#), der kun aftager meget langsomt.



Figur 3.8: De empiriske korrelationsfunktioner for renten på amerikanske statsobligationer i [Figur 1.5](#) på [side 5](#). Øverst for de 1-årige obligationer og nederst for de 10-årige.

Dickey-Fuller testet for de 1-årige obligationer i [SAS-udskrift 3.13](#) antyder lidt overraskende stationaritet (Single Mean).

SAS-udskrift 3.13

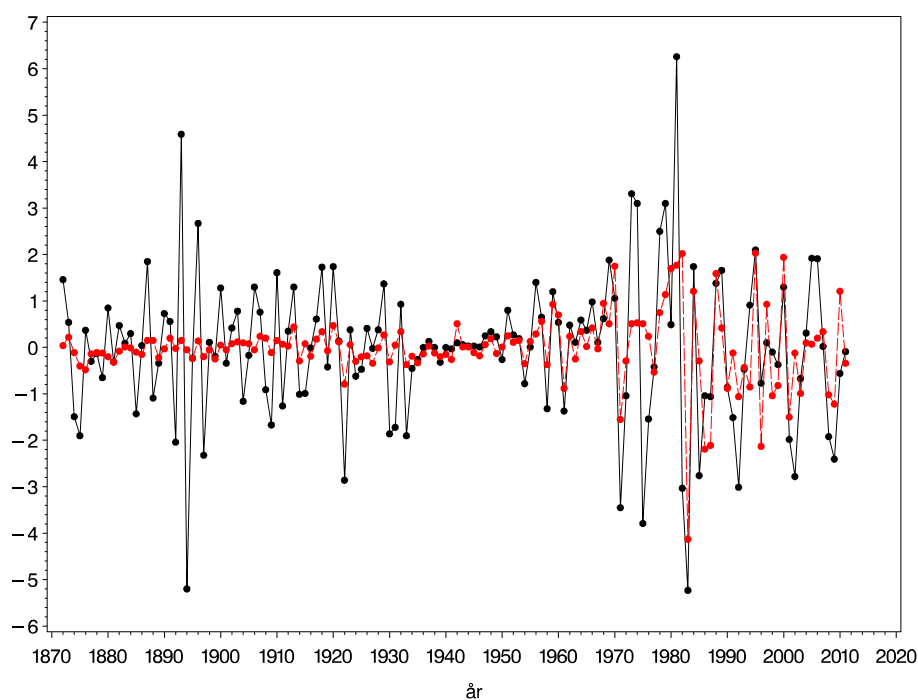
Dickey-Fuller Unit Root Tests					
Variable	Type	Rho	Pr < Rho	Tau	Pr < Tau
aar_1	Zero Mean	-6.25	0.0833	-1.93	0.0520
	Single Mean	-22.69	0.0046	-3.27	0.0186
	Trend	-22.88	0.0314	-3.26	0.0775
aar_10	Zero Mean	-1.38	0.4117	-0.94	0.3080
	Single Mean	-5.74	0.3625	-1.66	0.4491
	Trend	-7.07	0.6497	-1.81	0.6977

Tager vi førstedifferenser, får vi tidsrækkerne vist i [Figur 3.9](#), hvor stationariteten igen ser tvivlsom ud, da variansen ikke er konstant hen igennem rækken.

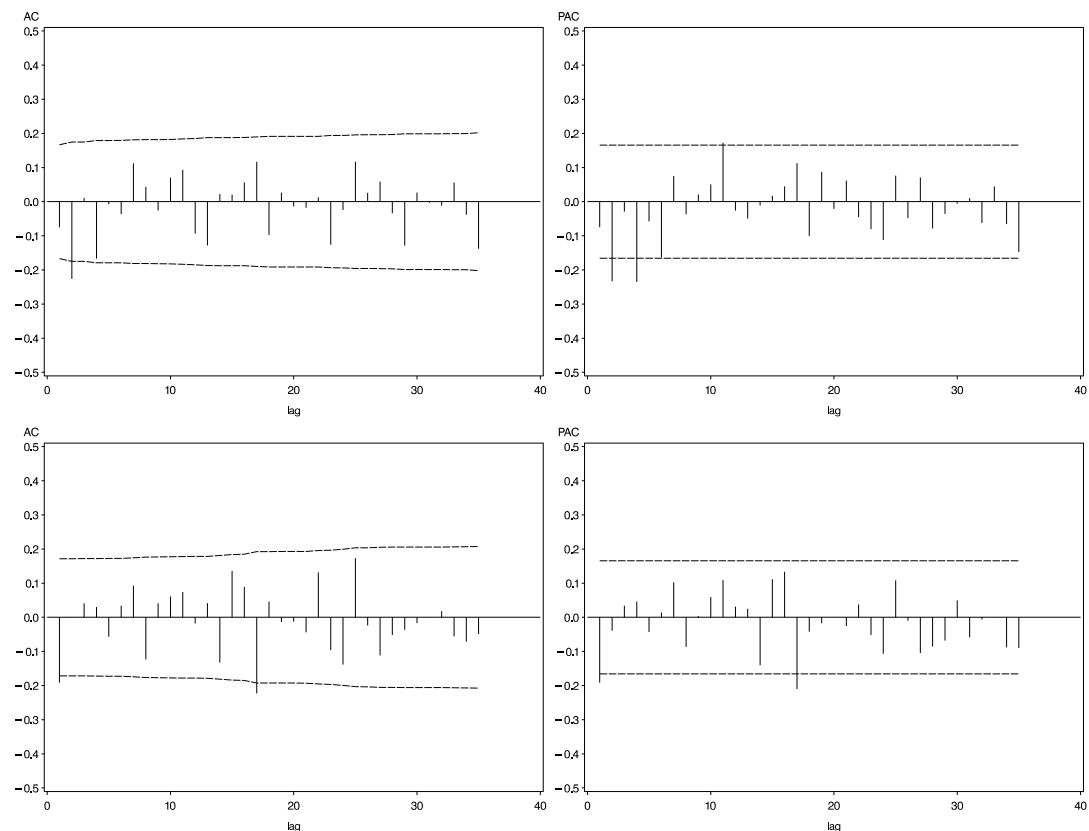
Dette ses ikke af Dickey-Fuller i [SAS-udskrift 3.14](#) eller af de empiriske korrelationsfunktioner i [Figur 3.10](#), som foreslår en MA(2) model for førstedifferenserne for de 1-årige obligationer og en MA(1) model for førstedifferenserne for de 10-årige obligationer. Disse valg bekræftes af MINIC i [SAS-udskrifterne 3.15](#) og [3.16](#).

SAS-udskrift 3.14

Observation(s) eliminated by differencing					1
Dickey-Fuller Unit Root Tests					
Variable	Type	Rho	Pr < Rho	Tau	Pr < Tau
aar_1	Zero Mean	-238.34	0.0001	-10.87	<.0001
	Single Mean	-239.21	0.0001	-10.85	<.0001
	Trend	-239.32	0.0001	-10.81	<.0001
aar_10	Zero Mean	-177.66	0.0001	-9.33	<.0001
	Single Mean	-177.98	0.0001	-9.30	<.0001
	Trend	-178.05	0.0001	-9.27	<.0001



Figur 3.9: Førstedifferenser af renten på 1-årige (sort) og af renten på 10-årige (rød) amerikanske statsobligationer i årene 1871–2011.



Figur 3.10: De empiriske korrelationsfunktioner for førstedifferenserne af renten på amerikanske statsobligationer i [Figur 1.5](#) på [side 5](#). Øverst for de 1-årige obligationer og nederst for de 10-årige.

SAS-udskrift 3.15

Minimum Information Criterion						
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	0.904642	0.926947	0.868303	0.901585	0.901558	0.932422
AR 1	0.933156	0.958074	0.902751	0.936404	0.936644	0.964363
AR 2	0.912682	0.943999	0.911859	0.944785	0.971333	0.999347
AR 3	0.945296	0.977655	0.945510	0.974736	0.995675	1.028981
AR 4	0.915921	0.949902	0.965799	0.992269	1.027148	1.060190
AR 5	0.945522	0.979828	0.987800	1.018277	1.050129	1.068114

SAS-udskrift 3.16

Minimum Information Criterion						
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	-0.52852	-0.53300	-0.49811	-0.46441	-0.43004	-0.39851
AR 1	-0.53257	-0.49795	-0.46290	-0.42933	-0.39593	-0.36358
AR 2	-0.49918	-0.46474	-0.43316	-0.40455	-0.37478	-0.34687

AR 3	-0.46529	-0.43022	-0.40455	-0.36958	-0.34836	-0.31434
AR 4	-0.43270	-0.39924	-0.37060	-0.34853	-0.31324	-0.27923
AR 5	-0.40114	-0.36598	-0.34207	-0.31451	-0.27922	-0.24780

Estimation

I dette afnit diskuterer vi estimation i ARMA-modeller. Vi betragter tre estimationsmetoder, maksimum likelihood estimation, betinget mindste kvadraters metode, ubetinget mindste kvadraters metode, og desuden for AR-modeller en metode, der omtales som Yule-Walker estimation. Metoderne kan også anvendes i ARIMA-modeller ved at erstatte de oprindelige observationer x_1, \dots, x_N med $\nabla^d x_{d+1}, \dots, \nabla^d x_N$, der fremkommer fra x_1, \dots, x_N ved at tage differenser et passende antal gange d .

Vi lader \mathbf{x} betegne søjlevektoren med observationerne x_1, \dots, x_N , dvs.

$$\mathbf{x} = (x_1, \dots, x_N)^*.$$

Først finder vi fordelingen af den tilsvarende stokastiske vektor $\mathbf{X} = (X_1, \dots, X_N)^*$ under antagelsen, at $\{X_t\}_{t \in \mathbb{Z}}$ er en ARMA(p, q)-proces med parametre $\mu = 0$, $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_p)$, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)$ and σ_W^2 . Ifølge [Sætning 2.19](#) kan processen repræsenteres som et baglæns glidende gennemsnit

$$X_t = \frac{\theta(B)}{\varphi(B)} W_t = \sum_{k=0}^{\infty} \psi_k W_{t-k}, \quad t \in \mathbb{Z}.$$

Lad $Q \in \mathbb{N}$ og betragt

$$\mathbf{X}_Q = (X_{Q1}, \dots, X_{QN})^*,$$

hvor

$$X_{Qt} = \sum_{k=0}^Q \psi_k W_{t-k}, \quad t = 1, \dots, N.$$

Da $\{W_t\}_{t \in \mathbb{Z}}$ er en følge af uafhængige og $N(0, \sigma_W^2)$ -fordelte stokastiske variable, og da desuden $(W_{1-Q}, \dots, W_0, X_{Q1}, \dots, X_{QN})$ er en lineær transformation af $(W_{1-Q}, \dots, W_0, W_1, \dots, W_N)$, fås det, at

$$\mathbf{X}_Q \sim N_N(0, \sigma_W^2 \Omega_Q),$$

hvor

$$\Omega_Q = \Omega_Q(\boldsymbol{\varphi}, \boldsymbol{\theta}) = \left\{ \sum_{k=0}^{Q-|i-j|} \psi_k \psi_{k+|i-j|} \right\}_{i,j=1}^N.$$

Lader vi $Q \rightarrow \infty$, finder vi, at

$$\mathbf{X} \sim N_N(0, \sigma_W^2 \Omega), \quad (3.20)$$

hvor

$$\Omega = \Omega(\boldsymbol{\varphi}, \boldsymbol{\theta}) = \left\{ \sum_{k=0}^{\infty} \psi_k \psi_{k+|i-j|} \right\}_{i,j=1}^N.$$

Bemærk, at vi med notationen fra [Kapitel 2](#) har, at

$$\sigma_W^2 \Omega = \mathbf{\Gamma}_N.$$

Af formel (3.20) ses, at log likelihood funktionen svarende til observationerne $\mathbf{x} = (x_1, \dots, x_N)^*$ er

$$l(\sigma_W^2, \boldsymbol{\varphi}, \boldsymbol{\theta}; \mathbf{x}) = -\frac{1}{2\sigma_W^2} \mathbf{x}^* \Omega^{-1} \mathbf{x} - \frac{1}{2} \log(|\Omega|) - \frac{N}{2} \log(\sigma_W^2) - \frac{N}{2} \log(2\pi), \quad (3.21)$$

hvor $|\cdot|$ betegner determinanten. Ofte skriver vi blot $l(\sigma_W^2, \boldsymbol{\varphi}, \boldsymbol{\theta})$ i stedet for $l(\sigma_W^2, \boldsymbol{\varphi}, \boldsymbol{\theta}; \mathbf{x})$.

For fast $(\boldsymbol{\varphi}, \boldsymbol{\theta})$ fås det, at maksimum likelihood estimatet $\hat{\sigma}_{W(\boldsymbol{\varphi}, \boldsymbol{\theta})}^2$ for σ_W^2 er

$$\hat{\sigma}_{W(\boldsymbol{\varphi}, \boldsymbol{\theta})}^2 = \frac{1}{N} \mathbf{x}^* \Omega^{-1} \mathbf{x}$$

og at profil log likelihood funktionen for $(\boldsymbol{\varphi}, \boldsymbol{\theta})$ er

$$\begin{aligned} \tilde{l}(\boldsymbol{\varphi}, \boldsymbol{\theta}) &= l(\hat{\sigma}_{W(\boldsymbol{\varphi}, \boldsymbol{\theta})}^2, \boldsymbol{\varphi}, \boldsymbol{\theta}) = -\frac{N}{2} - \frac{N}{2} \log\left(\frac{1}{N} \mathbf{x}^* \Omega^{-1} \mathbf{x}\right) - \frac{1}{2} \log(|\Omega|) - \frac{N}{2} \log(2\pi) \\ &\equiv -\frac{N}{2} \log(\mathbf{x}^* \Omega^{-1} \mathbf{x}) - \frac{1}{2} \log(|\Omega|), \end{aligned} \quad (3.22)$$

hvor \equiv antyder lighed på nær konstanter, der ikke afhænger af parametrene.

Lad $H = H(\boldsymbol{\varphi}, \boldsymbol{\theta})$ betegne den nedre trekantsmatriks med positive elementer i diagonalen i Cholesky dekomposition af $\Omega = \Omega(\boldsymbol{\varphi}, \boldsymbol{\theta})$, dvs.

$$\Omega = HH^* \quad (3.23)$$

og lad $\mathbf{e} = \mathbf{e}(\boldsymbol{\varphi}, \boldsymbol{\theta})$ være vektoren

$$\mathbf{e} = H^{-1} \mathbf{x}. \quad (3.24)$$

Ved hjælp af (3.20) fås det, at den tilsvarende stokastiske vektor \mathbf{E} har en N -dimensional normalfordeling med middelværdivektor $\mathbf{0}$ og med kovariansmatriks $\sigma_W^2 I_N$, hvor I_N betegner den N -dimensionale identitetsmatriks, dvs.

$$\mathbf{E} \sim N_N(\mathbf{0}, \sigma_W^2 I_N). \quad (3.25)$$

Desuden bliver profil log likelihood funktionen med denne notation

$$\begin{aligned}\tilde{l}(\boldsymbol{\varphi}, \boldsymbol{\theta}) &\equiv -\frac{N}{2} \log(\mathbf{e}^* \mathbf{e}) - \log(|H|) \\ &= -\frac{N}{2} \log(|H|^{1/N} \mathbf{e}^* \mathbf{e} |H|^{1/N}).\end{aligned}\quad (3.26)$$

Maksimum likelihood estimatet $(\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\theta}})$ for $(\boldsymbol{\varphi}, \boldsymbol{\theta})$ findes ved at minimere kvadratsummen

$$|H|^{1/N} \mathbf{e}^* \mathbf{e} |H|^{1/N}.$$

I SAS gøres dette ved hjælp af en Marquardt algoritme.

Maksimum likelihood estimatet $\hat{\sigma}_W^2$ for σ_W^2 er

$$\hat{\sigma}_W^2 = \hat{\sigma}_{W(\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\theta}})}^2 = \frac{1}{N} \hat{\mathbf{e}}^* \hat{\mathbf{e}},$$

hvor vektoren $\hat{\mathbf{e}} = \mathbf{e}(\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\theta}})$ kaldes *vektoren af residualer*.

Det ubetingede mindste kvadraters estimat – på engelsk *unconditional least square estimate* eller blot *uls estimate* – $(\check{\boldsymbol{\varphi}}, \check{\boldsymbol{\theta}})$ for $(\boldsymbol{\varphi}, \boldsymbol{\theta})$ findes ved at minimere kvadratsummen

$$\mathbf{e}^* \mathbf{e} = \mathbf{x}^* \boldsymbol{\Omega}^{-1} \mathbf{x},$$

dvs. ved at negligere leddet $-\log(|H|)$ i (3.26). Det tilsvarende estimat $\check{\sigma}_W^2$ for σ_W^2 er

$$\check{\sigma}_W^2 = \frac{1}{N} \check{\mathbf{e}}^* \check{\mathbf{e}}.$$

Vektoren $\check{\mathbf{e}} = \mathbf{e}(\check{\boldsymbol{\varphi}}, \check{\boldsymbol{\theta}})$ kaldes ofte *vektoren af ubetingede mindste kvadraters residualer* – *unconditional least square residuals* – eller blot *uls residualer*.

Bemærkning 3.9. Vi betragter nu igen ligningen (3.24) for at give et eksplicit udtryk for komponenterne e_t i \mathbf{e} ved hjælp af komponenterne i \mathbf{x} . For $t = 1, \dots, N$, lader vi \mathbf{x}_t betegne vektoren $(x_1, \dots, x_t)^*$ og lader $\sigma_W^2 \boldsymbol{\Omega}_t$ være kovariansmatricen for den stokastiske vektor \mathbf{X}_t . Betragt opdelingen af $\boldsymbol{\Omega}_t$ svarende til opdelingen $(\mathbf{x}_{t-1}^*, x_t^*)^*$ af \mathbf{x}_t^* , dvs.

$$\boldsymbol{\Omega}_t = \begin{Bmatrix} \boldsymbol{\Omega}_{t-1} & \mathbf{C}_t^* \\ \mathbf{C}_t & \boldsymbol{\Omega}_{tt} \end{Bmatrix},$$

hvor $\sigma_W^2 \mathbf{C}_t$ er $1 \times (t-1)$ -kovariansmatricen mellem x_t and \mathbf{x}_{t-1}^* . Lad desuden

$$\boldsymbol{\Omega}_{t,t-1} = \boldsymbol{\Omega}_{tt} - \mathbf{C}_t \boldsymbol{\Omega}_{t-1}^{-1} \mathbf{C}_t^*, \quad t = 2, \dots, N.$$

Med denne notation er

$$\mathbf{E}_1 = \boldsymbol{\Omega}_{11}^{-\frac{1}{2}} \mathbf{X}_1, \quad (3.27)$$

og

$$\mathbf{E}_t = \boldsymbol{\Omega}_{t,t-1}^{-\frac{1}{2}} (\mathbf{X}_t - \mathbf{C}_t \boldsymbol{\Omega}_{t-1}^{-1} \mathbf{X}_{t-1}), \quad t = 2, \dots, N. \quad (3.28)$$

For at vise disse formler lader vi $A = \{a_{ij}\}$ betegne matricen H^{-1} . Da A er en nedre trekantsmatriks, får vi fra (3.24), at

$$E_t = \sum_{j=1}^t a_{tj} X_j, \quad t = 1, \dots, N.$$

Formel (3.23) medfører, at $a_{11}^2 = \Omega_{11}^{-1}$ og (3.27) er vist.

For $t = 2, \dots, N$, medfører uafhængigheden af komponenterne i E (formel (3.25)), at

$$E(E_t | E_{t-1}) = 0$$

eller, ækvivalent hermed, idet E_{t-1} er i 1-1-forbindelse med X_{t-1} , at

$$0 = E(E_t | X_{t-1}) = a_{tt} E(X_t | X_{t-1}) + \sum_{j=1}^{t-1} a_{tj} X_j,$$

hvilket medfører, at

$$E_t = a_{tt} (X_t - E(X_t | X_{t-1})).$$

Ved hjælp af den velkendte formel $\text{Var } Y = \text{Var}\{E(Y|Z)\} + E\{\text{Var}(Y|Z)\}$ finder vi, at

$$\begin{aligned} \sigma_W^2 &= \text{Var } E_t = a_{tt}^2 \text{Var}(X_t - E(X_t | X_{t-1})) \\ &= a_{tt}^2 [E\{\text{Var}([X_t - E(X_t | X_{t-1})] | X_{t-1})\} \\ &\quad + \text{Var}\{E[X_t - E(X_t | X_{t-1}) | X_{t-1}]\}] \\ &= a_{tt}^2 E\{\text{Var}(X_t | X_{t-1})\} \\ &= a_{tt}^2 \sigma_W^2 \Omega_{t,t-1}, \end{aligned}$$

dvs. $a_{tt}^2 = \Omega_{t,t-1}^{-1}$, og da $E(X_t | X_{t-1}) = C_t \Omega_{t-1}^{-1} X_{t-1}$, er (3.28) vist.

Bemærk, at bruges (3.27), (3.28) og den velkendte identitet

$$\mathbf{x}_t^* \Omega_t^{-1} \mathbf{x}_t = (x_t - C_t \Omega_{t-1}^{-1} \mathbf{x}_{t-1})^* \Omega_{t,t-1}^{-1} (x_t - C_t \Omega_{t-1}^{-1} \mathbf{x}_{t-1}) + \mathbf{x}_{t-1}^* \Omega_{t-1}^{-1} \mathbf{x}_{t-1}$$

gentagne gange, får vi, at

$$\begin{aligned} \mathbf{e}^* \mathbf{e} &= x_1 \Omega_{11}^{-1} x_1 + \sum_{t=2}^N (x_t - C_t \Omega_{t-1}^{-1} \mathbf{x}_{t-1})^* \Omega_{t,t-1}^{-1} (x_t - C_t \Omega_{t-1}^{-1} \mathbf{x}_{t-1}) \\ &= \mathbf{x}^* \Omega^{-1} \mathbf{x}. \end{aligned} \quad \blacksquare$$

Til sidst omtaler vi *betinget mindste kvadraters estimator – conditional least square estimates* eller kort *cls estimator*. Her antages det, at de uobserverede tidligere værdier $x_t, t \leq 0$, alle er lig med 0. Fra repræsentationen af processen $\{X_t\}_{t \in \mathbb{Z}}$ som en baglæns autoregression

$$W_t = \frac{\varphi(B)}{\theta(B)} X_t = X_t + \sum_{k=1}^{\infty} \pi_k X_{t-k} \quad t \in \mathbb{Z},$$

finder vi ved at betinge med $\{X_t = 0, t \leq 0\}$, at elementerne af den hvide støj svarende til observationerne x_1, \dots, x_N er

$$w_t = x_t + \sum_{k=1}^{t-1} \pi_k x_{t-k}, \quad t = 1, \dots, N.$$

Lader $w(\varphi, \theta) = w = (w_1, \dots, w_N)^*$, så findes *cls* estimatet $(\tilde{\varphi}, \tilde{\theta})$ for (φ, θ) ved at minimere kvadratsummen

$$w^* w = \sum_{t=1}^N w_t^2.$$

Det tilsvarende estimat $\tilde{\sigma}_W^2$ for σ_W^2 er

$$\tilde{\sigma}_W^2 = \frac{1}{N} \tilde{w}^* \tilde{w},$$

hvor $\tilde{w} = w(\tilde{\varphi}, \tilde{\theta})$ er vektoren af *cls* residualer.

Hvis ARMA-processen har middelværdi μ er log likelihood funktionen

$$l(\mu, \sigma_W^2, \varphi, \theta; \mathbf{x}) \equiv -\frac{1}{2\sigma_W^2} (\mathbf{x} - \boldsymbol{\mu})^* \Omega^{-1} (\mathbf{x} - \boldsymbol{\mu}) - \frac{1}{2} \log(|\Omega|) - \frac{N}{2} \log(\sigma_W^2), \quad (3.29)$$

hvor $\boldsymbol{\mu}$ betegner $N \times 1$ vektoren $(\mu, \dots, \mu)^*$.

Ofte estimeres middelværdien μ ved den empiriske middelværdi, dvs.

$$\mu \leftarrow \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad (3.30)$$

Imidlertid er maksimum likelihood estimatet $\hat{\mu}_{(\varphi, \theta)}$ for μ for fast værdi af (φ, θ)

$$\hat{\mu}_{(\varphi, \theta)} = \frac{\mathbf{1}^* \Omega^{-1} \mathbf{x}}{\mathbf{1}^* \Omega^{-1} \mathbf{1}}, \quad (3.31)$$

hvor $\mathbf{1}$ er $N \times 1$ vektoren $(1, \dots, 1)^*$.

Ved hjælp af (3.20) fås, at

$$\bar{X} \sim N\left(\mu, \frac{\sigma_W^2}{N^2} \mathbf{1}^* \Omega^{-1} \mathbf{1}\right)$$

og

$$\hat{\mu}_{(\varphi, \theta)} \sim N\left(\mu, \frac{\sigma_W^2}{\mathbf{1}^* \Omega^{-1} \mathbf{1}}\right),$$

dvs. $\hat{\mu}_{(\varphi, \theta)}$ er et udfald af en normalfordelt stokastisk variabel med middelværdi μ og varians $\sigma_W^2/(\mathbf{1}^* \Omega^{-1} \mathbf{1})$.

Hvis μ estimeres ved \bar{x} , beregnes maksimum likelihood estimerne for de resterende parametre σ_W^2, φ and θ som i tilfældet $\mu = 0$ – betragtet ovenfor – idet \mathbf{x} erstattes med $\mathbf{x} - \bar{x}$, hvor \bar{x} er $N \times 1$ vektoren $(\bar{x}, \dots, \bar{x})^*$.

Hvis μ estimeres ved maksimum likelihood metoden kan en iterativ procedure anvendes. Ved hjælp af begyndelsesværdi μ_0 for μ , beregnes estimaterne φ_{μ_0} og θ_{μ_0} for φ og θ som i tilfældet $\mu = 0$ – betragtet ovenfor – ved at erstatte \mathbf{x} med $\mathbf{x} - \mu_0 \mathbf{1}$. Den næste værdi μ_1 af μ fås så fra (3.31) som $\mu_1 = \hat{\mu}_{(\varphi_{\mu_0}, \theta_{\mu_0})}$, og fra denne beregnes φ_{μ_1} og θ_{μ_1} og så videre. Proceduren gentages indtil konvergens indtræffer. Betegner $(\hat{\mu}, \hat{\varphi}, \hat{\theta})$ maksimum likelihood estimatet for (μ, φ, θ) , er maksimum likelihood estimatet for σ_W^2

$$\hat{\sigma}_W^2 = \frac{1}{N}(\mathbf{x} - \hat{\boldsymbol{\mu}})^* \hat{\boldsymbol{\Omega}}^{-1}(\mathbf{x} - \hat{\boldsymbol{\mu}}).$$

Som i den sædvanlige teori for lineære normale modeller, benyttes ofte et estimat, der fremkommer ved at dividere kvadratsummen med antallet af observationer minus antallet af parametre, der specificerer modellen, dvs.

$$\sigma_W^2 \leftarrow \frac{1}{N - (p + q + 1)}(\mathbf{x} - \hat{\boldsymbol{\mu}})^* \hat{\boldsymbol{\Omega}}^{-1}(\mathbf{x} - \hat{\boldsymbol{\mu}}). \quad (3.32)$$

Lignende bemærkninger, som den ovenfor, gælder også for cls estimation og for uls estimation.

For at angive fordelingen af maksimum likelihood estimatorerne skal vi bruge asymptotisk teori. For at reducere notationen i det følgende lader vi $l_{/\mu}$ betegne den partielt afledede af l med hensyn til μ , dvs.

$$l_{/\mu} = \frac{\partial}{\partial \mu} l(\mu, \sigma_W^2, \varphi, \theta; \mathbf{x}).$$

Tilsvarende lader vi

$$l_{/\mu \sigma_W^2} = \frac{\partial^2}{\partial \mu \partial \sigma_W^2} l(\mu, \sigma_W^2, \varphi, \theta; \mathbf{x}),$$

$$l_{/\mu \varphi} = \frac{\partial^2}{\partial \mu \partial \varphi} l(\mu, \sigma_W^2, \varphi, \theta; \mathbf{x})$$

osv.

Lad j betegne den observerede Fisher information, dvs. $(p + q + 2) \times (p + q + 2)$ matricen

$$j = j(\mu, \sigma_W^2, \varphi, \theta; \mathbf{x}) = -l_{/(\mu, \sigma_W^2, \varphi, \theta)(\mu, \sigma_W^2, \varphi, \theta)}$$

og sæt

$$\hat{j} = j(\hat{\mu}, \hat{\sigma}_W^2, \hat{\varphi}, \hat{\theta}; \mathbf{x}).$$

Idet $(\hat{\Omega}^{-1})_{/(\varphi, \theta)}$ betegner $(\Omega^{-1})_{/(\varphi, \theta)}$ beregnet i $(\hat{\varphi}, \hat{\theta})$ osv., er det vist i [Andersen og Blæsild \(2003\)](#), at \hat{j} har følgende elementer:

$$\begin{aligned}\hat{j}_{\mu\mu} &= \frac{1}{\hat{\sigma}_W^2} \mathbf{1}^* \hat{\Omega}^{-1} \mathbf{1}, \\ \hat{j}_{\mu\sigma_W^2} &= 0, \\ \hat{j}_{\mu(\varphi, \theta)} &= -\frac{1}{\hat{\sigma}_W^2} \mathbf{1}^* (\hat{\Omega}^{-1})_{/(\varphi, \theta)} (\mathbf{x} - \hat{\boldsymbol{\mu}}), \\ \hat{j}_{\sigma_W^2\sigma_W^2} &= \frac{N}{2(\hat{\sigma}_W^2)^2}, \\ \hat{j}_{\sigma_W^2(\varphi, \theta)} &= -\frac{1}{2\hat{\sigma}_W^2} (\log(|\hat{\Omega}|))_{/(\varphi, \theta)(\varphi, \theta)}, \\ \hat{j}_{(\varphi, \theta)(\varphi, \theta)} &= \frac{1}{2\hat{\sigma}_W^2} (\mathbf{x} - \hat{\boldsymbol{\mu}})^* (\hat{\Omega}^{-1})_{/(\varphi, \theta)(\varphi, \theta)} (\mathbf{x} - \hat{\boldsymbol{\mu}}) + \frac{1}{2} (\log(|\hat{\Omega}|))_{/(\varphi, \theta)(\varphi, \theta)}.\end{aligned}$$

Elementerne i den forventede Fisher information, dvs. i $(p + q + 2) \times (p + q + 2)$ matricen

$$i = i(\mu, \sigma_W^2, \varphi, \theta) = E_{(\mu, \sigma_W^2, \varphi, \theta)} \{j(\mu, \sigma_W^2, \varphi, \theta; \mathbf{X})\}$$

er også angivet i [Andersen og Blæsild \(2003\)](#). De er:

$$\begin{aligned}i_{\mu\mu} &= \frac{1}{\sigma_W^2} \mathbf{1}^* \Omega^{-1} \mathbf{1}, \\ i_{\mu\sigma_W^2} &= 0, \\ i_{\mu(\varphi, \theta)} &= 0, \\ i_{\sigma_W^2\sigma_W^2} &= \frac{N}{2(\sigma_W^2)^2}, \\ i_{\sigma_W^2(\varphi, \theta)} &= \frac{1}{2\sigma_W^2} (\log(|\Omega|))_{/(\varphi, \theta)}, \\ i_{(\varphi, \theta)(\varphi, \theta)} &= \frac{1}{2} \text{tr}((\Omega^{-1})_{/(\varphi, \theta)} \Omega) + \frac{1}{2} (\log(|\Omega|))_{/(\varphi, \theta)(\varphi, \theta)}.\end{aligned}$$

Da kovariansmatricen for den stokastiske vektor $\mathbf{X} = (X_1, \dots, X_N)^*$ ikke er en diagonalmatris, dvs. da komponenterne af \mathbf{X} ikke er uafhængige, kan den asymptotiske fordeling af maksimum likelihood estimatoren $(\hat{\mu}, \hat{\sigma}_W^2, \hat{\varphi}, \hat{\theta})$ for $(\mu, \sigma_W^2, \varphi, \theta)$ ikke findes ved den sædvanlige asymptotiske teori, som forudsætter at observationerne er uafhængige. Imidlertid er det ved hjælp af generalisationer af de centrale grænseværdisætninger til situationer, hvor de stokastiske variable er korrelerede, muligt for ARMA-processerne at vise asymptotiske resultater for maksimum likelihood estimatoren, som formelt er analoge med resultaterne i tilfældet med uafhængige observationer. Hvis $(\mu, \sigma_W^2, \varphi, \theta)$ betegner den "sande" værdi af den $(p + q + 2)$ -dimensionale parameter for $ARMA(p, q)$ -processen, kan

det vises, at

$$\left\{(\hat{\mu}, \hat{\sigma}_W^2, \hat{\phi}, \hat{\theta}) - (\mu, \sigma_W^2, \varphi, \theta)\right\} \left\{i(\mu, \sigma_W^2, \varphi, \theta)\right\}^{\frac{1}{2}} \sim N_{p+q+2}(\mathbf{0}, I_{p+q+2}) \quad \text{når } N \rightarrow \infty. \quad (3.33)$$

For store værdier af N kan fordelingen af $(\hat{\mu}, \hat{\sigma}_W^2, \hat{\phi}, \hat{\theta})$ derfor approksimeres med en $(p+q+2)$ -dimensional normalfordeling med middelværdivektor $(\mu, \sigma_W^2, \varphi, \theta)$ og kovariansmatriks

$$\text{Var}(\hat{\mu}, \hat{\sigma}_W^2, \hat{\phi}, \hat{\theta}) = \left\{i(\mu, \sigma_W^2, \varphi, \theta)\right\}^{-1}. \quad (3.34)$$

Af formlerne ovenfor fås, at

$$\begin{aligned} \text{Var}(\hat{\mu}) &= \frac{\sigma_W^2}{\mathbf{1}^* \Omega^{-1} \mathbf{1}}, \\ \text{Var}(\hat{\sigma}_W^2) &= \frac{2(\sigma_W^2)^2}{N} \end{aligned}$$

og desuden ses det, at $\hat{\mu}$ og $(\hat{\sigma}_W^2, \hat{\phi}, \hat{\theta})$ er asymptotisk uafhængige. Det kan desuden vises, se Box and Jenkins (1970), at $\hat{\sigma}_W^2$ og $(\hat{\phi}, \hat{\theta})$ er asymptotisk uafhængige. For $ARMA(p, q)$ -processer med $p+q < 3$ kan følgende udtryk for den asymptotiske (ko)varians for $(\hat{\phi}, \hat{\theta})$ findes i [Box og Jenkins \(1970\)](#):

$$\begin{aligned} \text{AR}(1): \quad & \text{Var}(\hat{\phi}) = \frac{1}{N}(1 - \varphi^2), \\ \text{MA}(1): \quad & \text{Var}(\hat{\theta}) = \frac{1}{N}(1 - \theta^2), \\ \text{AR}(2): \quad & \text{Var}(\hat{\phi}_1, \hat{\phi}_2) = \frac{1}{N} \begin{pmatrix} 1 - \varphi_2^2 & -\varphi_1(1 + \varphi_2) \\ -\varphi_1(1 + \varphi_2) & 1 - \varphi_2^2 \end{pmatrix}, \\ \text{MA}(2): \quad & \text{Var}(\hat{\theta}_1, \hat{\theta}_2) = \frac{1}{N} \begin{pmatrix} 1 - \theta_2^2 & -\theta_1(1 + \theta_2) \\ -\theta_1(1 + \theta_2) & 1 - \theta_2^2 \end{pmatrix}, \\ \text{ARMA}(1,1): \quad & \text{Var}(\hat{\phi}, \hat{\theta}) = \frac{1 - \varphi\theta}{N(\varphi - \theta)^2} \begin{pmatrix} (1 - \varphi^2)(1 - \varphi\theta) & (1 - \varphi^2)(1 - \theta^2) \\ (1 - \varphi^2)(1 - \theta^2) & (1 - \theta^2)(1 - \varphi\theta) \end{pmatrix}. \end{aligned}$$

Det kan vises, at disse asymptotiske resultater også gælder for såvel cls estimerer som for uls estimerer.

For at anvende disse resultater i praksis benyttes approksimationen

$$i(\mu, \sigma_W^2, \varphi, \theta) \approx i(\hat{\mu}, \hat{\sigma}_W^2, \hat{\phi}, \hat{\theta}),$$

hvor den “sande” men ukendte parameter erstattes med maksimum likelihood estimatet (eller et af de andre estimerer nævnt ovenfor). Undertiden erstattes

den forventede Fisher information med den observerede information, dvs. man benytter approksimationen

$$i(\mu, \sigma_W^2, \boldsymbol{\varphi}, \boldsymbol{\theta}) \approx j(\hat{\mu}, \hat{\sigma}_W^2, \hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\theta}}).$$

Vi afslutter de teoretiske overvejelser i dette afsnit med at omtale en alternativ estimationsmetode for autoregressive processer kendt under navnet Yule-Walker estimation. For en $AR(p)$ -proces findes Yule-Walker estimatet $\vec{\boldsymbol{\varphi}} = (\vec{\varphi}_1, \dots, \vec{\varphi}_p)$ for $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_p)$ ved at løse den empiriske version af Yule-Walker ligningerne (2.27). Mere præcist, hvis \mathbf{R}_p betegner $p \times p$ -matricen, der fremkommer ud fra den teoretiske korrelationsmatris \mathbf{P}_p ved at erstatte ρ_k med den empiriske autokorrelation r_k for $k = 1, \dots, p-1$, dvs.

$$\mathbf{R}_p = \begin{pmatrix} 1 & r_1 & \cdots & r_{p-2} & r_{p-1} \\ r_1 & 1 & \cdots & r_{p-3} & r_{p-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r_{p-2} & r_{p-3} & \cdots & 1 & r_1 \\ r_{p-1} & r_{p-2} & \cdots & r_1 & 1 \end{pmatrix},$$

så er $(\vec{\varphi}_1, \dots, \vec{\varphi}_p)$ givet ved

$$\vec{\boldsymbol{\varphi}} = (\vec{\varphi}_1, \dots, \vec{\varphi}_p) = (r_1, \dots, r_p) \mathbf{R}_p^{-1}. \quad (3.35)$$

Variansen σ_W^2 af den hvide støj estimeres ved

$$\vec{\sigma}_W^2 = c_0(1 - \vec{\boldsymbol{\varphi}} \mathbf{R}_p \vec{\boldsymbol{\varphi}}^*)$$

svarende til formlen

$$\sigma_W^2 = \gamma_0(1 - \boldsymbol{\varphi} \mathbf{P}_p \boldsymbol{\varphi}^*).$$

Endelig estimeres μ ved \bar{x} . Det kan vises, at asymptotisk (når $N \rightarrow \infty$) er maksimum likelihood estimatoren og Yule-Walker estimatoren ækvivalente.

Eksempel 3.10. I SAS-udskrift 3.3 på side 51, så vi, at MINIC foreslog en MA(1)-model til beskrivelse af differenserne af kursen på US\$. PROC ARIMA giver følgende information om estimation i denne model med $\mu = 0$:

SAS-udskrift 3.17

Name of Variable = DK	
Period(s) of Differencing	1
Mean of Working Series	-0.00342
Standard Deviation	0.179116
Number of Observations	492
Observation(s) eliminated by differencing	1

Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag
MA1,1	-0.33820	0.04250	-7.96	<.0001	1
Variance Estimate			0.028909		

Udskriften starter med noget forholdsvis indlysende information om tidsrækken af førstedifferenser, blandt andet at den empiriske middelværdi er -0.00342 og at den empiriske spredning er 0.179116 . Herefter ses, at maksimum likelihood estimation giver MA(1)-modellen for differenserne, eller ARIMA(0,1,1)-modellen for kurserne $\{X_t\}$,

$$\nabla X_t = W_t + 0.33820W_{t-1}, \quad (3.36)$$

hvor

$$\sigma_W^2 \leftarrow 0.028909.$$

Eksempel 3.11. Maksimum likelihood estimaterne i AR(3)-modellen for isotopkurven,

$$\varphi(B)(X_t - \mu) = (1 - \varphi_1 B - \varphi_2 B^2 - \varphi_3 B^3)(X_t - \mu) = W_t,$$

fremgår af [SAS-udskrift 3.18](#).

SAS-udskrift 3.18

Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag
MU	293.33436	0.66563	440.69	<.0001	0
AR1,1	0.46804	0.03555	13.16	<.0001	1
AR1,2	-0.11268	0.03908	-2.88	0.0039	2
AR1,3	0.08802	0.03555	2.48	0.0133	3
Constant Estimate			163.2765		
Variance Estimate			108.9995		
Std Error Estimate			10.44028		
AIC			5967.372		
SBC			5986.07		
Number of Residuals			792		

Correlations of Parameter Estimates				
Parameter	MU	AR1,1	AR1,2	AR1,3
MU	1.000	0.002	0.001	0.003
AR1,1	0.002	1.000	-0.421	0.072
AR1,2	0.001	-0.421	1.000	-0.421
AR1,3	0.003	0.072	-0.421	1.000

Bemærkning 3.12. Med udgangspunkt i SAS-udskrifterne 3.17 og 3.18 kommer der nu nogle generelle kommentarer om udskriften fra estimation i ARIMA(p, d, q)-modellen

$$(1 - \varphi_1 B - \dots - \varphi_p B^p)(\nabla^d X_t - \mu) = (1 - \theta_1 B - \dots - \theta_q B^q)W_t$$

ved hjælp af PROC ARIMA.

I tabellen Maximum Likelihood Estimation findes estimaterne for parametrene $\mu, \varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q$. I én række i tabellen angives foruden et estimat (Estimate) en Standard Error for estimatet, dvs. et estimat for den teoretiske spredning af den tilsvarende estimator – beregnet ved hjælp af $j(\hat{\mu}, \hat{\sigma}_W^2, \hat{\varphi}, \hat{\theta})$ – og et approksimativt t -test (t Value) og den tilsvarende testsandsynlighed (Approx Pr > |t|) for hypotesen at parameteren er 0.

Tallet i linjen Constant Estimate er

$$\hat{\phi}(1)\mu = (1 - \hat{\phi}_1 - \dots - \hat{\phi}_p)\mu.$$

Estimaterne for σ_W^2 og σ_W beregnet ved hjælp af (3.32) findes henholdsvis i linjerne Variance Estimate og Std Error Estimate.

Estimatoren for σ_W^2 er approksimativt $\sigma_W^2 \chi^2(f)/f$ -fordelt med $f = (N - d) - (p + q + 1)$ frihedsgrader, så standard error for estimatet for σ_W^2 kan beregnes som estimatet ganget med $\sqrt{2/f}$.

Hvis \hat{l} er den maksimale værdi af log likelihood funktionen og $n = N - d$, er de to informationskriterier for ARIMA(p, d, q)-modellen, der er givet på udskriften, Akaike's information criterion (AIC), hvor

$$AIC = -2\hat{l} + 2(p + q + 1),$$

og Schwartz's Bayesian criterion (SBC), hvor

$$SBC = -2\hat{l} + (p + q + 1)\log(n).$$

Number of Residuals er $N - d$, antallet af observationer i tidsrækken $\{\nabla^d x_t\}$ af d 'te differenser.

Endelig angives estimatet for $(p + q + 1) \times (p + q + 1)$ korrelationsmatricen for estimatorerne for parametrene $\mu, \varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q$. Da vi kender estimator for spredningen på estimatorerne, kan der beregnes et estimat for kovarianssmatricen for estimatorerne. ■

Eksempel 3.13. Af [SAS-udskrift 3.18](#) på [side 70](#) ses, at den estimerede AR(3)-model for isotopkurven er

$$X_t = 163.2765 + 0.46804X_{t-1} - 0.11268X_{t-2} + 0.08802X_{t-3} + W_t, \quad (3.37)$$

og

$$\sigma_W^2 \leftarrow 108.9995.$$

Det ses desuden af udskriften, at $\varphi_3(\text{AR1}, 3)$ er signifikant forskellig fra 0 ($t = 2.48$, $p_{\text{obs}} = 0.0133$). Vi kan med andre ord ikke reducere AR(3)-modellen til en AR(1)-model, som MINIC foreslog.

Eksempel 3.14. Option MINIC foreslog i [SAS-udskrift 3.7](#) på [side 53](#) en AR(2)-model for tidsrækken bestående af målingerne af nedbøren i [Eksempel 1.3](#), hvilket var lidt overraskende på baggrund af de empiriske autokorrelationsfunktioner øverst i [Figur 3.4](#) på [side 52](#), som antyder, at rækken kan modelleres ved en hvid støj. [SAS-udskrift 3.19](#) viser maksimum likelihood estimationen i AR(2)-modellen. Det ses, at den estimerede model er

$$X_t = 266.0528 + 0.08094X_{t-1} + 0.14670X_{t-2} + W_t.$$

Estimatet for variansen er

$$\sigma_W^2 \leftarrow 7437.076.$$

I 4. og 5. søjle ses approksimative t -test for hypoteserne, at $\varphi_1(\text{AR1}, 1)=0$ og $\varphi_2(\text{AR1}, 2)=0$. Da hypoteserne ikke forkastes, kan AR(2)-modellen reduceres til en hvid støj.

Eksemplet her viser, at de modeller, MINIC foreslår, ikke nødvendigvis er de endelige modeller.

SAS-udskrift 3.19

Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag
MU	344.46639	11.97630	28.76	<.0001	0
AR1,1	0.08094	0.11171	0.72	0.4687	1
AR1,2	0.14670	0.11212	1.31	0.1907	2
Constant Estimate			266.0528		
Variance Estimate			7437.076		

Modellen for nedbøren er altså

$$X_t = \mu + W_t,$$

hvor

$$\mu \leftarrow 345.2012 \quad \text{og} \quad \sigma_W^2 \leftarrow 7487.2004.$$

I en tilsvarende model for vandmængden er estimerne

$$\mu \leftarrow 1164.8332 \quad \text{og} \quad \sigma_W^2 \leftarrow 54473.6059.$$

Eksempel 3.15. Identifikationen startende på [side 54](#) gav følgende forslag til modeller for førstedifferenserne af logaritmerne til henholdsvis faste investeringer, AR(4), disponibel indkomst, ARMA(0,0), og udgifter til forbrug, AR(2). De estimerede modeller er:

Logaritmen til faste investeringer

$$(1 - \varphi_1 B - \varphi_2 B^2 - \varphi_3 B^3 - \varphi_4 B^4)(\nabla X_t - \mu) = W_t,$$

hvor estimerne ses i

SAS-udskrift 3.20

Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag
MU	0.01899	0.0063237	3.00	0.0027	0
AR1,1	-0.23692	0.11239	-2.11	0.0350	1
AR1,2	-0.03744	0.11408	-0.33	0.7428	2
AR1,3	0.16085	0.11452	1.40	0.1602	3
AR1,4	0.32959	0.11210	2.94	0.0033	4
Constant Estimate			0.014886		
Variance Estimate			0.001944		

Logaritmen til disponibel indkomst

$$\nabla X_t = \mu + W_t,$$

hvor

$$\mu \leftarrow 0.020711 \quad \text{og} \quad \sigma_W^2 \leftarrow 0.000146.$$

Logaritmen til udgifter til forbrug

$$(1 - \varphi_1 B - \varphi_2 B^2)(\nabla X_t - \mu) = W_t,$$

hvor estimerne ses i

SAS-udskrift 3.21

Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag
MU	0.01982	0.0014767	13.42	<.0001	0
AR1,1	-0.05664	0.11463	-0.49	0.6213	1
AR1,2	0.27287	0.11507	2.37	0.0177	2
The ARIMA Procedure					
Constant Estimate		0.015533			
Variance Estimate		0.000102			

Eksempel 3.16. Identifikationen startende på [side 58](#) førte frem til en MA(2)-model for førstedifferenserne for renten på de 1-årige obligationer og en MA(1)-model for førstedifferenserne for renten på de 10-årige obligationer. De estimate-rede modeller er, idet det antages at $\mu = 0$:

1-årige obligationer:

$$X_t = W_t - \theta_1 W_{t-1} - \theta_2 W_{t-2},$$

hvor estimaterne ses i

SAS-udskrift 3.22

Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag
MA1,1	0.11117	0.07963	1.40	0.1627	1
MA1,2	0.37302	0.07990	4.67	<.0001	2
Variance Estimate		2.30778			

10-årige obligationer:

$$X_t = W_t - \theta_1 W_{t-1},$$

hvor estimaterne ses i

SAS-udskrift 3.23

Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag
MA1,1	0.19239	0.08323	2.31	0.0208	1
Variance Estimate			0.574696		

Modelkontrol

Vurderingen af, om ARMA(p, q)-modellen

$$\varphi(B)(X_t - \mu) = \theta(B)W_t, \quad t \in \mathbb{Z}, \quad (3.38)$$

er en rimelig model til beskrivelse af den empiriske tidsrække x_1, \dots, x_N , baseres på residualerne fra den estimerede model. Residualerne afhænger naturligvis af hvilken slags estimation, der er benyttet. Diskussionen nedenfor gælder også for ARIMA-modeller, hvis den oprindelige tidsrække x_1, \dots, x_N erstattes af rækken $\nabla^d x_{d+1}, \dots, \nabla^d x_N$ af d 'te differenser.

Residualerne \hat{w} fra maksimum likelihood estimation defineres ved hjælp af formlerne (3.27) og (3.28). Hvis $(\hat{\mu}, \hat{\sigma}_W^2, \hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\theta}})$ betegner maksimum likelihood estimatet for $(\mu, \sigma_W^2, \boldsymbol{\varphi}, \boldsymbol{\theta})$, er maksimum likelihood residualerne

$$\hat{w}_1 = \hat{\Omega}_{11}^{-\frac{1}{2}}(x_1 - \hat{\mu}) \quad (3.39)$$

$$\hat{w}_t = \hat{\Omega}_{t,t-1}^{-\frac{1}{2}}(x_t - \hat{\mu} - \hat{C}_t \hat{\Omega}_{t-1}^{-1}(\mathbf{x} - \hat{\boldsymbol{\mu}})_{t-1}), \quad t = 2, \dots, N, \quad (3.40)$$

hvor $\hat{\Omega} = \Omega_{(\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\theta}})}$ osv.

Uls residualerne – ubetinget mindste kvadraters residualerne – \check{w} defineres som i (3.39) og (3.40) men med “ $\hat{\cdot}$ ” erstattes med “ $\check{\cdot}$ ”.

Cls residualerne – betinget mindste kvadraters residualerne – \tilde{w} beregnes som

$$\tilde{w}_t = x_t - \tilde{\mu} - \sum_{k=1}^t \tilde{\pi}_k(x_{t-k} - \tilde{\mu}), \quad t = 1, \dots, N,$$

hvor $(\tilde{\mu}, \tilde{\sigma}_W^2, \tilde{\boldsymbol{\varphi}}, \tilde{\boldsymbol{\theta}})$ er cls estimatet for $(\mu, \sigma_W^2, \boldsymbol{\varphi}, \boldsymbol{\theta})$ og hvor $\{\tilde{\pi}_k\}_{k \in N_0}$ – i en indlysende notation – er bestemt ved

$$\frac{\tilde{\varphi}(B)}{\tilde{\theta}(B)} = 1 - \sum_{k=1}^{\infty} \tilde{\pi}_k B^k.$$

Tilpasningen af modellen (3.38) baseret på maksimum likelihood estimation vurderes ved at undersøge, om \hat{w} er en hvid støj. Dette gøres ved plote residualerne og ved at beregne de empiriske autokorrelationsfunktioner for residualerne. De empiriske autokorrelationer for residualerne er

$$r_k(\hat{w}) = \frac{c_k(\hat{w})}{c_0(\hat{w})}, \quad k = 0, 1, \dots, N,$$

hvor

$$c_k(\hat{w}) = \frac{1}{N} \sum_{t=1}^{N-k} \hat{w}_t \hat{w}_{t+k}.$$

De empiriske partielle autokorrelationer for residualerne $\{\hat{\phi}_{k,k}(\hat{w})\}$ beregnes ud fra $\{r_k(\hat{w})\}$ på samme måde som de teoretiske partielle autokorrelationer $\{\phi_{k,k}\}$ beregnes ud fra de teoretiske autokorrelationer $\{\rho_k\}$.

Desuden kan tilpasningen af modellen bedømmes ved hjælp af *Ljung-Box* χ^2 statistics for testing lack of fit. Disse teststørrelser er

$$\chi_m^2 = N(N+2) \sum_{k=1}^m \frac{r_k^2(\hat{w})}{N-k} \sim \chi^2(m - (p+q)). \quad (3.41)$$

Resultaterne i (3.13) og (3.14) gælder ikke for de empiriske autokorrelationsfunktioner for residualerne.

For at formulere resultaterne om den asymptotiske fordeling af de empiriske autokorrelationer for residualerne er det nødvendigt at introducere den følgende notation. Lad $\tilde{\varphi}(z)$ betegne produktet af polynomierne $\varphi(z)$ og $\theta(z)$, dvs.

$$\tilde{\varphi}(z) = \varphi(z)\theta(z), \quad z \in \mathbb{C}.$$

Da $\tilde{\varphi}(z) \neq 0$ for $|z| \leq 1$, har vi

$$\tilde{\varphi}^{-1}(z) = \sum_{j=0}^{\infty} \tilde{\psi}_j z^j \quad \text{for } |z| \leq 1,$$

hvor $\tilde{\psi}_0 = 1$. Lad

$$\tilde{\sigma}^2 = \sum_{j=0}^{\infty} \tilde{\psi}_j^2,$$

$$\tilde{\rho}_k = \frac{1}{\tilde{\sigma}^2} \sum_{j=0}^{\infty} \tilde{\psi}_j \tilde{\psi}_{j+k}, \quad k \geq 0,$$

og lad desuden $(p+q) \times (p+q)$ matricen A være givet ved

$$A^{-1} = \begin{pmatrix} 1 & \tilde{\rho}_1 & \tilde{\rho}_2 & \cdots & \tilde{\rho}_{p+q-1} \\ \tilde{\rho}_1 & 1 & \tilde{\rho}_1 & \cdots & \tilde{\rho}_{p+q-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{\rho}_{p+q-1} & \tilde{\rho}_{p+q-2} & \tilde{\rho}_{p+q-3} & \cdots & 1 \end{pmatrix}.$$

Endelig, for $m \geq p + q$ lader vi G_m være $m \times (p + q)$ matricen givet ved

$$G_m = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \tilde{\psi}_1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{\psi}_{p+q-1} & \tilde{\psi}_{p+q-2} & \tilde{\psi}_{p+q-3} & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{\psi}_{m-1} & \tilde{\psi}_{m-2} & \tilde{\psi}_{m-3} & \cdots & \tilde{\psi}_{m-(p+q)} \end{pmatrix}$$

og lader H_m betegne $m \times m$ matricen

$$H_m = \frac{1}{\tilde{\sigma}^2} G_m A G_m^*.$$

Med denne notation gælder der følgende sætning, der skyldes [Box og Jenkins \(1970\)](#).

Sætning 3.17. *For de empiriske autokorrelationsfunktioner for residualerne gælder der følgende asymptotiske resultater for $N \rightarrow \infty$:*

$$\begin{aligned} \sqrt{N}(r_1(\hat{w}), \dots, r_m(\hat{w})) &\xrightarrow{\sim} N_m(\mathbf{0}, I_m - H_m) \\ \sqrt{N}(\hat{\phi}_{1,1}(\hat{w}), \dots, \hat{\phi}_{m,m}(\hat{w})) &\xrightarrow{\sim} N_m(\mathbf{0}, I_m - H_m), \end{aligned}$$

hvor I_m er $m \times m$ enhedsmatricen og $\xrightarrow{\sim}$ betegner konvergens i fordeling.

I anvendelser af disse resultater sættes $\tilde{\varphi}(z) = \hat{\varphi}(z)\hat{\theta}(z)$.

Bemærkning 3.18. Modelkontrollen baseret på de empiriske autokorrelationsfunktioner for residualerne er i modsætning til test baseret på Ljung-Box χ^2 -teststørrelserne i (3.41) konstruktiv i den forstand, at autokorrelationsfunktionerne kan antyde, hvordan modellen (3.38) skal udvides for at opnå en rimelig beskrivelse af den observerede tidsrække x_1, \dots, x_N .

Mere præcist antag, at residualerne \hat{w} fra maksimum likelihood estimation i ARMA(p, q)-modellen i (3.38) antyder, at $\{W_t\}_{t \in \mathbb{Z}}$ ikke er en hvid støj men kan beskrives ved hjælp af ARMA(p_W, q_W)-modellen

$$\varphi_W(B)W_t = \theta_W(B)V_t, \quad t \in \mathbb{Z}, \quad (3.42)$$

hvor $\{V_t\}_{t \in \mathbb{Z}}$ er en hvid støj. Formel (3.42) er ækvivalent med

$$W_t = \frac{\theta_W(B)}{\varphi_W(B)} V_t.$$

Indsættes dette i (3.38), fås

$$\varphi(B)(X_t - \mu) = \theta(B) \frac{\theta_W(B)}{\varphi_W(B)} V_t. \quad (3.43)$$

Residualerne \hat{w} foreslår altså at betragte ARMA($p_W + p, q_W + q$)-modellen

$$\varphi_W(B)\varphi(B)(X_t - \mu) = \theta_W(B)\theta(B)V_t, \quad t \in \mathbb{Z}, \quad (3.44)$$

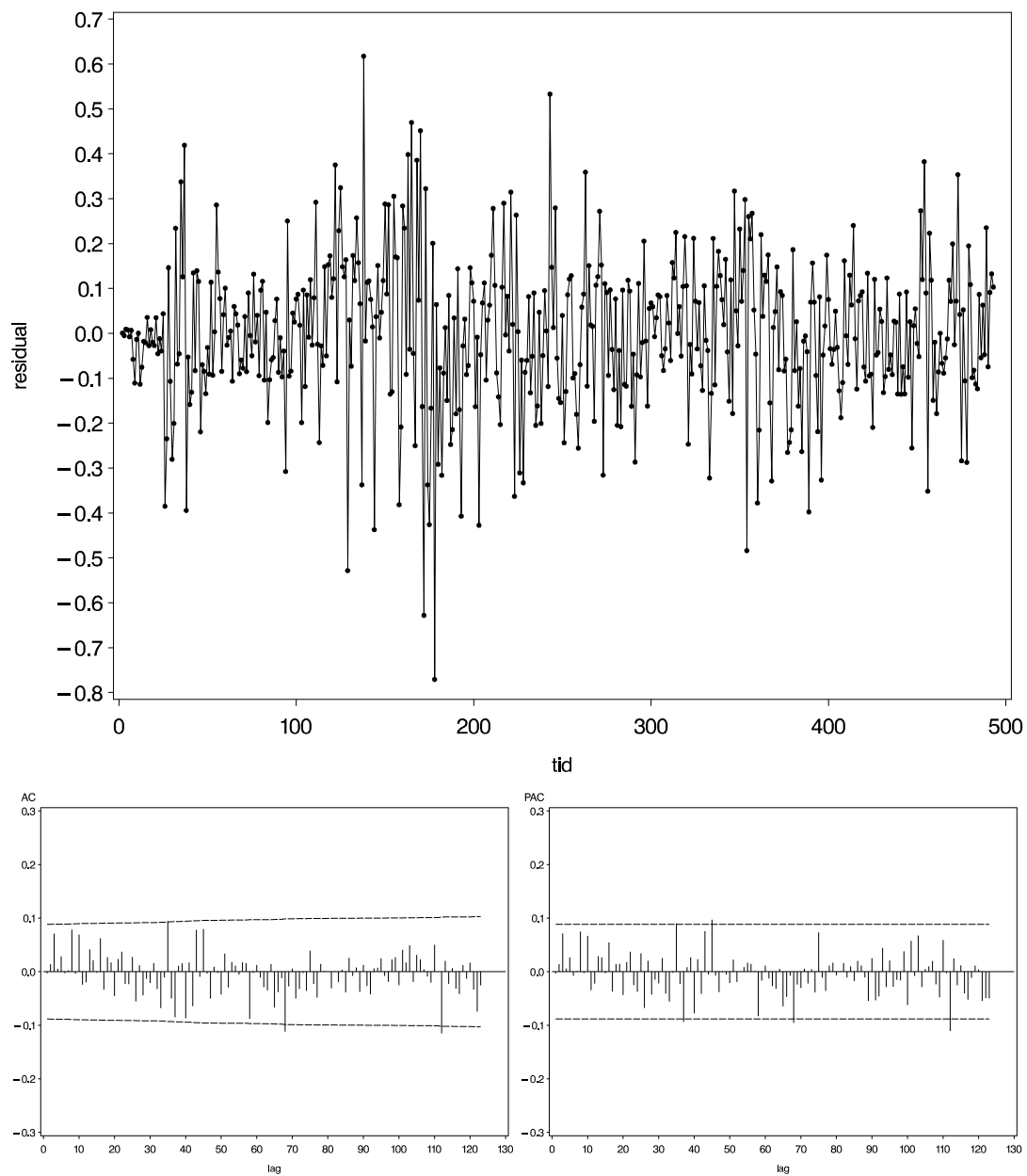
■

Eksempel 3.19. Residualerne fra MA(1)-modellen i formel (3.36) på side 70 for førstedifferenserne for kursen på US\$ er vist i øverst i Figur 3.11 og de tilsvarende empiriske korrelationsfunktioner nederst. Hverken disse figurer eller Ljung-Box χ^2 -teststørrelserne i SAS-udskrift 3.24 giver anledning til at tvivle på MA(1)-modellen for førstedifferenserne for kursen på US\$ eller – ækvivalent hermed – ARIMA(0, 1, 1)-modellen for kurserne.

SAS-udskrift 3.24

Autocorrelation Check of Residuals									
To Lag	Chi - Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	3.09	5	0.6865	-0.003	0.014	0.072	0.006	0.029	-0.002
12	9.12	11	0.6108	0.003	0.079	-0.003	0.069	-0.024	-0.020
18	13.21	17	0.7217	0.042	0.022	-0.002	0.063	-0.034	0.027
24	15.98	23	0.8561	0.017	-0.045	0.024	0.038	-0.023	-0.023
30	19.40	29	0.9107	0.028	-0.055	0.012	-0.044	-0.013	-0.021
36	28.66	35	0.7668	0.016	-0.032	-0.068	-0.011	0.094	-0.050
42	39.13	41	0.5539	-0.085	0.011	0.016	-0.087	0.017	-0.064
48	47.40	47	0.4564	0.078	-0.009	0.080	-0.003	-0.050	0.009
54	49.76	53	0.6012	-0.003	-0.042	0.034	-0.029	0.018	0.011
60	54.49	59	0.6421	-0.006	0.018	0.016	-0.088	-0.002	0.013
66	59.19	65	0.6797	-0.011	-0.029	-0.035	0.014	-0.066	-0.038
72	68.94	71	0.5472	-0.012	-0.112	-0.028	0.006	-0.050	-0.032
78	72.39	77	0.6275	-0.003	-0.036	0.039	-0.023	-0.048	0.014
84	73.19	83	0.7709	0.001	-0.000	-0.031	0.001	-0.019	0.004
90	75.51	89	0.8453	-0.039	0.026	-0.008	0.008	-0.038	0.012
96	77.50	95	0.9045	-0.027	-0.042	0.006	0.007	0.025	-0.008
102	80.11	101	0.9380	-0.019	0.023	0.027	-0.025	0.041	0.017
108	82.84	107	0.9598	0.049	-0.019	0.031	0.023	0.004	-0.009
114	93.82	113	0.9051	-0.021	0.050	0.000	-0.115	0.020	-0.023
120	96.03	119	0.9398	0.006	-0.032	-0.042	0.012	-0.013	0.017

Eksempel 3.20. Residualerne i AR(3)-modellen i formel (3.37) på side 72 for isotopkurven og de tilsvarende empiriske autokorrelationsfunktioner ses i Figur 3.12. Ingen af plottene i figuren antyder afvigelser fra AR(3)-modellen. Dette indtryk bakkes op af Ljung-Box χ^2 -teststørrelserne i SAS-udskrift 3.25, som af pladshensyn kun vises op til lag 30.



Figur 3.11: Øverst residualerne fra MA(1)-modellen i formel (3.36) for førstedifferenserne for kursen på US\$ og nederst de tilsvarende empiriske korrelationsfunktioner.

SAS-udskrift 3.25

Autocorrelation Check of Residuals									
To	Chi -	Pr >							
Lag	Square	DF	ChiSq	-----Autocorrelations-----					
6	2.87	3	0.4120	0.001	-0.006	0.008	-0.032	0.041	0.028
12	11.32	9	0.2542	0.008	-0.021	0.063	0.010	-0.043	0.064
18	23.07	15	0.0827	0.052	-0.003	0.065	-0.031	0.056	0.059
24	26.48	21	0.1888	0.043	0.012	0.016	-0.025	-0.003	0.036
30	27.65	27	0.4290	-0.006	-0.001	-0.018	-0.009	0.001	0.032

Uheldigvis beregner SAS ikke 95 %-konfidensgrænserne for korrelationerne i Figur 3.12 ifølge Sætning 3.17. SAS beregner grænserne som i (3.13) og (3.14), og betragter således blot residualerne som et nyt sæt af observationer.

Vi illustrerer nu brugen af Sætning 3.17 og Bemærkning 3.18 om, at modelkontrol baseret på de empiriske autokorrelationsfunktioner for residualerne kan være konstruktiv i den forstand, at den kan antyde, hvordan en forkert model skal modificeres for at opnå en bedre model.

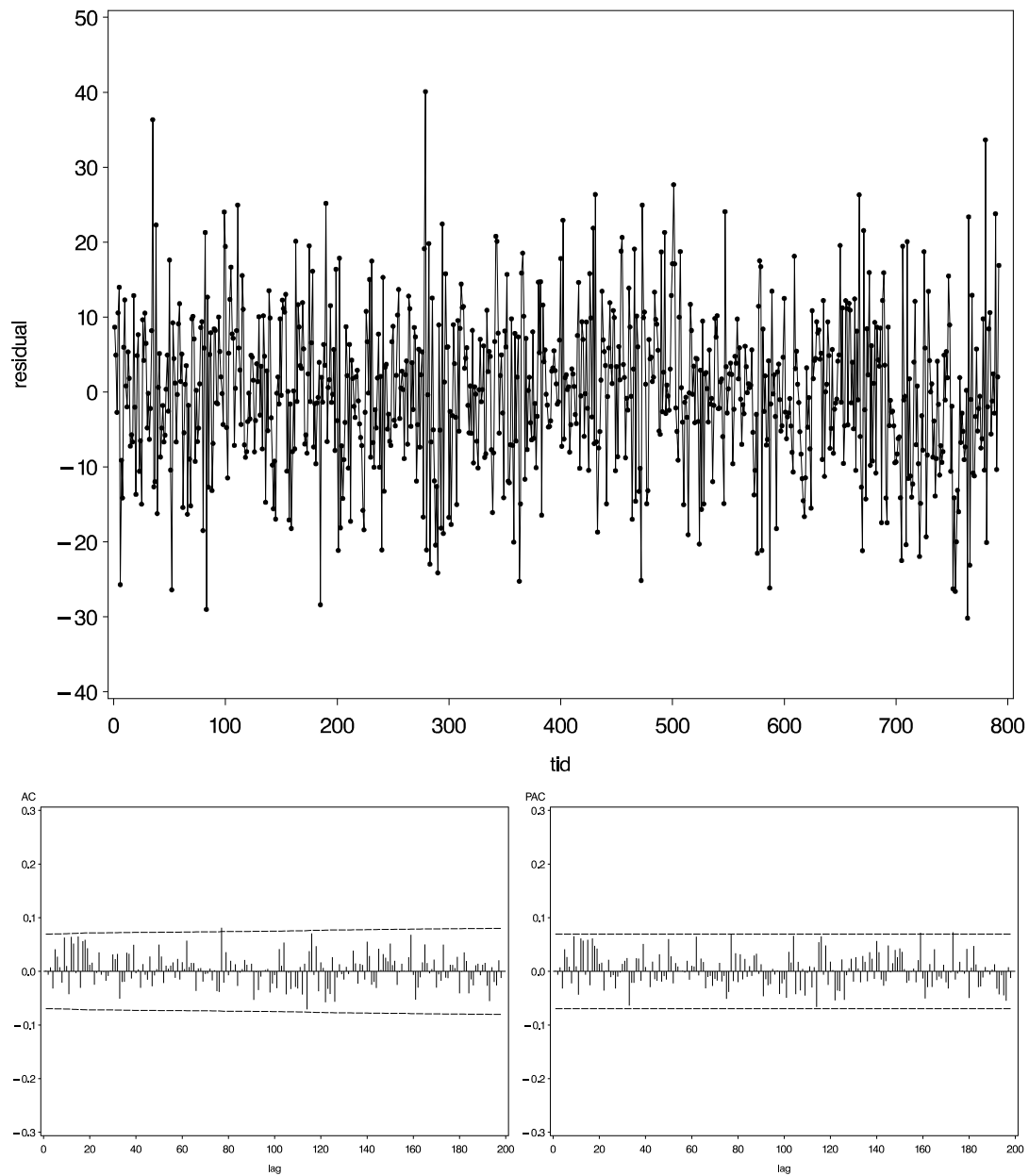
For at gøre dette, lad os vælge at overse at den partielle empiriske autokorrelation $\hat{\phi}_{3,3}$ i Figur 3.3 på side 51 falder udenfor 95 %-konfidensgrænsen, og betragte en AR(1)-model for isotopkurven. Maksimum likelihood estimer i denne model fremgår af SAS-udskrift 3.26.

SAS-udskrift 3.26

Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag
MU	293.33318	0.65457	448.13	<.0001	0
AR1,1	0.43065	0.03216	13.39	<.0001	1
Constant Estimate			167.0106		
Variance Estimate			110.1452		

Residualerne i AR(1)-modellen og de tilsvarende empiriske korrelationsfunktioner er vist i Figur 3.13.

Korrelationsfunktionerne antyder, at AR(1)-modellen er utilstrækkelig, idet $r_2(\hat{w})$ såvel som $\hat{\phi}_{2,2}(\hat{w})$ falder udenfor 95 %-konfidensintervallet. Denne indikation ses også i SAS-udskrift 3.27 med Ljung-Box χ^2 -teststørrelser:



Figur 3.12: Øverst residualerne i AR(3)-modellen i formel (3.37) for isotopkurven og nederst de tilsvarende empiriske korrelationsfunktioner.

SAS-udskrift 3.27

Autocorrelation Check of Residuals									
To	Chi -	Pr >							
Lag	Square	DF	ChiSq	-----Autocorrelations-----					
6	12.18	5	0.0325	0.031	-0.095	0.043	-0.001	0.044	0.038
12	22.47	11	0.0210	-0.001	-0.020	0.073	0.011	-0.049	0.068
18	33.66	17	0.0093	0.054	0.002	0.058	-0.026	0.049	0.066
24	36.98	23	0.0327	0.039	0.014	0.017	-0.025	-0.002	0.038
30	37.81	29	0.1267	-0.004	-0.004	-0.014	-0.010	0.002	0.026

Konfidensgrænserne fra [Sætning 3.17](#) beregnes let for en AR(1)-model. Med notationen fra sætningen har vi at

$$\tilde{\varphi}^{-1}(z) = (1 - \varphi_1 z)^{-1} = \sum_{j=0}^{\infty} \varphi_1^j z^j,$$

dvs. at $\tilde{\psi}_j = \varphi_1^j$, $j = 0, 1, \dots$, så

$$\tilde{\sigma}^2 = \sum_{j=0}^{\infty} \tilde{\psi}_j^2 = \sum_{j=0}^{\infty} \varphi_1^{2j} = \frac{1}{1 - \varphi_1^2}.$$

Da desuden $A = 1$ og $G_m = (1, \tilde{\psi}_1, \dots, \tilde{\psi}_{m-1})^* = (1, \varphi_1^1, \dots, \varphi_1^{m-1})^*$ fås det, at diagonalelementerne i $m \times m$ matricen

$$H_m = (1 - \varphi_1^2) G_m^* G_m$$

bliver

$$(H_m)_{j,j} = (1 - \varphi_1^2) \varphi_1^{2(j-1)}, \quad j = 1, \dots, m.$$

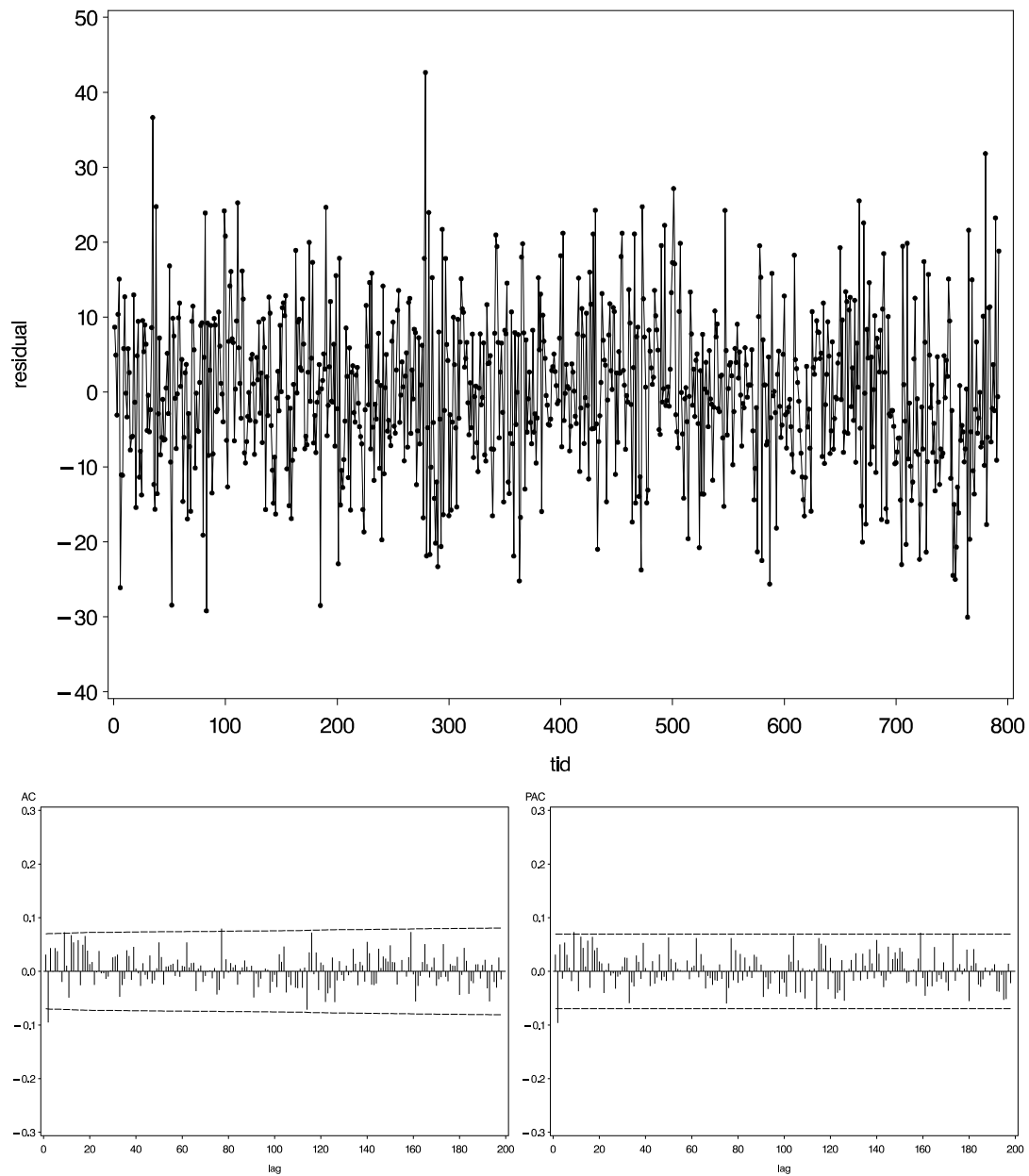
I [Figur 3.14](#) er de tilsvarende 95 %-konfidensintervaller for residualerne i AR(1)-modellen,

$$\left(j, \pm 1.96 \frac{\sqrt{1 - (H_m)_{j,j}}}{\sqrt{N}} \right), \quad j = 1, \dots, 25,$$

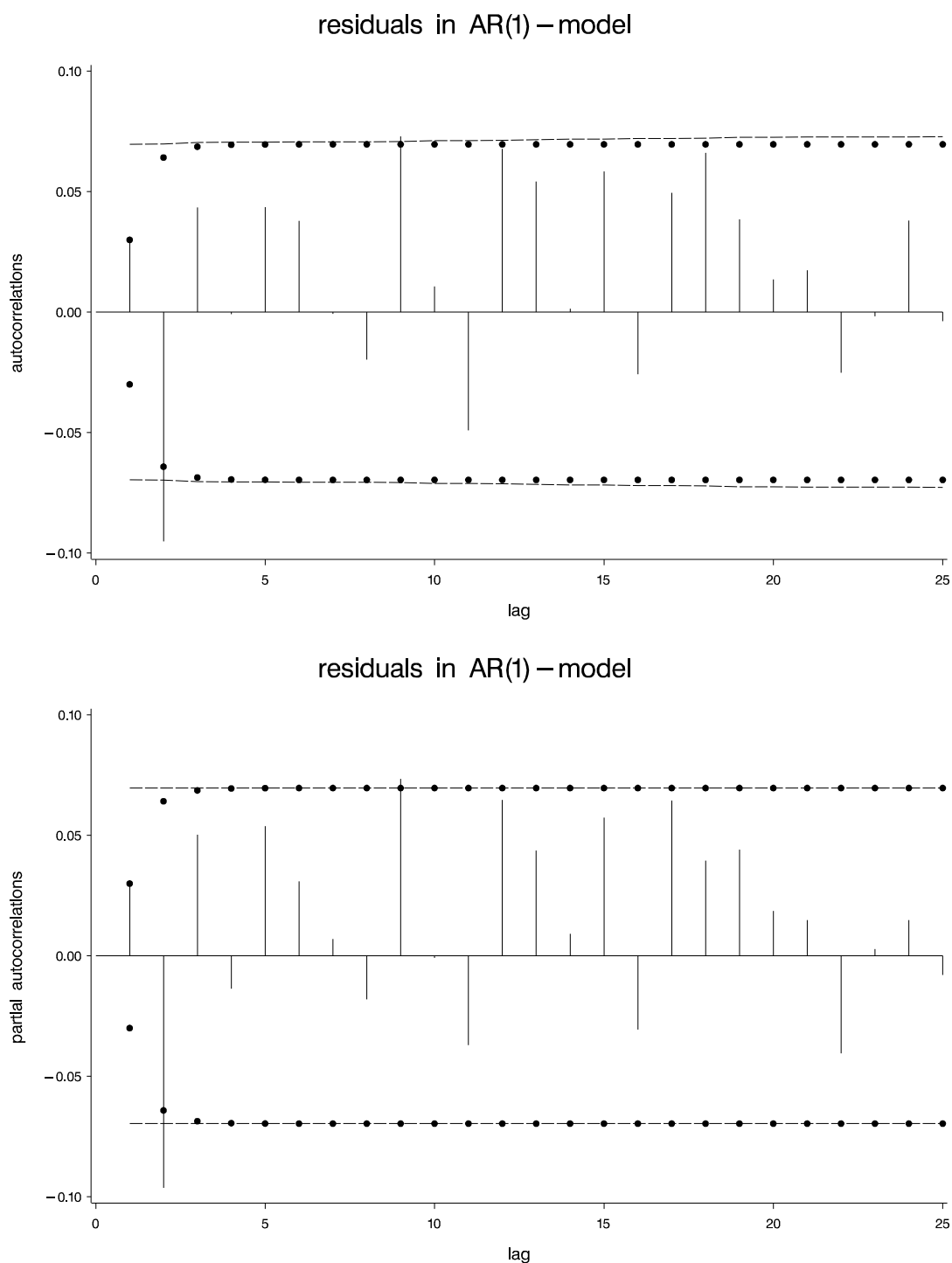
antydnet med prikker.

Vi afslutter her dette eksempel med at vise, hvorledes de empiriske autokorrelationsfunktioner for residualerne kan bruges til af foreslå en ny model i stedet for den ufyldstgørende AR(1)-model. Af [Figur 3.14](#) ses det, at for $k \geq 3$ er de empiriske partielle autokorrelationer $\hat{\phi}_{k,k}(\hat{\mathbf{w}})$ for residualerne fra den estimerede AR(1)-model,

$$(1 - 0.431B)(X_t - \hat{\mu}) = \hat{W}_t, \quad (3.45)$$



Figur 3.13: Øverst residualerne i AR(1)-modellen for isotopkurven og nederst de tilsvarende empiriske korrelationsfunktioner.



Figur 3.14: De empiriske autokorrelationsfunktioner for residualerne i AR(1)-modellen for isotopekurven. Prikkerne antyder de approksimative 95 % konfidensintervaller beregnet ved hjælp af [Sætning 3.17](#). Fra [Andersen og Blæsild \(2003\)](#).

med $\hat{\mu} = 293.33318$, alle lig med 0, hvilket antyder en AR(2)-model for residualerne. Grove estimater for parameterne φ_1 og φ_2 i denne model kan findes fra ligningerne

$$0.031 = \hat{\phi}_{1,1}(\hat{w}) = \frac{\varphi_1}{1 - \varphi_2}$$

og

$$-0.096 = \hat{\phi}_{2,2}(\hat{w}) = \varphi_2.$$

Løses disse fås $(\varphi_1, \varphi_2) = (0.034, -0.096)$, hvilket giver følgende model for residualerne

$$\varphi_W(B)\hat{W}_t = (1 - 0.034B + 0.096B^2)\hat{W}_t = V_t,$$

hvor $\{V_t\}_{t \in \mathbb{Z}}$ er en hvid støj.

Formlerne (3.44) og (3.45) giver

$$(1 - 0.034B + 0.096B^2)(1 - 0.431B)(X_t - \hat{\mu}) = Z_t,$$

eller

$$(1 - 0.465B + 0.111B^2 - 0.041B^3)(X_t - \hat{\mu}) = Z_t,$$

dvs. en AR(3)-model, der til forveksling ligner den estimerede AR(3)-model

$$(1 - 0.468B + 0.112B^2 - 0.088B^3)(X_t - \mu) = W_t$$

med $\mu = 293.33436$.

De næste to eksempler vedrører flerdimensionale tidsrækker, hvor vi indtil nu kun har identificeret og estimeret i de marginale tidsrækker. Da vi senere skal studere disse rækker ved hjælp af flerdimensionale modeller, indskrænker vi her modelkontrollen af de marginale rækker til blot at omfatte Ljung-Box testet.

Eksempel 3.21. I SAS-udskrift 3.20 på side 73 ses estimaterne i ARIMA(4, 1, 0)-modellen for *logaritmen til faste investeringer*. Ljung-Box testene, der ses i SAS-udskrift 3.28, giver ikke anledning til at betvivle modellen.

SAS-udskrift 3.28

Autocorrelation Check of Residuals									
To	Chi -	Pr >							
Lag	Square	DF	ChiSq	-----Autocorrelations-----					
6	0.41	2	0.8144	0.012	0.002	-0.019	0.032	0.051	0.030
12	5.79	8	0.6706	-0.007	-0.126	-0.066	0.058	0.064	-0.179
18	14.01	14	0.4487	-0.187	-0.149	-0.099	-0.104	0.088	0.024
24	18.80	20	0.5351	-0.208	0.026	0.004	-0.007	-0.037	-0.025

For *logaritmen til udgifter til forbrug* ses estimaterne i ARIMA(2, 1, 0)-modellen i SAS-udskrift 3.21. Bedømt ud fra Ljung-Box testene i SAS-udskrift 3.29 giver modellen en lidt tvivlsom beskrivelse af data.

SAS-udskrift 3.29

Autocorrelation Check of Residuals									
To	Chi-	Pr >							
Lag	Square	DF	ChiSq	-----Autocorrelations-----					
6	5.73	4	0.2203	-0.056	0.034	0.162	-0.024	0.161	-0.115
12	18.31	10	0.0500	-0.006	-0.180	-0.012	0.129	-0.288	0.096
18	23.72	16	0.0959	-0.002	-0.082	0.068	-0.138	0.136	-0.078
24	38.71	22	0.0152	0.105	0.004	-0.207	0.187	-0.219	-0.032

Eksempel 3.22. Ljung-Box testene for ARIMA(0,1,2)-modellen og ARIMA(0,1,1)-modellen fra side 74 for renten på henholdsvis 1-årige og 10-årige amerikanske statsobligationer, som ses i [SAS-udskrift 3.30](#) og [SAS-udskrift 3.31](#) nedenfor, forkaster ikke modellerne.

SAS-udskrift 3.30

Autocorrelation Check of Residuals									
To	Chi-	Pr >							
Lag	Square	DF	ChiSq	-----Autocorrelations-----					
6	4.99	4	0.2881	-0.009	0.052	0.005	-0.166	0.024	-0.056
12	10.88	10	0.3671	0.138	0.060	0.049	0.068	0.070	-0.066
18	17.05	16	0.3823	-0.091	0.009	0.039	0.040	0.141	-0.084
24	20.77	22	0.5349	0.047	-0.055	-0.050	-0.027	-0.113	-0.029

SAS-udskrift 3.31

Autocorrelation Check of Residuals									
To	Chi-	Pr >							
Lag	Square	DF	ChiSq	-----Autocorrelations-----					
6	1.01	5	0.9616	-0.005	0.007	0.048	0.030	-0.044	0.041
12	6.10	11	0.8663	0.082	-0.101	0.037	0.084	0.090	0.003
18	18.52	17	0.3565	0.022	-0.106	0.129	0.073	-0.209	0.003
24	25.39	23	0.3307	-0.018	-0.022	-0.030	0.107	-0.101	-0.131

Kapitel 4

VARMA-processer

I dette kapitel skal vi betragte flerdimensionale tidsrækker og introducere klassen af VARMA-processer, som generaliserer ARMA-processerne for endimensionale tidsrækker. Vi starter med at indføre lidt notation for svagt stationære K -dimensionale tidsrækker.

4.1 Svagt stationære K -dimensionale tidsrækker

Hvis $\{X_{it}\}_{t \in \mathbb{Z}}$ er endimensionale tidsrækker for $i = 1, \dots, K$, og hvis \mathbf{X}_t er den K -dimensionale vektor $(X_{1t}, \dots, X_{Kt})^*$, siges $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ at være en K -dimensional tidsrække.

Processen $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ er svagt stationær, hvis middelværdi(vektor)en for \mathbf{X}_t ikke afhænger af t , og hvis kovariansen mellem \mathbf{X}_t og \mathbf{X}_{t-h} ikke afhænger af t . Vi benytter betegnelserne

$$E(\mathbf{X}_t) = \boldsymbol{\mu}$$

og

$$E[(\mathbf{X}_t - \boldsymbol{\mu})(\mathbf{X}_{t-h} - \boldsymbol{\mu})^*] = \Gamma_X(h) = \begin{Bmatrix} \gamma_{11}(h) & \gamma_{12}(h) & \cdots & \gamma_{1K}(h) \\ \gamma_{21}(h) & \gamma_{22}(h) & \cdots & \gamma_{2K}(h) \\ \vdots & \vdots & & \vdots \\ \gamma_{K1}(h) & \gamma_{K2}(h) & \cdots & \gamma_{KK}(h) \end{Bmatrix} \quad (4.1)$$

for lag $h = 0, \pm 1, \pm 2, \dots$. Vi kalder

$$\gamma_{ij}(h) = \text{Cov}(X_{it}, X_{j,t-h}) \quad i, j = 1, \dots, K,$$

for krydskovariansfunktionen mellem $\{X_{it}\}_{t \in \mathbb{Z}}$ og $\{X_{jt}\}_{t \in \mathbb{Z}}$, og vi kalder

$$\gamma_{ii}(h) = \text{Cov}(X_{it}, X_{i,t-h}) \quad i = 1, \dots, K,$$

for autokovariansfunktionen for $\{X_{it}\}_{t \in \mathbb{Z}}$.

Bemærk, at

$$\Gamma_X(-h) = \Gamma_X(h)^*. \quad (4.2)$$

Formel (4.3) nedenfor er en generalisation af formel (2.2) i Sætning 2.8 på side 10. For alle $n \geq 0$ og for alle K -dimensionale vektorer $\mathbf{a}_0, \dots, \mathbf{a}_n$ gælder der, at

$$\sum_{i=0}^n \sum_{j=0}^n \mathbf{a}_i^* \Gamma_X(i-j) \mathbf{a}_j \geq 0. \quad (4.3)$$

Lad D være diagonalmatricen givet ved

$$D = \begin{Bmatrix} \sqrt{\gamma_{11}(0)} & 0 & \cdots & 0 \\ 0 & \sqrt{\gamma_{22}(0)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\gamma_{KK}(0)} \end{Bmatrix}.$$

Korrelationsmatricen svarende til lag h for en svagt stationær proces er givet ved

$$\rho_X(h) = D^{-1} \Gamma_X(h) D^{-1} = \begin{Bmatrix} \rho_{11}(h) & \rho_{12}(h) & \cdots & \rho_{1K}(h) \\ \rho_{21}(h) & \rho_{22}(h) & \cdots & \rho_{2K}(h) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{K1}(h) & \rho_{K2}(h) & \cdots & \rho_{KK}(h) \end{Bmatrix} \quad (4.4)$$

for $h = 0, \pm 1, \pm 2, \dots$, og hvor det ij 'te element

$$\rho_{ij}(h) = \text{Cor}(X_{it}, X_{j,t-h}) = \frac{\gamma_{ij}(h)}{\sqrt{\gamma_{ii}(0)} \sqrt{\gamma_{jj}(0)}}$$

er krydskorrelationsfunktionen for $\{X_{it}\}_{t \in \mathbb{Z}}$ og $\{X_{jt}\}_{t \in \mathbb{Z}}$. Diagonalelementet

$$\rho_{ii}(h) = \frac{\gamma_{ii}(h)}{\gamma_{ii}(0)}, \quad i = 1, 2, \dots, K,$$

i $\rho_X(h)$ kaldes for *autokorrelationsfunktionen* for $\{X_{it}\}_{t \in \mathbb{Z}}$. Da $\Gamma_X(-h) = \Gamma_X(h)^*$ ses det, at $\rho_X(-h) = \rho_X(h)^*$, og dermed at $\rho_{ij}(-h) = \rho_{ji}(h)$.

De partielle autokorrelationer blev i Definition 2.14 på side 12 indført som komponenten $\phi_{k,k}$ i vektoren $\phi_k = (\phi_{k,1}, \dots, \phi_{k,k})$ bestemt ved ligningen $\phi_k = \rho_k P_k^{-1}$ eller ækvivalent hermed i

$$\phi_k = \gamma_k \Gamma_k^{-1}. \quad (4.5)$$

De *partielle autoregressionsmatricer* bliver nu indført ved hjælp af en K -dimensional version af (4.5). Lad nu Γ_k være den symmetriske $kK \times kK$ matris

$$\Gamma_k = \begin{bmatrix} \Gamma_X(0) & \Gamma_X(1) & \cdots & \Gamma_X(k-1) \\ \Gamma_X(-1) & \Gamma_X(0) & \cdots & \Gamma_X(k-2) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_X(1-k) & \Gamma_X(2-k) & \cdots & \Gamma_X(0) \end{bmatrix},$$

hvis elementer er $K \times K$ matricer og lad $K \times K$ matricerne $\Phi_{1,k}, \dots, \Phi_{k,k}$ være bestemt ved

$$[\Phi_{1,k} : \dots : \Phi_{k,k}] = [\Gamma_X(1) : \dots : \Gamma_X(k)] \Gamma_k^{-1}. \quad (4.6)$$

Den sidste af disse matricer $\Phi_{k,k}$ kaldes *den kte partielle autoregressionsmatriks*. Til senere brug noterer vi, at ligningerne i (4.6) kan skrives som

$$[\Gamma_X(1) : \dots : \Gamma_X(k)] = [\Phi_{1,k} : \dots : \Phi_{k,k}] \begin{bmatrix} \Gamma_X(0) & \Gamma_X(1) & \dots & \Gamma_X(k-1) \\ \Gamma_X(-1) & \Gamma_X(0) & \dots & \Gamma_X(k-2) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_X(1-k) & \Gamma_X(2-k) & \dots & \Gamma_X(0) \end{bmatrix}. \quad (4.7)$$

4.2 VAR-processer

For at indføre VARMA-processerne starter vi med at studere VAR-processer, *vektor autoregressive processer*, idet vi senere viser, at de K -dimensionale VARMA-processer kan opfattes som VAR-processer i en højere dimension.

En VAR-proces af orden p , kort en VAR(p)-proces, er en K -dimensional tidsrække $\{X_t\}_{t \in \mathbb{Z}}$, som løser ligningerne

$$X_t = \nu + \Phi_1 X_{t-1} + \dots + \Phi_p X_{t-p} + w_t, \quad t \in \mathbb{Z}, \quad (4.8)$$

hvor Φ_1, \dots, Φ_p er $K \times K$ koefficientmatricer, $\Phi_p \neq 0$, og $\nu = (\nu_1, \dots, \nu_K)^*$ er et K -dimensionalt konstantled. Desuden er processen $\{w_t\}_{t \in \mathbb{Z}}$, hvor $w_t = (w_{1t}, \dots, w_{Kt})^*$, en K -dimensional *hvid støj*, dvs. at

$$E(w_t) = 0, \quad (4.9)$$

$$E(w_t w_t^*) = \Sigma_w, \quad (4.10)$$

$$E(w_s w_t^*) = 0 \quad \text{for } s \neq t. \quad (4.11)$$

I det følgende gøres der rede for, hvornår processen i (4.8) er veldefineret og dens middelværdi og kovariansmatricer bestemmes. Vi betragter indledningsvis en VAR(1)-proces

$$X_t = \nu + \Phi_1 X_{t-1} + w_t, \quad t \in \mathbb{Z}. \quad (4.12)$$

Ifølge [Regneregler \(A.3\)](#) på side 154 i [Appendiks A](#) gælder der, at hvis alle egenverdierne for koefficientmatricen Φ_1 har modulus mindre end 1, vil følgen Φ_1^i , $i = 0, 1, 2, \dots$, være absolut summabel. Da gælder der ifølge [Sætning A.1](#), at den uendelige sum

$$\sum_{i=0}^{\infty} \Phi_1^i w_{t-i}$$

eksisterer i kvadratisk middel. Vi vil derfor betragte værdien af den uendelige sum som grænseværdien i kvadratisk middel.

Regneregel (A.2) giver også, at matricen $I_K - \Phi_1$ er invertibel. Lad $\mu = (I_K - \Phi_1)^{-1} \nu$. Vi viser nu, at

$$X_t = \mu + \sum_{i=0}^{\infty} \Phi_1^i w_{t-i}, \quad t \in \mathbb{Z}, \quad (4.13)$$

løser ligningerne i (4.12). Vi har

$$\begin{aligned} X_t &= \mu + \sum_{i=0}^{\infty} \Phi_1^i w_{t-i} = \mu + \sum_{i=1}^{\infty} \Phi_1^i w_{t-i} + w_t = \mu + \Phi_1 \sum_{i=1}^{\infty} \Phi_1^{i-1} w_{t-i} + w_t \\ &= \mu + \Phi_1 \sum_{j=0}^{\infty} \Phi_1^j w_{(t-1)-j} + w_t = \mu + \Phi_1 (X_{t-1} - \mu) + w_t \\ &= (I_K - \Phi_1) \mu + \Phi_1 X_{t-1} + w_t = \nu + \Phi_1 X_{t-1} + w_t, \end{aligned}$$

hvilket skulle vises.

Ved at bruge Sætning A.3 fås det, at

$$E(X_t) = \mu + \lim_{n \rightarrow \infty} \sum_{i=0}^n \Phi_1^i E(w_{t-i}) = \mu \quad \text{for alle } t, \quad (4.14)$$

da $E(w_t) = 0$ for alle t . Sætning A.3 giver også, at

$$\begin{aligned} \text{Cov}(X_t, X_{t-h}) &= E[(X_t - \mu)(X_{t-h} - \mu)^*] \\ &= \lim_{n \rightarrow \infty} \sum_{i=0}^n \sum_{j=0}^n \Phi_1^i E(w_{t-i} w_{t-h-j}^*) (\Phi_1^j)^*. \end{aligned}$$

Da $E(w_t w_s^*) = 0$ for $s \neq t$, er leddene i dobbeltsummen kun forskellige fra nul, når $i = h + j$. Desuden er $E(w_t w_t^*) = \Sigma_w$ for alle t , så

$$\begin{aligned} \text{Cov}(X_t, X_{t-h}) &= \lim_{n \rightarrow \infty} \sum_{j=0}^n \Phi_1^{h+j} \Sigma_w (\Phi_1^j)^* \\ &= \sum_{i=0}^{\infty} \Phi_1^{h+i} \Sigma_w (\Phi_1^i)^*. \end{aligned}$$

Da koefficientmatricerne er absolut summable, eksisterer kovariansen. Alt-så har vi set, at VAR(1)-processen har endelig middelværdi og kovarians. Det bemærkes også, at middelværdien af X_t ikke afhænger af t og at kovariansen mellem X_t og X_{t-h} ikke afhænger af t , men kun af afstanden i tid mellem de to observationer, dvs.

$$\Gamma_X(h) = \text{Cov}(X_t, X_{t-h}) = \sum_{i=0}^{\infty} \Phi_1^{h+i} \Sigma_w (\Phi_1^i)^*. \quad (4.15)$$

Vi siger, at VAR(1)-processen er *stabil*, såfremt alle egenverdierne for matricen Φ_1 har modulus mindre end 1. Dette er ifølge [Regneregler \(A.4\)](#) i [Appendiks A](#) ækvivalent med, at

$$\det(I_K - \Phi_1 z) = |I_K - \Phi_1 z| \neq 0 \quad \text{for} \quad |z| \leq 1. \quad (4.16)$$

Af formlerne (4.14) og (4.15) ses det, at en stabil VAR(1)-proces er svagt stationær, ifølge [Definition 2.5](#) på [side 9](#).

I (4.13) er processen repræsenteret som et baglæns glidende gennemsnit og (4.13) omtales som VMA-repræsentationen (*vector moving average*) af processen.

Stabilitetsbetingelsen (4.16) er for en VAR(1)-proces, men kan udvides til VAR(p)-processen ved at opskrive denne som en Kp -dimensional VAR(1)-proces og anvende de netop fundne resultater for en VAR(1)-proces på denne og derefter oversætte resultaterne til VAR(p)-processen. Vi laver opskrivningen ved at definere den Kp -dimensionale stokastiske vektor \mathbf{Z}_t og det Kp -dimensionale konstantled \mathbf{V} som

$$\mathbf{Z}_t = \begin{bmatrix} \mathbf{X}_t \\ \mathbf{X}_{t-1} \\ \vdots \\ \mathbf{X}_{t-p+1} \end{bmatrix} \quad \text{og} \quad \mathbf{V} = \begin{bmatrix} \mathbf{v} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Desuden defineres ($Kp \times Kp$)-koefficientmatricen Φ og den Kp -dimensionale stokastiske vektor \mathbf{W}_t som

$$\Phi = \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{p-1} & \Phi_p \\ I_K & 0 & \cdots & 0 & 0 \\ 0 & I_K & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_K & 0 \end{bmatrix} \quad \text{og} \quad \mathbf{W}_t = \begin{bmatrix} \mathbf{w}_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (4.17)$$

Med denne notation kan vi nu opskrive den K -dimensionale VAR(p)-proces i (4.8) som en Kp -dimensional VAR(1)-proces

$$\mathbf{Z}_t = \mathbf{V} + \Phi \mathbf{Z}_{t-1} + \mathbf{W}_t. \quad (4.18)$$

Her $\{\mathbf{W}_t\}_{t \in \mathbb{Z}}$ er en Kp -dimensional hvid støj, fordi

$$\begin{aligned} \mathbb{E}(\mathbf{W}_t) &= 0, \\ \Sigma_W &= \mathbb{E}(\mathbf{W}_t \mathbf{W}_t^*) = \begin{bmatrix} \Sigma_w & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \end{aligned}$$

og

$$E(\mathbf{W}_s \mathbf{W}_t^*) = 0 \quad \text{for } s \neq t.$$

Når vi skriver (4.18) ud, fås

$$\begin{bmatrix} \mathbf{X}_t \\ \mathbf{X}_{t-1} \\ \mathbf{X}_{t-2} \\ \vdots \\ \mathbf{X}_{t-p+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\nu} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{p-1} & \Phi_p \\ I_K & 0 & \cdots & 0 & 0 \\ 0 & I_K & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_K & 0 \end{bmatrix} \begin{bmatrix} \mathbf{X}_{t-1} \\ \mathbf{X}_{t-2} \\ \mathbf{X}_{t-3} \\ \vdots \\ \mathbf{X}_{t-p} \end{bmatrix} + \begin{bmatrix} \mathbf{w}_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

som består af de p ligninger

$$\begin{aligned} \mathbf{X}_t &= \boldsymbol{\nu} + \Phi_1 \mathbf{X}_{t-1} + \cdots + \Phi_p \mathbf{X}_{t-p} + \mathbf{w}_t, \\ \mathbf{X}_{t-1} &= \mathbf{X}_{t-1}, \\ \mathbf{X}_{t-2} &= \mathbf{X}_{t-2}, \\ &\vdots \\ \mathbf{X}_{t-p+1} &= \mathbf{X}_{t-p+1}, \end{aligned}$$

hvor vi ser udtrykket for VAR(p)-processen i (4.8) i den første ligning.

Når den K -dimensionale VAR(p)-proces er skrevet op som en Kp -dimensional VAR(1)-proces, ses det direkte af (4.16), at stabilitetsbetingelsen for VAR(p)-processen er

$$|I_{Kp} - \Phi z| \neq 0 \quad \text{for } |z| \leq 1.$$

Da det karakteristiske polynomium for matricen Φ er $|\Phi - zI_{Kp}|$, kaldes $|I_{Kp} - \Phi z|$ for det *omvendte karakteristiske polynomium*.

Som det fremgår af følgende sætning, kan betingelsen ovenfor udtrykkes ved hjælp af koefficientmatricerne Φ_1, \dots, Φ_p i (4.8).

Sætning 4.1. Der gælder, at

$$|I_{Kp} - \Phi z| = |I_K - \Phi_1 z - \cdots - \Phi_p z^p|. \quad (4.19)$$

Bevis. Se Opgave 11. □

Vi siger derfor, at VAR(p)-processen i (4.8) er stabil, hvis

$$|I_K - \Phi_1 z - \cdots - \Phi_p z^p| \neq 0 \quad \text{for } |z| \leq 1. \quad (4.20)$$

Bemærkning 4.2. Lad $\Phi(z) = I_K - \Phi_1 z - \cdots - \Phi_p z^p$. Betingelsen i (4.20), som sikrer, at VAR(p)-processen i (4.8) er stationær og har en VMA-repræsentation, er da

$$|\Phi(z)| \neq 0 \quad \text{for } |z| \leq 1.$$

I det endimensionale tilfælde var det betingelsen

$$\varphi(z) \neq 0 \quad \text{for } |z| \leq 1,$$

hvor $\varphi(z) = 1 - \varphi_1 z - \dots - \varphi_p z^p$, der medførte, at en $\text{AR}(p)$ -proces var stationær og kunne repræsenteres som et baglæns glidende gennemsnit, se (a) og (b) i [Sætning 2.19](#) på side 19.

Den flerdimensionale version af betingelsen $\varphi(z) \neq 0$ for $|z| \leq 1$ er altså $|\Phi(z)| \neq 0$ for $|z| \leq 1$. ■

Vi antager nu, at betingelsen i (4.20) er opfyldt.

Middelværdien \mathbf{M} for den Kp -dimensionale $\text{VAR}(1)$ -proces er bestemt på samme måde som middelværdien for den K -dimensionale $\text{VAR}(1)$ -proces, nemlig som

$$\mathbf{M} = (I_{Kp} - \Phi)^{-1} \mathbf{V}.$$

Tilsvarende findes kovariansmatricerne til at være

$$\Sigma_Z(h) = \sum_{i=0}^{\infty} \Phi^{h+i} \Sigma_W (\Phi^i)^*,$$

hvor $\Sigma_W = E(\mathbf{W}_t \mathbf{W}_t^*)$, og $h = 0, \pm 1, \pm 2, \dots$.

Vi kan få fat på $\text{VAR}(p)$ -processen ved at gange \mathbf{Z}_t med $(K \times Kp)$ -matricen J , som består af den K -dimensionale enhedsmatriks I_K efterfulgt af $p - 1$ versioner af $K \times K$ matricen med 0'er på alle pladser, dvs.

$$J = [I_K : 0 : \dots : 0], \quad (4.21)$$

så

$$\mathbf{X}_t = J \mathbf{Z}_t. \quad (4.22)$$

Da vi tidligere viste, at $\text{VAR}(1)$ -processen var en veldefineret stokastisk proces, er $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ også veldefineret. Vi finder middelværdi og kovariansmatricerne for $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ ud fra de tilsvarende for $\{\mathbf{Z}_t\}_{t \in \mathbb{Z}}$, idet

$$E \mathbf{X}_t = E(J \mathbf{Z}_t) = J E(\mathbf{Z}_t) = J \mathbf{M},$$

som er en konstant vektor, og

$$\Sigma_X(h) = \text{Cov}(J \mathbf{Z}_t, J \mathbf{Z}_{t-h}) = J \text{Cov}(\mathbf{Z}_t, \mathbf{Z}_{t-h}) J^* = J \Sigma_Z(h) J^*.$$

Fra [side 90](#) ved vi, at når $\text{VAR}(1)$ -processen er stabil, har den VMA-repræsentationen i (4.13). Tilsvarende gælder det for \mathbf{Z}_t , at hvis denne er stabil, så har den VMA-repræsentationen

$$\mathbf{Z}_t = \mathbf{M} + \sum_{i=0}^{\infty} \Phi^i \mathbf{W}_{t-i}.$$

Hvis vi ganger dette med J , findes VMA-repræsentationen for VAR(p)-processen $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ som

$$\mathbf{X}_t = J\mathbf{Z}_t = J\mathbf{M} + J \sum_{i=0}^{\infty} \mathbf{\Phi}^i \mathbf{W}_{t-i}. \quad (4.23)$$

Sættes $\boldsymbol{\mu} = J\mathbf{M}$ og $\Psi_i = J\mathbf{\Phi}^i J^*$, fås

$$\begin{aligned} \mathbf{X}_t &= J\mathbf{M} + J \sum_{i=0}^{\infty} \mathbf{\Phi}^i \mathbf{W}_{t-i} = J\mathbf{M} + \sum_{i=0}^{\infty} J\mathbf{\Phi}^i J^* J \mathbf{W}_{t-i} \\ &= \boldsymbol{\mu} + \sum_{i=0}^{\infty} \Psi_i \mathbf{w}_{t-i}. \end{aligned} \quad (4.24)$$

Her er det brugt, at

$$J^* J \mathbf{W}_t = \begin{bmatrix} I_K & 0 & \dots & 0 \\ 0 & \dots & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w}_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{w}_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{W}_t,$$

og $J\mathbf{W}_t = \mathbf{w}_t$. Desuden bemærkes det, at $\Psi_0 = J\mathbf{\Phi}^0 J^* = I_K$.

På grund af stabilitetsbetingelsen er følgen $\mathbf{\Phi}^i$, $i = 0, 1, \dots$, absolut summabel. Det vil følgen Ψ_i , $i = 1, 2, \dots$, dermed også være. Det fås nu ved hjælp af [Sætning A.3](#), at

$$\mathbb{E}(\mathbf{X}_t) = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \Psi_i \mathbb{E}(\mathbf{w}_{t-i}) = \boldsymbol{\mu} \quad \text{for alle } t, \quad (4.25)$$

og at

$$\begin{aligned} \Gamma_X(h) &= \mathbb{E}[(\mathbf{X}_t - \boldsymbol{\mu})(\mathbf{X}_{t-h} - \boldsymbol{\mu})^*] \\ &= \mathbb{E}\left[\left(\sum_{j=0}^{\infty} \Psi_j \mathbf{w}_{t-j}\right)\left(\sum_{i=0}^{\infty} \Psi_i \mathbf{w}_{t-h-i}\right)^*\right] \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \Psi_j \mathbb{E}(\mathbf{w}_{t-j} \mathbf{w}_{t-h-i}^*) \Psi_i^*. \end{aligned}$$

Da $\mathbb{E}(\mathbf{w}_t \mathbf{w}_s^*) = 0$ for $t \neq s$, fås

$$\Gamma_X(h) = \sum_{i=0}^{\infty} \Psi_{h+i} \Sigma_w \Psi_i^*. \quad (4.26)$$

Koefficientmatricerne Ψ_i $i = 0, 1, \dots$, i processens VMA-repræsentation kan bestemmes på flere måder. De kan findes ud fra koefficientmatricerne Φ_1, \dots, Φ_p fra processens VAR-repræsentation, idet vi jo har defineret $\Psi_i = J\mathbf{\Phi}^i J^*$, $i = 0, 1, \dots$.

Der findes dog en nemmere metode, hvor de bestemmes rekursivt. Denne tager udgangspunkt i, at VAR(p)-processen i (4.8) kan formuleres som

$$\Phi(B)\mathbf{X}_t = \mathbf{v} + \mathbf{w}_t, \quad (4.27)$$

hvor $\Phi(B) = I_K - \Phi_1 B - \dots - \Phi_p B^p$, og $B^i \mathbf{X}_t = \mathbf{X}_{t-i}$. Definér

$$\Psi(B) = \sum_{i=0}^{\infty} \Psi_i B^i$$

således, at

$$\Psi(B)\Phi(B) = I_K. \quad (4.28)$$

Operatoren $\Psi(B)$ kaldes derfor den inverse af $\Phi(B)$, og den betegnes ofte $\Phi(B)^{-1}$. Operatoren $\Phi(B)$ siges at være *invertibel*, hvis

$$|\Phi(z)| = |I_K - \Phi_1 z - \dots - \Phi_p z^p| \neq 0 \quad \text{for } |z| \leq 1,$$

dvs. hvis VAR(p)-processen (4.27) er stabil. Fra [Regneregler \(A.5\)](#) ses det, at den inverse er på formen

$$\Psi(z) = \Phi(z)^{-1} = \frac{1}{|\Phi(z)|} \Phi^{\text{adj}}(z),$$

hvor $\Phi^{\text{adj}}(z)$ er den adjungerede af $\Phi(z)$.

Bemærkningen efter [Sætning A.1](#) i [Appendiks A](#) giver, at koefficientmatricerne Ψ_i i $\Psi(z)$ er absolut summable, når processen er stabil.

Vi ganger nu (4.27) fra venstre med $\Psi(B)$ og får

$$\mathbf{X}_t = \Psi(B)\mathbf{v} + \Psi(B)\mathbf{w}_t = \left(\sum_{i=0}^{\infty} \Psi_i \right) \mathbf{v} + \sum_{i=0}^{\infty} \Psi_i \mathbf{w}_{t-i} = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \Psi_i \mathbf{w}_{t-i},$$

hvor $\boldsymbol{\mu} = \Psi(1)\mathbf{v}$. Her er det brugt, at $B^i \mathbf{v} = \mathbf{v}$, $i = 0, 1, 2, \dots$. Da følgen Ψ_i er absolut summabel giver [Sætning A.1](#), at processen $\Psi(B)\mathbf{w}_t = \sum_{i=0}^{\infty} \Psi_i \mathbf{w}_{t-i}$ er veldefineret.

Koefficientmatricerne Ψ_i bestemmes nu rekursivt ud fra Φ_1, \dots, Φ_p . Skrives (4.28) ud, fås

$$\begin{aligned} I_K &= (\Psi_0 + \Psi_1 B + \Psi_2 B^2 + \dots)(I_K - \Phi_1 B - \dots - \Phi_p B^p) \\ &= \Psi_0 + (\Psi_1 - \Psi_0 \Phi_1)B + (\Psi_2 - \Psi_1 \Phi_1 - \Psi_0 \Phi_2)B^2 + \dots + \left(\Psi_i - \sum_{j=1}^i \Psi_{i-j} \Phi_j \right) B^i + \dots \end{aligned}$$

Heraf ses det, at

$$\begin{aligned} I_K &= \Psi_0, \\ 0 &= \Psi_1 - \Psi_0 \Phi_1, \\ 0 &= \Psi_2 - \Psi_1 \Phi_1 - \Psi_0 \Phi_2, \\ &\vdots \\ 0 &= \Psi_i - \sum_{j=1}^i \Psi_{i-j} \Phi_j, \\ &\vdots \end{aligned}$$

hvor $\Phi_j = 0$ for $j > p$, dvs.

$$\begin{aligned} \Psi_0 &= I_K \\ \Psi_i &= \sum_{j=1}^i \Psi_{i-j} \Phi_j \quad \text{for } i = 1, 2, \dots \end{aligned} \quad (4.29)$$

Specielt gælder der for en VAR(1)-proces, at $\Psi_0 = I_K$, og $\Psi_i = \Phi_1^i$.

4.3 Kovarianser og korrelationer

Vi har i (4.26) set kovariansmatricerne $\Gamma_X(h)$ for en stabil VAR(p)-proces udtrykt ved koefficientmatricerne Ψ_i fra VMA-repræsentationen, idet

$$\Gamma_X(h) = \sum_{i=0}^{\infty} \Psi_{h+i} \Sigma_w \Psi_i^*.$$

Denne formel er dog ikke særlig anvendelig i praksis, da kovariansen her er en uendelig sum. Derfor udtrykker vi nu kovarianserne ved hjælp af de p koefficientmatricer Φ_1, \dots, Φ_p fra VAR-repræsentationen. Først gøres det for en VAR(1)-proces og udvides derefter til en VAR(p)-proces ved at opskrive denne som en VAR(1)-proces.

Så betragt den stabile VAR(1)-proces

$$\mathbf{X}_t = \boldsymbol{\nu} + \Phi_1 \mathbf{X}_{t-1} + \mathbf{w}_t = (I_K - \Phi_1) \boldsymbol{\mu} + \Phi_1 \mathbf{X}_{t-1} + \mathbf{w}_t, \quad (4.30)$$

hvor $E \mathbf{X}_t = \boldsymbol{\mu}$, og den hvide støj har kovariansmatrix Σ_w . Vi kan også formulere denne model ud fra de centrerede variable $\mathbf{X}_t - \boldsymbol{\mu}$ ved at trække $\boldsymbol{\mu}$ fra på begge sider af lighedstegnet i (4.30). Da fås det, at

$$\mathbf{X}_t - \boldsymbol{\mu} = \Phi_1 (\mathbf{X}_{t-1} - \boldsymbol{\mu}) + \mathbf{w}_t.$$

Denne ganges nu med $(X_{t-h} - \mu)^*$ fra højre og derefter tager vi middelværdien. Dette giver

$$\begin{aligned}\Gamma_X(h) &= E[(X_t - \mu)(X_{t-h} - \mu)^*] \\ &= E[(\Phi_1(X_{t-1} - \mu) + w_t)(X_{t-h} - \mu)^*] \\ &= \Phi_1 E[(X_{t-1} - \mu)(X_{t-h} - \mu)^*] + E[w_t(X_{t-h} - \mu)^*].\end{aligned}$$

For $h = 0$ ses, at

$$\begin{aligned}\Gamma_X(0) &= \Sigma_X \\ &= \Phi_1 E[(X_{t-1} - \mu)(X_t - \mu)^*] + E[w_t(X_t - \mu)^*] \\ &= \Phi_1 \Gamma_X(-1) + \Sigma_w \\ &= \Phi_1 \Gamma_X(1)^* + \Sigma_w.\end{aligned}$$

Her er det brugt, at når processen er stabil, kan den udtrykkes som det glidende gennemsnit (4.24), og dermed er

$$\begin{aligned}E[w_t(X_t - \mu)^*] &= E\left[w_t \left(\sum_{i=0}^{\infty} \Psi_i w_{t-i}\right)^*\right] \\ &= E(w_t w_t^*) + E\left[w_t \left(\sum_{i=1}^{\infty} \Psi_i w_{t-i}\right)^*\right] \\ &= \Sigma_w.\end{aligned}$$

Tilsvarende fås det, når $h > 0$, at

$$E[w_t(X_{t-h} - \mu)^*] = 0,$$

da $X_{t-h} - \mu$ kan skrives som et glidende gennemsnit af støjvariable fra tid $t - h$ eller før, og disse er ukorrelerede med w_t . Dette giver, at

$$\begin{aligned}\Gamma_X(h) &= \Phi_1 E[(X_{t-1} - \mu)(X_{t-h} - \mu)^*] \\ &= \Phi_1 E[(X_{t-1} - \mu)(X_{t-1-(h-1)} - \mu)^*] \\ &= \Phi_1 \Gamma_X(h-1).\end{aligned}\tag{4.31}$$

Ligningerne kaldes for *Yule-Walker ligningerne af orden 1*. Så hvis vi kender koeficientmatricen Φ_1 og kovariansmatricen Σ_X , kan vi beregne $\Gamma_X(h)$ rekursivt.

Det ses, at

$$\Gamma_X(h) = \Phi_1 \Gamma_X(h-1) = \Phi_1^2 \Gamma_X(h-2) = \dots = \Phi_1^h \Gamma_X(0)$$

for $h > 0$, og

$$\Gamma_X(0) = \Phi_1 \Gamma_X(1)^* + \Sigma_w = \Phi_1 (\Phi_1 \Gamma_X(0))^* + \Sigma_w = \Phi_1 \Gamma_X(0) \Phi_1^* + \Sigma_w.$$

Vi ser nu på den stabile VAR(p)-proces $\{X_t\}_{t \in \mathbb{Z}}$, hvor vi ligesom før benytter de centrerede variable, dvs.

$$X_t - \mu = \Phi_1(X_{t-1} - \mu) + \cdots + \Phi_p(X_{t-p} - \mu) + w_t.$$

Vi ganger $(X_{t-h} - \mu)^*$ på fra højre og tager middelværdien af udtrykket. Dermed fås

$$\begin{aligned} \Gamma_X(h) &= E[(X_t - \mu)(X_{t-h} - \mu)^*] \\ &= \Phi_1 E[(X_{t-1} - \mu)(X_{t-h} - \mu)^*] + \cdots + \Phi_p E[(X_{t-p} - \mu)(X_{t-h} - \mu)^*] \\ &\quad + E[w_t(X_{t-h} - \mu)^*]. \end{aligned} \quad (4.32)$$

Som ovenfor fås, at når $h = 0$, er

$$\begin{aligned} \Gamma_X(0) &= \Phi_1 \Gamma_X(-1) + \cdots + \Phi_p \Gamma_X(-p) + \Sigma_w \\ &= \Phi_1 \Gamma_X(1)^* + \cdots + \Phi_p \Gamma_X(p)^* + \Sigma_w, \end{aligned} \quad (4.33)$$

og når $h > 0$, er

$$\Gamma_X(h) = \Phi_1 \Gamma_X(h-1) + \cdots + \Phi_p \Gamma_X(h-p). \quad (4.34)$$

Dette er *Yule-Walker ligningerne af orden p* . Bemærk, at i (4.7) er Yule-Walker ligningerne af orden k skrevet op for $h = 1, \dots, k$.

Når vi kender Φ_1, \dots, Φ_p og $\Gamma_X(p-1), \dots, \Gamma_X(0)$, kan vi ud fra formel (4.34) rekursivt beregne $\Gamma_X(h)$ for $h \geq p$. Så først skal vi have bestemt $\Gamma_X(p-1), \dots, \Gamma_X(0)$. Derfor opskrives VAR(p)-processen som en VAR(1)-proces ligesom i (4.18), men udtrykt ved de centrerede variable, dvs.

$$Z_t - M = \Phi(Z_{t-1} - M) + W_t.$$

Vi har netop vist, hvordan kovariansen ser ud for en VAR(1)-proces og får tilsvarende, at

$$\Gamma_Z(0) = \Phi \Gamma_Z(1)^* + \Sigma_W \quad \text{og} \quad \Gamma_Z(1) = \Phi \Gamma_Z(0).$$

Dette giver, at

$$\Gamma_Z(0) = \Phi \Gamma_Z(0) \Phi^* + \Sigma_W, \quad (4.35)$$

hvor

$$\begin{aligned} \Gamma_Z(0) &= E[(Z_t - M)(Z_t - M)^*] \\ &= E \left[\begin{bmatrix} X_t - \mu \\ \vdots \\ X_{t-p+1} - \mu \end{bmatrix} \begin{bmatrix} (X_t - \mu)^*, \dots, (X_{t-p+1} - \mu)^* \end{bmatrix} \right] \\ &= \begin{bmatrix} \Gamma_X(0) & \Gamma_X(1) & \cdots & \Gamma_X(p-1) \\ \Gamma_X(-1) & \Gamma_X(0) & \cdots & \Gamma_X(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_X(-p+1) & \Gamma_X(-p+2) & \cdots & \Gamma_X(0) \end{bmatrix}. \end{aligned} \quad (4.36)$$

Her ses det, at de kovarianser, som vi var interesserede i, er elementerne i den første række af $\Gamma_Z(0)$, som ved hjælp af (4.35) kan udtrykkes ved Φ_1, \dots, Φ_p og Σ_w .

Bemærkning 4.3. Det viser sig, at VAR(p)-processen faktisk er en ret generel proces i den forstand, at der ikke skal lægges mange bånd på en svagt stationær proces for, at den kan beskrives som en VAR-proces. Der gælder nemlig, at enhver svagt stationær proces $\{X_t\}_{t \in \mathbb{Z}}$ kan approksimeres med en VAR-model af endelig orden. Dette er en konsekvens af *Wold's Decomposition Theorem*.

Wold viste, at enhver svagt stationær proces $\{X_t\}_{t \in \mathbb{Z}}$ kan opslittes i to ukorrelerede processer $\{Z_t\}_{t \in \mathbb{Z}}$ og $\{Y_t\}_{t \in \mathbb{Z}}$, som er henholdsvis en deterministisk proces og en VMA-proces, dvs.

$$X_t = Z_t + Y_t,$$

hvor $\{Z_t\}_{t \in \mathbb{Z}}$ er ikke-stokastisk, og

$$Y_t = \sum_{i=0}^{\infty} \Psi_i u_{t-i} = \Psi(B)u_t, \quad (4.37)$$

med $\Psi(B) = \sum_{i=0}^{\infty} \Psi_i B^i$, og hvor $\{u_t\}_{t \in \mathbb{Z}}$ er en hvid støj. Den uendelige sum skal forstås som grænseværdien i kvadratisk middelværdi. I modsætning til hvad vi har set tidligere, er der her ingen antagelse om, at koefficientmatricerne Ψ_i i denne VMA-repræsentation skal være absolut summable.

Vi begrænser os til at se på processer, hvis deterministiske del er middelværdiledet μ , dvs. $Z_t = \mu$ for alle t . Da giver sætningen, at hvis blot den betragtede proces er svagt stationær, så har den VMA-repræsentationen

$$X_t = Z_t + Y_t = \mu + \sum_{i=0}^{\infty} \Psi_i u_{t-i}.$$

Vi antager, at koefficientmatricerne Ψ_i i VMA-repræsentationen er absolut summable. Det antages yderligere, at der eksisterer en operator $\Phi(B)$ med absolut summable koefficientmatricer Φ_i så $\Phi(B)\Psi(B) = I_K$. Det vil sige, at $\Phi(B) = \Psi(B)^{-1}$, så $\Psi(B)$ er invertibel, og VMA-processen kan repræsenteres som en VAR-proces. Da $\Phi(B)$ i modsætning til tidligere kan bestå af uendelig mange led, er det dog ikke sikkert, at processens VAR-repræsentation er af endelig orden. Ved at gange (4.37) med $\Phi(B)$ fra venstre findes VAR-repræsentationen

$$\Phi(B)Y_t = u_t$$

af $\{Y_t\}_{t \in \mathbb{Z}}$. Det fås altså, at

$$Y_t = \sum_{i=1}^{\infty} \Phi_i Y_{t-i} + u_t,$$

hvor

$$\Phi(z) = I_K - \sum_{i=1}^{\infty} \Phi_i z^i = \left(\sum_{i=0}^{\infty} \Psi_i z^i \right)^{-1} \quad \text{for } |z| \leq 1.$$

Vi kan bestemme matricerne Φ_i rekursivt ud fra matricerne Ψ_i på grund af relationen $\Phi(B)\Psi(B) = I_K$. Dette gøres på samme måde, som vi tidligere bestemte Ψ_i ud fra Φ_i i (4.29). Vi antog, at koefficientmatricerne Φ_i var absolut summable. Derfor gælder der, at $\Phi_i \rightarrow 0$ for $i \rightarrow \infty$. Så fra et vist trin, f.eks. for $i > p$, er det rimeligt at antage, at $\Phi_i = 0$. Vi kan dermed approksimere den uendelige ordens VAR-proces med en endelig ordens VAR(p)-proces.

Betragt nu en svagt stationær stokastisk proces $\{\mathbf{Y}_t\}_{t \in \mathbb{Z}}$, der ikke indeholder en deterministisk del. Da giver Wold's Decomposition Theorem, at enhver delproces $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ af denne også vil have en VMA-repræsentation. Når $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ kaldes for en delproces betyder det, at den består af nogle af komponenterne fra $\{\mathbf{Y}_t\}_{t \in \mathbb{Z}}$. F.eks. kunne \mathbf{X}_t bestå af de første M komponenter af \mathbf{Y}_t

$$\mathbf{X}_t = \begin{bmatrix} \mathbf{Y}_{1t} \\ \vdots \\ \mathbf{Y}_{mt} \end{bmatrix} = F \mathbf{Y}_t,$$

hvor F er $(M \times K)$ -matricen $[I_M : 0]$. Da fås det, at

$$E(\mathbf{X}_t) = F E(\mathbf{Y}_t) = F \boldsymbol{\mu},$$

og

$$\Gamma_X(h) = F \Gamma_Y(h) F^*.$$

Processen $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ er dermed svagt stationær, og Wold's sætning giver, at den har en VMA-repræsentation. Dette resultat vil vi få brug for senere. ■

Vi har indtil nu set på K -dimensionale VAR(p)-processer

$$\mathbf{X}_t = \boldsymbol{\nu} + \Phi_1 \mathbf{X}_{t-1} + \cdots + \Phi_p \mathbf{X}_{t-p} + \boldsymbol{\epsilon}_t, \quad t \in \mathbb{Z},$$

hvor $\{\boldsymbol{\epsilon}_t\}_{t \in \mathbb{Z}}$ er en hvid støj. Vi skal nu se på VAR-processer, hvor $\{\boldsymbol{\epsilon}_t\}_{t \in \mathbb{Z}}$ ikke længere er en hvid støj, men en VMA-proces af endelig orden q , dvs.

$$\boldsymbol{\epsilon}_t = \mathbf{w}_t - \Theta_1 \mathbf{w}_{t-1} - \cdots - \Theta_q \mathbf{w}_{t-q}, \quad t \in \mathbb{Z}, \quad (4.38)$$

hvor $\Theta_1, \dots, \Theta_q$ er $(K \times K)$ -koefficientmatricer, og $\{\mathbf{w}_t\}_{t \in \mathbb{Z}}$ er en hvid støj med middelværdi nul og kovariansmatriks Σ_w . Repræsentationen af den K -dimensionale proces $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ har altså både en VAR(p)-del og en VMA(q)-del. Vi kalder sådanne processer for VARMA-processer (*vector autoregressive moving average*) af orden p og q , kort VARMA(p, q)-processer. En VARMA(p, q)-proces $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ er således en løsning til ligningerne

$$\mathbf{X}_t = \boldsymbol{\nu} + \Phi_1 \mathbf{X}_{t-1} + \cdots + \Phi_p \mathbf{X}_{t-p} + \mathbf{w}_t - \Theta_1 \mathbf{w}_{t-1} - \cdots - \Theta_q \mathbf{w}_{t-q}, \quad t \in \mathbb{Z}. \quad (4.39)$$

Bemærkning 4.4. En VARMA($p, 0$)-proces er en VAR(p)-proces, mens man i litteraturen kan se betegnelserne VMA(q) og MA(q) for VARMA($0, q$)-processen. ■

4.4 VMA-processer

Tidligere har vi set på VMA-processer af uendelig orden, men nu skal vi se på VMA-processer (4.38) af endelig orden q . Vi starter med at se på egenskaber ved VMA(1)-processen og antager uden tab af generalitet, at $\nu = 0$, dvs. vi betragter processen

$$w_t = X_t + \Theta_1 w_{t-1}, \quad t \in \mathbb{Z}, \quad (4.40)$$

eller

$$X_t = (I_K - \Theta_1 B)w_t, \quad t \in \mathbb{Z}. \quad (4.41)$$

En VMA(1)-proces er per definition altid en stationær proces. Det følgende viser, at en VMA(1)-proces undertiden også kan betragtes som en VAR-proces af uendelig orden.

Hvis alle egenverdierne for Θ_1 har modulus mindre end 1, gælder ifølge [Regneregler \(A.3\)](#) i [Appendiks A](#), at følgen af Θ_1^i , $i = 0, 1, 2, \dots$, er absolut summabel. Da fås det fra [Sætning A.1](#), at den uendelige sum

$$\sum_{i=0}^{\infty} \Theta_1^i X_{t-i}$$

eksisterer i kvadratisk middel. Vi viser nu, at

$$w_t = \sum_{i=0}^{\infty} \Theta_1^i X_{t-i}, \quad t \in \mathbb{Z}, \quad (4.42)$$

løser ligningerne (4.40), idet

$$\begin{aligned} w_t &= \sum_{i=0}^{\infty} \Theta_1^i X_{t-i} \\ &= X_t + \Theta_1 \sum_{i=1}^{\infty} \Theta_1^{i-1} X_{t-i} \\ &= X_t + \Theta_1 \sum_{j=0}^{\infty} \Theta_1^j X_{t-1-j} \\ &= X_t + \Theta_1 w_{t-1}. \end{aligned}$$

Repræsentationen af processen i (4.42) er en VAR-repræsentation af uendelig orden.

Betingelsen, at alle egenverdierne for Θ_1 skal have modulus mindre end 1, er ifølge [Regneregul \(A.4\)](#) ækvivalent med, at

$$|I_K - \Theta_1 z| \neq 0 \quad \text{for } |z| \leq 1. \quad (4.43)$$

Denne betingelse, som er analog med stabilitetsbetingelsen for en VAR(1)-proces, sikrer altså, at processen i (4.42) er veldefineret, så hvis stabilitetsbetingelsen (4.43) er opfyldt, har VMA(1)-processen en VAR-repræsentation af uendelig orden. Dette er ækvivalent med, hvad vi så for VAR(1)-processen tidligere. Denne havde en VMA-repræsentation af uendelig orden, når (4.16) på side 91 var opfyldt. Vi kunne udvide dette resultat til at gælde for en VAR(p)-proces, hvis stabilitetsbetingelsen (4.20) på side 92 var opfyldt.

Vi kan på tilsvarende måde udvide det netop viste til at gælde for VMA(q)-processen

$$\mathbf{X}_t = \mathbf{w}_t - \Theta_1 \mathbf{w}_{t-1} - \cdots - \Theta_q \mathbf{w}_{t-q} \quad (4.44)$$

med middelværdi nul. Denne vil have en VAR-repræsentation af uendelig orden,

$$\sum_{i=0}^{\infty} \Pi_i \mathbf{X}_{t-i} = \mathbf{w}_t, \quad (4.45)$$

hvis betingelsen

$$|I_K - \Theta_1 z - \cdots - \Theta_q z^q| \neq 0 \quad \text{for } |z| \leq 1 \quad (4.46)$$

er opfyldt. Er dette tilfældet, kaldes VMA(q)-processen for *invertibel*.

Vi kan finde koefficientmatricerne Π_i hørende til VAR-repræsentationen rekursivt ud fra koefficientmatricerne $\Theta_1, \dots, \Theta_q$. Først skriver vi (4.44) som

$$\mathbf{X}_t = \Theta(B) \mathbf{w}_t,$$

hvor $\Theta(B) = I_K - \Theta_1 B - \cdots - \Theta_q B^q$ kaldes for VMA-operatoren. Denne operator er invertibel, hvis (4.46) er opfyldt. Vi antager, at dette er tilfældet. Da eksisterer $\Theta(B)^{-1}$, og vi finder VAR-repræsentationen som

$$\Theta(B)^{-1} \mathbf{X}_t = \mathbf{w}_t.$$

Lad $\Pi(B) = I_K + \Pi_1 B + \Pi_2 B^2 + \cdots$ være operatoren, der opfylder, at

$$\Pi(B) \Theta(B) = I_K,$$

dvs. $\Pi(B) = \Theta(B)^{-1}$, og vi ser, at

$$\Pi(B) \mathbf{X}_t = \mathbf{w}_t.$$

hvor $\Pi(B) \mathbf{X}_t = \sum_{i=0}^{\infty} \Pi_i \mathbf{X}_{t-i}$ er veldefineret, da $\Theta(B)$ er invertibel.

Vi finder koefficientmatricerne Π_i , $i = 1, 2, \dots$, ud fra relationen

$$\Pi(B)\Theta(B) = (I_K + \Pi_1 B + \Pi_2 B^2 + \dots)(I_K - \Theta_1 B - \Theta_2 B^2 - \dots - \Theta_q B^q) = I_K.$$

Vi bestemmer matricerne Π_i , $i = 1, 2, \dots$ på samme måde, som vi fandt Ψ_i , $i = 1, 2, \dots$, ud fra Φ_1, \dots, Φ_p i (4.29) på side 96. Det ses, at

$$\Pi_1 = \Theta_1 \quad \text{og} \quad \Pi_i = \Theta_i + \sum_{j=1}^{i-1} \Pi_{i-j} \Theta_j$$

for $i = 2, 3, \dots$, hvor vi sætter $\Theta_j = 0$ for $j > q$.

Vi så tidligere i (4.26) på side 94, at kovariansmatricerne for det glidende gennemsnit $X_t = \sum_{i=0}^{\infty} \Psi_i w_{t-i}$, er bestemt ved

$$\Gamma_X(h) = \sum_{i=0}^{\infty} \Psi_{h+i} \Sigma_w \Psi_i^*.$$

Sammenlignes dette med VMA(q)-processen i (4.44), fås det, at

$$\Gamma_X(h) = E(X_t X_{t-h}^*) = \begin{cases} \sum_{i=0}^q \Theta_i \Sigma_w \Theta_i^*, & h = 0 \\ -\Theta_h \Sigma_w + \sum_{i=1}^{q-h} \Theta_{h+i} \Sigma_w \Theta_i^*, & h = 1, \dots, q, \\ 0, & h = q+1, q+2, \dots, \end{cases} \quad (4.47)$$

hvor det er brugt, at $\Theta_i = 0$ for $i > q$. Vi kan se, at $\Gamma_X(h)$ ikke afhænger af tidspunktet t , og da middelværdien $E(X_t) = 0$ for alle t , er processen i (4.44) stationær. Det bemærkes også, at X_t og X_{t-h} er ukorrelerede for $h > q$.

4.5 VAR- og VMA-repræsentationer af VARMA-processen

Vi skal nu se nærmere på VARMA(p, q)-processen $\{X_t\}_{t \in \mathbb{Z}}$ fra (4.39), når både $p > 0$ og $q > 0$ og se, at processen kan opskrives dels som et glidende gennemsnit af uendelig orden og dels som en autoregression af uendelig orden.

Lad

$$\Phi(B) = I_K - \Phi_1 B - \dots - \Phi_p B^p$$

og

$$\Theta(B) = I_K - \Theta_1 B - \dots - \Theta_q B^q.$$

Processen kan da skrives som

$$\Phi(B)X_t = v + \Theta(B)w_t, \quad (4.48)$$

Vi siger, at VARMA(p, q)-processen er stabil, hvis VAR(p)-delen er stabil, så $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ er stabil, hvis

$$|\Phi(z)| = |I_K - \Phi_1 z - \dots - \Phi_p z^p| \neq 0 \quad \text{for } |z| \leq 1. \quad (4.49)$$

Vi antager, at dette er tilfældet. Da er $\Phi(B)^{-1}$ veldefineret, kan vi gange ligningen i (4.48) fra venstre med $\Phi(B)^{-1}$. Vi får

$$\mathbf{X}_t = \Phi(B)^{-1} \boldsymbol{\nu} + \Phi(B)^{-1} \Theta(B) \mathbf{w}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \Psi_i \mathbf{w}_{t-i}, \quad (4.50)$$

hvor

$$\Phi(B)^{-1} \Theta(B) = \Psi(B) = \sum_{i=0}^{\infty} \Psi_i B^i \quad (4.51)$$

og $\Phi(B)^{-1} \boldsymbol{\nu} = \Phi(1)^{-1} \boldsymbol{\nu} = \boldsymbol{\mu}$. Her er det brugt, at operatoren B ikke virker på en konstant $\boldsymbol{\nu}$, så derfor er $\Phi(B)^{-1} \boldsymbol{\nu} = \Phi(1)^{-1} \boldsymbol{\nu}$.

Vi kan bestemme koefficientmatricerne Ψ_i fra VARMA-processens VMA-repræsentation af uendelig orden i (4.50) rekursivt ud fra Φ_1, \dots, Φ_p og $\Theta_1, \dots, \Theta_q$. Af (4.51) ses, at

$$\begin{aligned} I_K - \Theta_1 B - \dots - \Theta_q B^q &= \Theta(B) \\ &= \Phi(B) \Psi(B) \\ &= (I_K - \Phi_1 B - \dots - \Phi_p B^p) \left(\sum_{i=0}^{\infty} \Psi_i B^i \right) \\ &= \sum_{i=0}^{\infty} \Psi_i B^i - \Phi_1 \sum_{i=0}^{\infty} \Psi_i B^{i+1} - \dots - \Phi_p \sum_{i=0}^{\infty} \Psi_i B^{i+p} \\ &= I_K + \sum_{i=1}^{\infty} \left(\Psi_i - \sum_{j=1}^i \Phi_j \Psi_{i-j} \right) B^i. \end{aligned}$$

Herfra fås det for $i = 1, 2, \dots$, at

$$-\Theta_i = \Psi_i - \sum_{j=1}^i \Phi_j \Psi_{i-j},$$

hvor $\Psi_0 = I_K$, og vi sætter $\Phi_j = 0$ for $j > p$, og $\Theta_j = 0$ for $j > q$. Vi finder så koefficientmatricerne Ψ_i rekursivt ved formlen

$$\Psi_i = -\Theta_i + \sum_{j=1}^i \Phi_j \Psi_{i-j}, \quad \text{for } i = 1, 2, \dots \quad (4.52)$$

Vi har altså set, at hvis VARMA(p, q)-processen er stabil, har den en VMA-repræsentation af uendelig orden, hvis koefficientmatricer er givet ved (4.52).

Vi skal nu se, at hvis VARMA(p, q)-processen er invertibel, dvs. hvis VMA-operatoren $\Theta(B)$ er invertibel, så har processen også en VAR-repræsentation af uendelig orden. Vi betragter igen VARMA(p, q)-processen på formen i (4.48). Da processen var invertibel, er $\Theta(B)^{-1}$ veldefineret. Vi lader $\Theta(B)^{-1}\Phi(B) = \Pi(B) = \sum_{i=0}^{\infty} \Pi_i B^i$ og ved at gange $\Theta(B)^{-1}$ på (4.48) fra venstre fås,

$$\sum_{i=0}^{\infty} \Pi_i X_{t-i} = \Theta(B)^{-1}\Phi(B)X_t = \Theta(1)^{-1}\mathbf{v} + \mathbf{w}_t. \quad (4.53)$$

Vi skal nu have bestemt koefficientmatricerne Π_i . Dette gøres ved at sammenligne koefficienterne i operatorerne $\Phi(B)$ og $\Theta(B)\Pi(B)$, idet der gælder, at

$$(I_K - \Theta_1 B - \dots - \Theta_q B^q) \left(I_K + \sum_{i=1}^{\infty} \Pi_i B^i \right) = \Theta(B)\Pi(B) = \Phi(B) = (I_K - \Phi_1 B - \dots - \Phi_p B^p).$$

Heraf fås det, at

$$\Pi_i = -\Phi_i + \Theta_i + \sum_{j=1}^{i-1} \Theta_{i-j} \Pi_j \quad (4.54)$$

for $i = 1, 2, \dots$. Her er $\Theta_0 = I_K$, og vi sætter igen $\Phi_j = 0$ for $j > p$, og $\Theta_j = 0$ for $j > q$. Desuden sættes $\sum_{j=1}^{i-1} \Theta_{i-j} \Pi_j = 0$, hvis $i < 2$.

Vi så i Afsnit 4.2, hvordan en K -dimensional VAR(p)-proces kan skrives som en Kp -dimensional VAR(1)-proces. Dette var nyttigt, da vi så nemt kunne udvide viste resultater for en VAR(1)-proces til at gælde for en VAR(p)-proces. På tilsvarende måde kan en K -dimensional VARMA(p, q)-proces udtrykkes som en $K(p+q)$ -dimensional VAR(1)-proces, se Opgave 14.

4.6 Kovarianser for en VARMA(p, q)-proces

I dette afsnit betragtes den stabile VARMA(p, q)-proces

$$\mathbf{X}_t = \Phi_1 \mathbf{X}_{t-1} + \dots + \Phi_p \mathbf{X}_{t-p} + \mathbf{w}_t - \Theta_1 \mathbf{w}_{t-1} - \dots - \Theta_q \mathbf{w}_{t-q}, \quad (4.55)$$

som har middelværdi $\boldsymbol{\mu} = 0$. Vi vil bestemme dennes kovariansmatricer $\Gamma_X(h) = E(\mathbf{X}_t \mathbf{X}_{t-h}^*)$ for $h = 0, \pm 1, \pm 2, \dots$. Dette er ikke særlig vanskeligt, idet vi lige har vist, at når $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ er stabil, kan processen udtrykkes som det glidende gennemsnit af uendelig orden i (4.50). På side 94 så vi i formel (4.26), hvordan man bestemmer $\Gamma_X(h)$ ud fra koefficientmatricerne Ψ_i i et sådant glidende gennemsnit. Dermed har vi, at

$$\Gamma_X(h) = \sum_{i=0}^{\infty} \Psi_{h+i} \Sigma_w \Psi_i^*.$$

I praksis er det dog ikke særlig anvendeligt at bestemme $\Gamma_X(h)$ ud fra denne formel. I stedet for bruger man

$$\Gamma_X(h) = \Phi_1 \Gamma_X(h-1) + \cdots + \Phi_p \Gamma_X(h-p), \quad \text{for } h > q. \quad (4.56)$$

Denne formel er fremkommet ved at gange (4.55) med \mathbf{X}_{t-h}^* fra højre og derefter tage middelværdi, dvs.

$$\begin{aligned} E(\mathbf{X}_t \mathbf{X}_{t-h}^*) &= \Phi_1 E(\mathbf{X}_{t-1} \mathbf{X}_{t-h}^*) + \cdots + \Phi_p E(\mathbf{X}_{t-p} \mathbf{X}_{t-h}^*) \\ &\quad + E(\mathbf{w}_t \mathbf{X}_{t-h}^*) - \Theta_1 E(\mathbf{w}_{t-1} \mathbf{X}_{t-h}^*) - \cdots - \Theta_q E(\mathbf{w}_{t-q} \mathbf{X}_{t-h}^*). \end{aligned}$$

Da vi kan skrive \mathbf{X}_t som et glidende gennemsnit af elementer fra den hvide støj op til og med tid t , gælder der, at $E(\mathbf{w}_t \mathbf{X}_s^*) = 0$ for $s < t$. Dette medfører, at

$$E(\mathbf{w}_{t-q} \mathbf{X}_{t-h}^*) = 0$$

for $h > q$, og vi har (4.56). Denne formel kan bruges til at beregne $\Gamma_X(p), \Gamma_X(p+1), \dots$, hvis $p > q$, og vi kender $\Gamma_X(0), \dots, \Gamma_X(p-1)$.

Vi finder korrelationsmatricerne $\rho_X(h)$ ligesom tidligere ved hjælp af formel (4.4) på side 88.

Bemærk, at (4.56) kan skrives på formen

$$\Phi(B) \Gamma_X(h) = 0, \quad \text{for } h > q. \quad (4.57)$$

Specielt for en $\text{VMA}(q) = \text{VARMA}(0, q)$ er $\Phi(B) = I_k$, så $\Gamma_X(h) = 0$ for $h > q$.

4.7 Forecasting

Et vigtigt aspekt i arbejdet med tidsrækker er forecasting. Hvis vi kender parametrene i en model for tidsrækken $\mathbf{X} = \{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ og kender værdierne af processens variable op til og med et bestemt tidspunkt t , kan vi på baggrund af disse give kvalificerede bud på fremtidige værdier af processen. Forecasting er derfor meget nyttigt inden for f.eks. økonomi, hvor man gerne vil have et godt billede af et bestemt markeds fremtid, inden man foretager en investering. Inden for hydrologien kunne man f.eks. være interesseret i at undersøge, hvordan nedbørsmængderne kan forventes at udvikle sig på baggrund af, hvad man har observeret indtil nu.

Vi antager, at vi kender værdierne, $\mathbf{X}_{\leq t} = \{\mathbf{X}_s : s \leq t\}$, af \mathbf{X} op til og med tid t . Vi ønsker at finde den mest optimale prediktor, dvs. det bedste estimat, af den fremtidige observation \mathbf{X}_{t+h} for $h > 0$. En h -trins prediktor betegnes $\tilde{\mathbf{X}}_t(h)$, hvor det nedre t angiver, at til tidspunkt t laves et forecast h tidsenheder frem. Prediktoren skal være en funktion af de allerede observerede værdier, altså en funktion af elementerne i $\mathbf{X}_{\leq t}$.

Vi ønsker at finde en prediktor $\bar{X}_t(h)$, således at prediktionsfejlen $X_{t+h} - \bar{X}_t(h)$, dvs. forskellen mellem observationen X_{t+h} til tidspunkt $t+h$ og den predikterede værdi $\bar{X}_t(h)$ er mindst mulig. Det synes naturligt at betragte *mean squared error* (MSE) matricen,

$$\text{MSE}[\bar{X}_t(h)] = E[(X_{t+h} - \bar{X}_t(h))(X_{t+h} - \bar{X}_t(h))^* | X_{\leq t}]$$

som et mål for, hvor stor forskellen er mellem den predikterede og den observerede værdi af processen til tid $t+h$.

Definition 4.5. *Minimum mean squared prediktoren* er den prediktor $\hat{X}_t(h)$, således at det for alle prediktorer $\bar{X}_t(h)$ gælder, at

$$\text{MSE}[\bar{X}_t(h)] \geq \text{MSE}[\hat{X}_t(h)],$$

dvs. at $\text{MSE}[\bar{X}_t(h)] - \text{MSE}[\hat{X}_t(h)]$ er en ikke-negativ definit matris.

Bemærkning 4.6. Minimum mean squared prediktoren $\hat{X}_t(h)$ er med andre ord den prediktor, der har “mindst varians”. ■

Sætning 4.7. Betragt en proces $\{X_t\}_{t \in \mathbb{Z}}$ med VMA-repræsentationen

$$X_t = \sum_{i=0}^{\infty} \Psi_i w_{t-i},$$

hvor $\Psi_0 = I_k$. Antag, at

- (i) der er en 1-1 sammenhæng mellem $X_{\leq t}$ og $w_{\leq t}$
- (ii) w_s og w_t er uafhængige for $s \neq t$.

Så er minimum mean squared prediktoren $\hat{X}_t(h)$ den betingede middelværdi af X_{t+h} givet $X_{\leq t}$, dvs.

$$\hat{X}_t(h) = E(X_{t+h} | X_{\leq t}). \quad (4.58)$$

Bemærkning 4.8. Betingelsen (i) er opfyldt for stabile og invertible VARMA-processer. ■

Bevis. Betingelserne (i) og (ii) medfører, at

$$E(w_{t+h} | X_{\leq t}) = E(w_{t+h} | w_{\leq t}) = \begin{cases} w_{t+h}, & \text{hvis } h \leq 0 \\ 0, & \text{hvis } h > 0. \end{cases} \quad (4.59)$$

Da

$$X_{t+h} = \sum_{i=0}^{\infty} \Psi_i w_{t+h-i}$$

medfører (4.59), at

$$\hat{X}_t(h) = E(X_{t+h} | X_{\leq t}) = \sum_{i=0}^{\infty} \Psi_i E(w_{t+h-i} | X_{\leq t}) = \sum_{i=h}^{\infty} \Psi_i w_{t+h-i} \quad (4.60)$$

ses det, at forecastfejlen

$$X_{t+h} - \hat{X}_t(h) = \sum_{i=0}^{h-1} \Psi_i w_{t+h-i} \quad (4.61)$$

er uafhængig af $w_{\leq t}$ og dermed af $X_{\leq t}$ og at

$$E(X_{t+h} - \hat{X}_t(h)) = \sum_{i=0}^{h-1} \Psi_i E(w_{t+h-i}) = 0.$$

Da endvidere $\hat{X}_t(h) - \bar{X}_t(h)$ kun afhænger af $w_{\leq t}$ ses, at

$$E[(X_{t+h} - \hat{X}_t(h))(\hat{X}_t(h) - \bar{X}_t(h))^* | X_{\leq t}] = E[X_{t+h} - \hat{X}_t(h)](\hat{X}_t(h) - \bar{X}_t(h))^* = 0. \quad (4.62)$$

Ved hjælp af (4.62) fås, at

$$\begin{aligned} \text{MSE}[\bar{X}_t(h)] &= E[(X_{t+h} - \bar{X}_t(h))(X_{t+h} - \bar{X}_t(h))^* | X_{\leq t}] \\ &= E\left[(X_{t+h} - \hat{X}_t(h) + \hat{X}_t(h) - \bar{X}_t(h))(X_{t+h} - \hat{X}_t(h) + \hat{X}_t(h) - \bar{X}_t(h))^* | X_{\leq t}\right] \\ &= E[(X_{t+h} - \hat{X}_t(h))(X_{t+h} - \hat{X}_t(h))^* | X_{\leq t}] \\ &\quad + E[(X_{t+h} - \hat{X}_t(h))(\hat{X}_t(h) - \bar{X}_t(h))^* | X_{\leq t}] \\ &\quad + E[(\hat{X}_t(h) - \bar{X}_t(h))(X_{t+h} - \hat{X}_t(h))^* | X_{\leq t}] \\ &\quad + E[(\hat{X}_t(h) - \bar{X}_t(h))(\hat{X}_t(h) - \bar{X}_t(h))^* | X_{\leq t}] \\ &= \text{MSE}[\hat{X}_t(h)] + (\hat{X}_t(h) - \bar{X}_t(h))(\hat{X}_t(h) - \bar{X}_t(h))^*. \end{aligned}$$

Da det endvidere for alle $z \in \mathbb{R}^K$ gælder, at

$$z^*(\hat{X}_t(h) - \bar{X}_t(h))(\hat{X}_t(h) - \bar{X}_t(h))^* z = \|(\hat{X}_t(h) - \bar{X}_t(h))^* z\|^2 \geq 0,$$

viser ovenstående, at $\text{MSE}[\bar{X}_t(h)] - \text{MSE}[\hat{X}_t(h)]$ er en ikke-negativ definit matrix, hvilket skulle vises. \square

Bemærkning 4.9. Af beviset for [Sætning 4.7](#) fremgik det, at prediktionsfejlen eller “forecast error” er

$$X_{t+h} - \hat{X}_t(h) = X_{t+h} - E(X_{t+h} | X_{\leq t}) = \sum_{i=0}^{h-1} \Psi_i w_{t+h-i}, \quad (4.63)$$

som har middelværdi

$$E(\mathbf{X}_{t+h} - \hat{\mathbf{X}}_t(h)) = E(\mathbf{X}_{t+h} - E(\mathbf{X}_{t+h} | \mathbf{X}_{\leq t})) = \sum_{i=0}^{h-1} \Psi_i E(\mathbf{w}_{t+h-i}) = 0$$

og kovariansmatricks

$$\begin{aligned} \Gamma_X(h | \mathbf{X}_{\leq t}) &= \text{MSE}[\hat{\mathbf{X}}_t(h)] = E[(\mathbf{X}_{t+h} - \hat{\mathbf{X}}_t(h))(\mathbf{X}_{t+h} - \hat{\mathbf{X}}_t(h))^*] \\ &= E\left[\left(\sum_{i=0}^{h-1} \Psi_i \mathbf{w}_{t+h-i}\right)\left(\sum_{j=0}^{h-1} \Psi_j \mathbf{w}_{t+h-j}\right)^*\right] \\ &= \sum_{i=0}^{h-1} \Psi_i \Sigma_w \Psi_i^*. \end{aligned} \quad (4.64)$$

Bemærk, at kovariansmatricen for forecast error ikke afhænger af det tidspunkt t , hvor forecastet foretages.

Lader vi $h \rightarrow \infty$ i (4.64), fås det, at

$$\Sigma_X(h | \mathbf{X}_{\leq t}) \rightarrow \sum_{i=0}^{\infty} \Psi_i \Sigma_w \Psi_i^* = \Gamma_X(0) = \Sigma_X, \quad (4.65)$$

hvor Σ_X er kovariansmatricen for \mathbf{X}_t .

Havde vi i stedet for $\hat{\mathbf{X}}_t$ brugt processens middelværdi $\boldsymbol{\mu}$ som prediktor, havde MSE-matricen for prediktoren været

$$E[(\mathbf{X}_{t+h} - \boldsymbol{\mu})(\mathbf{X}_{t+h} - \boldsymbol{\mu})^*],$$

som jo bare er kovariansmatricen Σ_X for \mathbf{X}_t . Når vi til tidspunkt t vil lave et forecast af processen uendelig langt ud i fremtiden, bliver det optimale forecast altså bare processens middelværdi $\boldsymbol{\mu}$. Da vil informationen $\mathbf{X}_{\leq t}$, som vi har til tid t , ikke spille nogen rolle for værdier så langt ude i fremtiden, og det bedste bud, vi kan give for de fremtidige værdier, er derfor middelværdien. ■

Bemærkning 4.10. Antagelsen (ii) i Sætning 4.7, om at elementerne af den hvide støj skal være uafhængige, er noget restriktiv. Uden denne antagelse er elementerne kun ukorreleerede, så derfor er det nødvendigt at lave nogle ekstra antagelser for at bestemme den optimale prediktor, der minimerer MSE.

Vi begrænser os nu til at finde den prediktor, der giver den mindste værdi af MSE blandt alle prediktorer $\bar{\mathbf{X}}_t(h)$, der er *lineære* funktioner af de variable i $\mathbf{X}_{\leq t}$. Vi betragter altså prediktorer af formen

$$\bar{\mathbf{X}}_t(h) = \sum_{i=0}^{\infty} \Delta_i \mathbf{X}_{t-i},$$

eller på grund af antagelsen (i) i Sætning 4.7

$$\bar{X}_t(h) = \sum_{i=0}^{\infty} \Lambda_i w_{t-i}.$$

Forecastfejlen bliver da

$$\begin{aligned} X_{t+h} - \bar{X}_t(h) &= \sum_{i=0}^{\infty} \Psi_i w_{t+h-i} - \sum_{i=0}^{\infty} \Lambda_i w_{t-i} \\ &= \sum_{i=0}^{h-1} \Psi_i w_{t+h-i} - \sum_{i=0}^{\infty} (\Psi_{i+h} - \Lambda_i) w_{t-i}. \end{aligned}$$

Da $E(w_s w_t^*) = 0$ for $s \neq t$, finder vi, at

$$E((X_{t+h} - \bar{X}_t(h))(X_{t+h} - \bar{X}_t(h))^*) = \sum_{i=0}^{h-1} \Psi_i \Sigma_w \Psi_i^* + \sum_{i=0}^{\infty} (\Psi_{i+h} - \Lambda_i) \Sigma_w (\Psi_{i+h} - \Lambda_i)^*.$$

Da det for $z \in \mathbb{R}^K$ gælder, at

$$z^* (\Psi_{i+h} - \Lambda_i) \Sigma_w (\Psi_{i+h} - \Lambda_i)^* z \geq 0,$$

ses det, at $\text{MSE}[\bar{X}_t(h)]$ bliver mindst, hvis $\Lambda_i = \Psi_{i+h}$, dvs.

$$\bar{X}_t(h) = \sum_{i=0}^{\infty} \Lambda_i w_{t-i} = \sum_{i=0}^{\infty} \Psi_{i+h} w_{t-i} = \hat{X}_t(h). \quad \blacksquare$$

Forecasting i VARMA-processer

De tre repræsentationer af en stabil og invertibel VARMA(p, q)-proces, giver alternative formler for minimum mean square prediktoren $\hat{X}_t(h)$ af X_{t+h} . I diskussionen heraf antager vi uden tab af generalitet, at $\mu = 0$, og sætter $\hat{X}_t(h) = X_{t+h}$ for $h \leq 0$.

Den blandede repræsentation af processen er

$$\Phi(B)X_{t+h} = \Theta(B)w_{t+h}, \quad (4.66)$$

mens VAR(∞)-repræsentationen og VMA(∞)-repræsentationen er henholdsvis

$$w_{t+h} = \Theta(B)^{-1} \Phi(B)X_{t+h} = \Pi(B)X_{t+h} = \sum_{i=0}^{\infty} \Pi_i X_{t+h-i} \quad (4.67)$$

og

$$X_{t+h} = \Phi(B)^{-1} \Theta(B)w_{t+h} = \Psi(B)w_{t+h} = \sum_{i=0}^{\infty} \Psi_i w_{t+h-i}. \quad (4.68)$$

Tager vi betinget middelværdi givet $\mathbf{X}_{\leq t}$ i (4.66) finder vi

$$\Phi(B)\hat{\mathbf{X}}_t(h) = \Theta(B)E(\mathbf{w}_{t+h}|\mathbf{X}_{\leq t}),$$

hvor B -operatoren nu virker på h . Er betingelsen (ii) i Sætning 4.7 opfyldt, fås af formel (4.59), at

$$\Phi(B)\hat{\mathbf{X}}_t(h) = 0 \quad \text{for } h > q. \quad (4.69)$$

Hvis processen er endimensional, dvs. hvis $K = 1$, opfylder forecastene altså en differensligning.

Tager vi betinget middelværdi givet $\mathbf{X}_{\leq t}$ i (4.67) fås for $h > 0$, at

$$0 = \sum_{i=0}^{\infty} \Pi_i \hat{\mathbf{X}}_t(h-i) = \sum_{i=0}^{h-1} \Pi_i \hat{\mathbf{X}}_t(h-i) + \sum_{i=h}^{\infty} \Pi_i \mathbf{X}_{t+h-i}. \quad (4.70)$$

Endelig giver VMA(∞)-repræsentationen i (4.68) som tidligere vist h -trins forecastet

$$\hat{\mathbf{X}}_t(h) = \sum_{i=h}^{\infty} \Psi_i \mathbf{w}_{t+h-i}. \quad (4.71)$$

4.8 Strukturel analyse af VARMA-processer

I den strukturelle analyse undersøges forholdet mellem komponenterne X_{1t}, \dots, X_{Kt} i den flerdimensionale VARMA-proces $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$. Hvis det viser sig, at en gruppe af komponenterne påvirker en anden gruppe, kan denne information bruges til at forbedre forecast. Man kan "måle" afhængigheden mellem grupper af variable på flere måder. Vi skal se på Granger-kausalitet, øjeblikkelig kausalitet og impuls-responser. Granger-kausaliteten defineres først for en helt generel proces, men diskuteres senere for en VARMA-proces. Impuls-responserne defineres for to forskellige repræsentationer af VARMA-processen, nemlig den simple og den ortogonale repræsentation.

Granger-kausalitet

Definition 4.11. Lad $\mathbf{X}_t = (\mathbf{Y}_t^*, \mathbf{Z}_t^*)^*$ være en opdeling af den K -dimensionale tidsrække $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ i en M -dimensional tidsrække $\{\mathbf{Y}_t\}_{t \in \mathbb{Z}}$ og en $(K-M)$ -dimensional $\{\mathbf{Z}_t\}_{t \in \mathbb{Z}}$. Vi siger da, at $\{\mathbf{Z}_t\}_{t \in \mathbb{Z}}$ **ikke er Granger kausal** for $\{\mathbf{Y}_t\}_{t \in \mathbb{Z}}$, hvis det for alle t og for alle h gælder, at

$$\Sigma_Y(h|\mathbf{Y}_{\leq t} \cup \mathbf{Z}_{\leq t}) = \Sigma_Y(h|\mathbf{Y}_{\leq t}),$$

dvs. hvis MSE-matricen $\Sigma_Y(h|\mathbf{Y}_{\leq t} \cup \mathbf{Z}_{\leq t})$ for forecastet af \mathbf{Y}_{t+h} baseret på $\mathbf{Y}_{\leq t} \cup \mathbf{Z}_{\leq t}$ er den samme som MSE-matricen $\Sigma_Y(h|\mathbf{Y}_{\leq t})$ for forecastet af \mathbf{Y}_{t+h} baseret udelukkende på $\mathbf{Y}_{\leq t}$.

Ækvivalent hermed, siges $\{Z_t\}_{t \in \mathbb{Z}}$ **at være Granger kausal** for $\{Y_t\}_{t \in \mathbb{Z}}$, hvis der eksisterer et tidspunkt t og et $h > 0$, så

$$\Sigma_Y(h|Y_{\leq t} \cup Z_{\leq t}) \neq \Sigma_Y(h|Y_{\leq t})$$

og

$$\Sigma_Y(h|Y_{\leq t} \cup Z_{\leq t}) \leq \Sigma_Y(h|Y_{\leq t}).$$

Bemærkning 4.12. Hvis $\{Z_t\}_{t \in \mathbb{Z}}$ er Granger kausal for $\{Y_t\}_{t \in \mathbb{Z}}$, påvirker processen $\{Z_t\}_{t \in \mathbb{Z}}$ processen $\{Y_t\}_{t \in \mathbb{Z}}$ i den forstand, at der eksisterer et tidspunkt t og et $h > 0$, så MSE-matricen for h -trins forecastet af Y_{t+h} baseret på $Y_{\leq t} \cup Z_{\leq t}$ er “mindre” end MSE-matricen for h -trins forecastet af Y_{t+h} baseret udelukkende på kendskab til $Y_{\leq t}$; kendskabet til både $Z_{\leq t}$ og $Y_{\leq t}$ giver altså et bedre forecast af Y_{t+h} end forecastet baseret udelukkende på $Y_{\leq t}$. Omvendt, hvis $\{Z_t\}_{t \in \mathbb{Z}}$ ikke er Granger kausal for $\{Y_t\}_{t \in \mathbb{Z}}$, forbedres h -trins forecastet af Y_{t+h} baseret på $Y_{\leq t}$ ikke, når vi også tager $Z_{\leq t}$ i betragtning ved beregning af forecastet. ■

Vi diskuterer Granger-kausalitet for svagt stationære K -dimensionale processer $\{X_t\}_{t \in \mathbb{Z}}$, der har en VMA(∞)-repræsentation,

$$X_t = \mu + \sum_{i=0}^{\infty} \Psi_i w_{t-i} = \mu + \Psi(B)w_t, \quad (4.72)$$

hvor $\Psi(B) = \sum_{i=0}^{\infty} \Psi_i B^i$ og $\Psi_0 = I_K$.

Vi deler processen op i to underprocesser $\{Y_t\}_{t \in \mathbb{Z}}$ og $\{Z_t\}_{t \in \mathbb{Z}}$, som er henholdsvis M - og $(K - M)$ -dimensionale, dvs.

$$X_t = \begin{bmatrix} Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \Psi_{11}(B) & \Psi_{12}(B) \\ \Psi_{21}(B) & \Psi_{22}(B) \end{bmatrix} \begin{bmatrix} w_{1t} \\ w_{2t} \end{bmatrix}, \quad (4.73)$$

hvor

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad \text{og} \quad w_t = \begin{bmatrix} w_{1t} \\ w_{2t} \end{bmatrix}.$$

Hvis for $i = 0, 1, \dots$

$$\Psi_i = \begin{bmatrix} \Psi_{11,i} & \Psi_{12,i} \\ \Psi_{21,i} & \Psi_{22,i} \end{bmatrix}$$

er en opdeling af $K \times K$ matricen Ψ_i , så $\Psi_{11,i}$ er en $M \times M$ matriks, er $\Psi_{rs}(B)$ i (4.73) givet ved

$$\Psi_{rs}(B) = \sum_{i=0}^{\infty} \Psi_{rs,i} B^i, \quad r, s = 1, 2.$$

Udskrives (4.73), fås

$$\mathbf{Y}_t = \boldsymbol{\mu}_1 + \Psi_{11}(B)\mathbf{w}_{1t} + \Psi_{12}(B)\mathbf{w}_{2t} = \boldsymbol{\mu}_1 + \sum_{i=0}^{\infty} \Psi_{11,i}\mathbf{w}_{1,t-i} + \sum_{i=0}^{\infty} \Psi_{12,i}\mathbf{w}_{2,t-i}, \quad (4.74)$$

$$\mathbf{Z}_t = \boldsymbol{\mu}_2 + \Psi_{21}(B)\mathbf{w}_{1t} + \Psi_{22}(B)\mathbf{w}_{2t} = \boldsymbol{\mu}_2 + \sum_{i=0}^{\infty} \Psi_{21,i}\mathbf{w}_{1,t-i} + \sum_{i=0}^{\infty} \Psi_{22,i}\mathbf{w}_{2,t-i}. \quad (4.75)$$

Følgende sætning fortæller, hvordan vi ud fra koefficientmatricerne i $\Psi(B)$ kan afgøre, om \mathbf{Z}_t ikke er Granger-kausal for \mathbf{Y}_t .

Sætning 4.13. For processen i (4.72), der opfylder !betingelserne (i) og (ii) i Sætning 4.7 på side 107, er følgende udsagn ækvivalente:

- (i) $\Sigma_Y(1|\mathbf{Y}_{\leq t} \cup \mathbf{Z}_{\leq t}) = \Sigma_Y(1|\mathbf{Y}_{\leq t})$ for $t \in \mathbb{Z}$
- (ii) $\Psi_{12,i} = 0$ for $i = 0, 1, \dots$
- (iii) $\Sigma_Y(h|\mathbf{Y}_{\leq t} \cup \mathbf{Z}_{\leq t}) = \Sigma_Y(h|\mathbf{Y}_{\leq t})$, for $t \in \mathbb{Z}$ og $h = 1, 2, \dots$

Bevis. Vi betragter processen med VMA(∞)-repræsentationen (4.73). Fra formel (4.71) ses det, at det optimale 1-trins forecast for \mathbf{X}_t er

$$\hat{\mathbf{X}}_t(1) = \boldsymbol{\mu} + \sum_{i=1}^{\infty} \Psi_i \mathbf{w}_{t+1-i}.$$

Ved at gange dette med $(M \times K)$ -matricen $[I_M : 0]$, fås det optimale 1-trins forecast for \mathbf{Y}_t baseret på $\mathbf{Y}_{\leq t} \cup \mathbf{Z}_{\leq t}$ til at være

$$\begin{aligned} \hat{\mathbf{Y}}_t(1|\mathbf{Y}_{\leq t} \cup \mathbf{Z}_{\leq t}) &= \hat{\mathbf{Y}}_t(1|\mathbf{X}_{\leq t}) = [I_M : 0]\hat{\mathbf{X}}_t(1) \\ &= \boldsymbol{\mu}_1 + \sum_{i=1}^{\infty} \Psi_{11,i}\mathbf{w}_{1,t+1-i} + \sum_{i=1}^{\infty} \Psi_{12,i}\mathbf{w}_{2,t+1-i}. \end{aligned} \quad (4.76)$$

Fejlen på dette forecast bliver

$$\begin{aligned} \mathbf{Y}_{t+1} - \hat{\mathbf{Y}}_t(1|\mathbf{Y}_{\leq t} \cup \mathbf{Z}_{\leq t}) &= \sum_{i=0}^{\infty} \Psi_{11,i}\mathbf{w}_{1,t+1-i} + \sum_{i=0}^{\infty} \Psi_{12,i}\mathbf{w}_{2,t+1-i} - \sum_{i=1}^{\infty} \Psi_{11,i}\mathbf{w}_{1,t+1-i} - \sum_{i=1}^{\infty} \Psi_{12,i}\mathbf{w}_{2,t+1-i} \\ &= \mathbf{w}_{1,t+1}, \end{aligned}$$

hvor det er brugt, at $\Psi_{12,0} = 0$, og $\Psi_{11,0} = I_M$, da $\Psi_0 = I_K$.

Vi har antaget, at $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ er stationær. Wold's Decomposition Theorem omtalt i Bemærkning 4.3 på side 99 giver så, at en delproces af denne har en VMA-repræsentation. Så udover, at

$$\mathbf{Y}_t = \boldsymbol{\mu}_1 + \sum_{i=0}^{\infty} \Psi_{11,i}\mathbf{w}_{1,t-i} + \sum_{i=1}^{\infty} \Psi_{12,i}\mathbf{w}_{2,t-i}, \quad (4.77)$$

hvor \mathbf{Y}_t udtrykkes ved den hvide støj hørende til $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$, kan \mathbf{Y}_t også repræsenteres som

$$\mathbf{Y}_t = \boldsymbol{\mu}_1 + \sum_{i=0}^{\infty} F_i \mathbf{v}_{t-i},$$

hvor $F_0 = I_M$, og $\{\mathbf{v}_t\}_{t \in \mathbb{Z}}$ er en hvid støj.

Vi kan ud fra denne repræsentation bestemme det optimale 1-trins forecast for \mathbf{Y}_t , hvor det kun er informationen om $\mathbf{Y}_{\leq t}$, der tages med i betragtning. Denne er ifølge formel (4.71) givet ved

$$\hat{\mathbf{Y}}_t(1|\mathbf{Y}_{\leq t}) = \boldsymbol{\mu}_1 + \sum_{i=1}^{\infty} F_i \mathbf{v}_{t+1-i},$$

og fejlen ved dette forecast bliver

$$\mathbf{Y}_{t+1} - \hat{\mathbf{Y}}_t(1|\mathbf{Y}_{\leq t}) = \sum_{i=0}^{\infty} F_i \mathbf{v}_{t+1-i} - \sum_{i=1}^{\infty} F_i \mathbf{v}_{t+1-i} = \mathbf{v}_{t+1}.$$

Det ses,

$$\begin{aligned} & \hat{\mathbf{Y}}_t(1|\mathbf{Y}_{\leq t} \cup \mathbf{Z}_{\leq t}) = \hat{\mathbf{Y}}_t(1|\mathbf{Y}_{\leq t}) \\ \Updownarrow & \\ & \mathbf{Y}_{t+1} - \hat{\mathbf{Y}}_t(1|\mathbf{Y}_{\leq t} \cup \mathbf{Z}_{\leq t}) = \mathbf{Y}_{t+1} - \hat{\mathbf{Y}}_t(1|\mathbf{Y}_{\leq t}) \\ \Updownarrow & \\ & \mathbf{w}_{1,t} = \mathbf{v}_t \end{aligned}$$

for alle t . Så hvis de to prediktorer er identiske, fås det, at

$$\begin{aligned} \mathbf{Y}_t &= \boldsymbol{\mu}_1 + \sum_{i=0}^{\infty} F_i \mathbf{v}_{t-i} \\ &= \boldsymbol{\mu}_1 + \sum_{i=0}^{\infty} F_i \mathbf{w}_{1,t-i} \\ &= \boldsymbol{\mu}_1 + \sum_{i=0}^{\infty} [F_i : 0] \mathbf{w}_{t-i} \\ &= \boldsymbol{\mu}_1 + \sum_{i=0}^{\infty} [\Psi_{11,i} : \Psi_{12,i}] \mathbf{w}_{t-i} \\ &= \boldsymbol{\mu}_1 + \sum_{i=0}^{\infty} \Psi_{11,i} \mathbf{w}_{1,t-i} + \sum_{i=0}^{\infty} \Psi_{12,i} \mathbf{w}_{2,t-i}, \end{aligned}$$

idet \mathbf{Y}_t også har repræsentationen (4.77). Da VMA-repræsentationen er entydig, må der nødvendigvis gælde, at $F_i = \Psi_{11,i}$ og $\Psi_{12,i} = 0$ for $i = 1, 2, \dots$. Altså har vi vist, at (i) \Rightarrow (ii).

Antag nu, at $\Psi_{12,i} = 0$ for $i = 1, 2, \dots$. Da er

$$\hat{Y}_t(h|Y_{\leq t} \cup Z_{\leq t}) = \mu_1 + \sum_{i=h}^{\infty} \Psi_{11,i} w_{1,t+h-i}. \quad (4.78)$$

Ifølge (4.74) er

$$Y_t = \mu_1 + \sum_{i=0}^{\infty} \Psi_{11,i} w_{1,t-i}.$$

og (4.71) og (4.78) giver, at

$$\hat{Y}_t(h|Y_{\leq t}) = \mu_1 + \sum_{i=h}^{\infty} \Psi_{11,i} w_{1,t+h-i} = \hat{Y}_t(h|Y_{\leq t} \cup Z_{\leq t}),$$

så (ii) \Rightarrow (iii). Da (iii) trivielt medfører (i), er sætningen vist. \square

For VAR(p)-processen (4.8) på side 89 kan spørgsmålet om Granger-kausaltet afgøres ved at betragte koefficientmatricerne

$$\Phi_i = \begin{bmatrix} \Phi_{11,i} & \Phi_{12,i} \\ \Phi_{21,i} & \Phi_{22,i} \end{bmatrix}, \quad i = 1, \dots, p,$$

idet der, som vist i Opgave 16, gælder

Sætning 4.14. *For en VAR(p)-proces er betingelsen, at delmatricerne $\Psi_{12,i} = 0$ for $i = 0, 1, \dots$ i processens VMA-repræsentation, ækvivalent med, at delmatricerne $\Phi_{12,i} = 0$ for $i = 0, 1, \dots, p$ i processens VAR(p)-repræsentation.*

I praksis skal man være forsigtig med at drage for mange konklusioner af Granger-kausalteten. Fordi man har vist, at to grupper af variable ikke er Granger-kausale, betyder det ikke nødvendigvis, at de to grupper ikke påvirker hinanden på anden vis. Desuden har vi i disse beregninger antaget, at informationen til tid t består af processen op til og med tid t . I virkeligheden kunne man have endnu mere information. Det kunne f.eks. dreje sig om information fra en ydre variabel, hvorom vi kender dennes indvirkning på processen.

Øjeblikkelig kausalitet

En anden form for kausalitet mellem to grupper af variable er *øjeblikkelig kausalitet*. Betragt de to processer $\{Z_t\}_{t \in \mathbb{Z}}$ og $\{Y_t\}_{t \in \mathbb{Z}}$ fra før. Der siges at være øjeblikkelig kausalitet mellem de to processer, hvis

$$\Sigma_Z(1|Y_{\leq t} \cup Z_{\leq t}) \neq \Sigma_Z(1|Y_{\leq t+1} \cup Z_{\leq t}).$$

Dette betyder, at 1-trins forecastet til tid t af Z_{t+1} på baggrund af informationen $X_{\leq t} = Y_{\leq t} \cup Z_{\leq t}$ bliver forbedret, hvis observationen Y_{t+1} også tages med i betragtning. Vi siger, at $\{Y_t\}_{t \in \mathbb{Z}}$ er *øjeblikkelig kausal* for $\{Z_t\}_{t \in \mathbb{Z}}$.

Til forskel fra Granger-kausalitet, har øjeblikkelig kausalitet ingen "retning", men er symmetrisk, hvis den hvide støj $\{w_t\}_{t \in \mathbb{Z}}$ er Gaussisk. Hvis dette er tilfældet, gælder der, at hvis $\{Y_t\}_{t \in \mathbb{Z}}$ er øjeblikkelig kausal for $\{Z_t\}_{t \in \mathbb{Z}}$, så er det omvendte også tilfældet, se [Sætning 4.14](#).

I diskussionen af begrebet øjeblikkelig kausalitet betragter vi processer $\{X_t\}_{t \in \mathbb{Z}}$ med VMA(∞)-repræsentationen

$$X_t = \mu + \sum_{i=0}^{\infty} \Psi_i w_{t-i},$$

der opfylder betingelserne (i) og (ii) i [Sætning 4.7](#) på [side 107](#).

Da $\Psi_0 = I_K$ er

$$X_t = \begin{bmatrix} Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix} \Psi_{11,i} & \Psi_{12,i} \\ \Psi_{21,i} & \Psi_{22,i} \end{bmatrix} \begin{bmatrix} w_{1,t-i} \\ w_{2,t-i} \end{bmatrix}. \quad (4.79)$$

Formel (4.71) medfører, at

$$\hat{X}_t(1) = \begin{bmatrix} \hat{Y}_t(1) \\ \hat{Z}_t(1) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix} \Psi_{11,i} & \Psi_{12,i} \\ \Psi_{21,i} & \Psi_{22,i} \end{bmatrix} \begin{bmatrix} w_{1,t+1-i} \\ w_{2,t+1-i} \end{bmatrix}. \quad (4.80)$$

Det ses, at

$$Y_{t+1} = w_{1,t+1} + \hat{Y}_t(1)$$

og

$$Z_{t+1} = w_{2,t+1} + \hat{Z}_t(1) = w_{2,t+1} + Z_t(1|X_{\leq t}), \quad (4.81)$$

som viser, at en 1-1-sammenhæng mellem $X_{\leq t} \cup Y_{t+1}$ og $w_{\leq t} \cup w_{1,t+1}$, så

$$Z_t(1|X_{\leq t} \cup Y_{t+1}) = Z_t(1|w_{\leq t} \cup w_{1,t+1}).$$

Det fås nu ved at tage den betingede middelværdi i (4.81) givet $w_{\leq t} \cup w_{1,t+1}$, at

$$Z_t(1|X_{\leq t} \cup Y_{t+1}) = E(w_{2,t+1}|w_{\leq t} \cup w_{1,t+1}) + Z_t(1|X_{\leq t}).$$

Da w_s og w_t er uafhængige hvis $s \neq t$, er

$$E(w_{2,t+1}|w_{\leq t} \cup w_{1,t+1}) = E(w_{2,t+1}|w_{1,t+1})$$

Dermed har vi, at

$$Z_t(1|X_{\leq t} \cup Y_{t+1}) = Z_t(1|X_{\leq t}) \quad \Leftrightarrow \quad E(w_{2,t+1}|w_{1,t+1}) = 0.$$

Hvis den hvide støj er Gaussisk, dvs.

$$\mathbf{w}_t = \begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix} \sim N_K(\mathbf{0}, \Sigma_w) = N_K\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right),$$

gælder der, at

$$E(w_{2,t} | w_{1,t}) = \Sigma_{21} \Sigma_{11}^{-1} w_{1,t} = 0 \Leftrightarrow \Sigma_{21} = 0.$$

Vi har nu vist

Sætning 4.15. *Betragt en proces $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$, der opfylder betingelserne i [Sætning 4.7](#) på side 107.*

Lad $\mathbf{X}_t = (\mathbf{Y}_t^, \mathbf{Z}_t^*)^*$ være en opdeling af \mathbf{X}_t , så \mathbf{Y}_t og \mathbf{Z}_t er henholdsvis M og $K - M$ -dimensionale, og lad $\mathbf{w}_t = (w_{1,t}^*, w_{2,t}^*)^*$ være en tilsvarende opdeling af elementerne i den hvide støj.*

Da er $\{\mathbf{Y}_t\}_{t \in \mathbb{Z}}$ ikke øjeblikkelig kausal for $\{\mathbf{Z}_t\}_{t \in \mathbb{Z}}$ hvis og kun hvis $E(w_{2,t} | w_{1,t}) = 0$.

Hvis den hvide støj er Gaussisk, er $\{\mathbf{Y}_t\}_{t \in \mathbb{Z}}$ ikke øjeblikkelig kausal for $\{\mathbf{Z}_t\}_{t \in \mathbb{Z}}$ hvis og kun hvis komponenterne $w_{1,t}$ og $w_{2,t}$ er uafhængige. (I så tilfælde er $\{\mathbf{Z}_t\}_{t \in \mathbb{Z}}$ ikke øjeblikkelig kausal for $\{\mathbf{Y}_t\}_{t \in \mathbb{Z}}$.)

I arbejdet med Granger-kausalitet og øjeblikkelig kausalitet, har vi betragtet en proces $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ med en VMA(∞)-repræsentation, som vi har delt op i to delprocesser $\{\mathbf{Z}_t\}_{t \in \mathbb{Z}}$ og $\{\mathbf{Y}_t\}_{t \in \mathbb{Z}}$. Vi har så undersøgt det indbyrdes forhold mellem de to delprocesser. Den information, som har været tilgængelig til det pågældende tidspunkt, har kun bestået af information om processen $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$. Dvs. den har kun indeholdt viden om de to delprocesser, hvis indbyrdes forhold, vi undersøgte. Vi viste i [Sætning 4.13](#), at hvis 1-trins forecastet for \mathbf{Y}_t ikke blev forbedret ved at tage $\{\mathbf{Z}_s\}_{s \leq t}$ med i betragtning, så var det samme gældende for h -trins forecastet, $h = 2, 3, \dots$

Man kunne også forestille sig, at processen $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ kunne deles op i tre eller flere delprocesser, og at man kun ønskede at finde det indbyrdes forhold mellem bare to af processerne. Hvis igen informationen til tid t består af viden om $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ op til og med tid t , vil vi dermed være i besiddelse af information udover de processer, hvis forhold vi skal bestemme. Dette giver nogle problemer. Da vil vi nemlig ikke længere kunne sige, som vi gjorde i [Sætning 4.13](#), at hvis ikke 1-trins forecastet for \mathbf{Y}_t bliver bedre af medtage $\mathbf{Z}_{s \leq t}$ i informationen, så er det samme gældende for h -trins forecastet. I disse tilfælde kan man betragte en anden form for flertrins kausalitet, som vi ikke kommer nærmere ind på her. Se eventuelt [Lütkepohl \(2007\)](#) side 49 og fremefter.

Impuls-response analyse

Vi skal nu se på en anden måde at undersøge forholdet mellem komponenterne i en VARMA-proces på, nemlig ved at se på impuls-responserne. Her er det i modsætning til ved kausabiliteten muligt at betragte forholdet mellem kun nogle af komponenterne i processen. Vi skal undersøge, hvordan en impuls i en af komponenterne kan påvirke de andre komponenter i processen. Det viser sig, at disse impuls-responser kan aflæses i koefficientmatricerne Ψ_i fra VMA-repræsentationen for processen.

Vi betragter som tidligere en proces $\{X_t\}_{t \in \mathbb{Z}}$ med VMA-repræsentationen

$$X_t = \sum_{i=0}^{\infty} \Psi_i w_{t-i}.$$

Antag, at $w_t = 0$, hvis $t \neq 0$, så det eneste element i den hvide støj, der ikke er 0 er altså w_0 . Det ses, at

$$X_i = \Psi_i w_0.$$

Da $X_i = (X_{1i}, \dots, X_{ji}, \dots, X_{Ki})^*$, er X_{ji} den j 'te komponent i $\Psi_i w_0$. Hvis

$$\Psi_i = \{\psi_{jk,i}\}$$

og $w_0 = e_k$, hvor e_k er den K -dimensionale enhedsvektor med et 1-tal på den k 'te plads og nuller ellers, fås

$$X_{ji} = \psi_{jk,i},$$

som altså er responsen i den j 'te komponent til tid i på en impuls i den k 'te komponent til tid 0.

Da impuls-responserne er elementerne i Ψ_i , hænger disse også sammen med Granger-kausaltet. Hvis en impuls i den j 'te komponent ikke giver nogen respons i de øvrige komponenter, er den j 'te variabel ikke Granger-kausal for gruppen bestående af de resterende variable. Dette ses ved at betragte det glidende gennemsnit for processen $\{X_t\}_{t \in \mathbb{Z}}$ med middelværdi nul

$$\begin{bmatrix} Y_t \\ Z_t \end{bmatrix} = \sum_{i=0}^{\infty} \begin{bmatrix} \Psi_{11,i} & \Psi_{12,i} \\ \Psi_{21,i} & \Psi_{22,i} \end{bmatrix} \begin{bmatrix} w_{1t} \\ w_{2t} \end{bmatrix}.$$

Lad her $Z_t = X_{Kt}$ være den K 'te variabel og $Y_t = (X_{1t}, \dots, X_{K-1,t})^*$ gruppen af de øvrige variable. Ifølge [Sætning 4.13](#) er Z_t ikke Granger-kausal for Y_t , hvis og kun hvis $\Psi_{12,i} = 0$. I dette tilfælde er $\Psi_{12,i}$ vektoren

$$\Psi_{12,i} = \begin{bmatrix} \psi_{1K,i} \\ \psi_{2K,i} \\ \vdots \\ \psi_{(K-1)K,i} \end{bmatrix} = 0,$$

hvoraf det ses, at en impuls i den K 'te variabel, dvs. i \mathbf{Z}_t , ingen respons giver i de øvrige.

Når vi i praksis skal undersøge forholdet mellem komponenterne i \mathbf{X}_t , er vi interesserede i at kunne afgøre, hvorvidt $\psi_{jk,i} = 0$ for $j \neq k$ og $i = 1, 2, \dots$. Umidledbart ser det ud til, at man er nødt til at beregne de uendelig mange koefficientmatricer Ψ_i for at finde svaret, men dette er ikke altid nødvendigt. Vi har følgende resultat for en VAR(p)-proces:

Sætning 4.16. Hvis $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ er en stabil K -dimensional VAR(p)-proces, gælder der for $j \neq k$, at

$$\psi_{jk,i} = 0 \quad \text{for } i = 1, 2, \dots$$

er ækvivalent med

$$\psi_{jk,i} = 0 \quad \text{for } i = 1, \dots, p(K-1).$$

Dermed er det kun nødvendigt at bestemme de første $p(K-1)$ impuls-responser. Hvis disse er lig nul, vil de resterende også være det.

Bevis. [Lütkepohl \(2007, side 55\)](#). □

Ved beregningen af de simple impuls-responser blev det antaget, at der kun skete en impuls i én af de variable, og at der ikke skete noget i de øvrige. Dermed vil den respons, som vi så sporede senere, kunne tolkes som en respons grundet en impuls i en enkelt variabel. Dette er i virkeligheden en streng antagelse, da det godt kunne tænkes, at en impuls i en variabel gav en øjeblikkelig impuls i en af de andre variable. Responserne vil dermed ikke kunne tolkes som værende et resultat udelukkende af en variabels impuls, men af flere. Dette vil være tilfældet, hvis komponenterne i \mathbf{w}_t ikke er uafhængige eller ukorreleerede. Hvis der er en stor korrelation mellem nogle af komponenterne i \mathbf{w}_t , vil de impuls-responser, som vi netop har set på, være misvisende for de indbyrdes forhold mellem komponenterne i processen.

Vi erstatter derfor den hvide støj $\{\mathbf{w}_t\}_{t \in \mathbb{Z}}$ med en hvid støj $\{\mathbf{u}_t\}_{t \in \mathbb{Z}}$, hvis komponenter er ukorreleerede. Dette gøres således:

Kovariansmatricen Σ_w for den hvide støj $\{\mathbf{w}_t\}_{t \in \mathbb{Z}}$ er positiv definit, og kan derfor Cholesky-dekomponeres. Dvs. der eksisterer en ikke-singulær nedre trekantsmatris P , hvis diagonalelementer er positive, så

$$\Sigma_w = PP^*.$$

Dermed kan VMA-repræsentationen skrives som

$$\mathbf{X}_t = \sum_{i=0}^{\infty} \Psi_i PP^{-1} \mathbf{w}_{t-i} = \sum_{i=0}^{\infty} \Lambda_i \mathbf{u}_{t-i},$$

hvor $\Lambda_i = \Psi_i P$, og $\mathbf{u}_t = P^{-1} \mathbf{w}_t$. Processen $\{\mathbf{u}_t\}_{t \in \mathbb{Z}}$ er en hvid støj, idet $E(\mathbf{u}_t) = 0$ for alle t , og kovariansmatricen Σ_w er

$$\Sigma_u = P^{-1} \Sigma_w (P^{-1})^* = P^{-1} P P^* (P^{-1})^* = I_K.$$

Heraf ses det, at komponenterne i \mathbf{u}_t er ukorrelerede, og vi kalder den hvide støj $\{\mathbf{u}_t\}_{t \in \mathbb{Z}}$ *ortogonal*. Når dette er tilfældet, kaldes elementerne i den hvide støj for *innovationer*.

I stedet for at se på de simple impuls-responser, som kan være misvisende, bestemmes i stedet for de *ortogonale impuls-responser* ud fra VMA-repræsentationen

$$\mathbf{X}_t = \sum_{i=0}^{\infty} \Lambda_i \mathbf{u}_{t-i}, \quad (4.82)$$

hvor komponenterne i den hvide støj $\mathbf{u}_t = P^{-1} \mathbf{w}_t$ er ukorrelerede og har varians 1, $\Sigma_u = I_K$. Da komponenterne i \mathbf{u}_t er ukorrelerede, kan vi antage, at en impuls i f.eks. den k 'te komponent ikke påvirker de øvrige komponenter. Det var netop dette, som vi gerne ville sikre os.

Ligesom elementerne i Ψ_i er de simple impuls-responser, er elementerne i koefficientmatricerne Λ_i de ortogonale impuls-responser. Dermed repræsenterer det jk -te element i Λ_i , $\lambda_{jk,i}$, responsen i den j 'te komponent i tidsenheder efter en impuls i den k 'te komponent af \mathbf{u}_t . Da elementerne i \mathbf{u}_t har varians 1 vil en impuls af enhedsstørrelse svare til en impuls af størrelse en standardafvigelse.

Ligesom for de simple impuls-responser er det ikke altid nødvendigt at bestemme alle matricerne Λ_i , $i = 1, 2, \dots$, for at afgøre, hvorvidt en impuls i en komponent slet ikke påvirker nogle af de øvrige. Vi har nemlig følgende resultat, som er helt tilsvarende det for de simple impuls-responser.

Sætning 4.17. Hvis $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ er en stabil K -dimensional VAR(p)-proces, gælder der for $j \neq k$, at

$$\lambda_{jk,i} = 0 \quad \text{for } i = 1, 2, \dots$$

er ækvivalent med

$$\lambda_{jk,i} = 0 \quad \text{for } i = 1, \dots, p(K-1).$$

Bevis. Lütkepohl (2007, side 61). □

Ligesom med Granger-kausaliteten skal man være varsom med at overfortolke analysen af impuls-responserne. I fortolkningen af de ortogonale impuls-responser har vi set, at $\lambda_{jk,i}$ kan betragtes som værende responsen i den j 'te variabel efter en impuls i den k 'te komponent af \mathbf{u}_t . Man skal være varsom med at antage, at denne impuls svarer til en impuls alene i den k 'te variabel $X_{k,t}$. Hvis systemet også påvirkes af andre variable, som ikke er med i VARMA-modellen, vil denne

påvirkning ske gennem støjleddene. Dermed bliver fortolkningen af impuls-responserne misvisende, idet en impuls i den k 'te komponent af \mathbf{u}_t nu ikke kan forklares som en impuls i den k 'te variabel alene.

Opsplitning af variansen på forecast error

Vi betragter en stabil VARMA-proces med den ortogonale VMA-repræsentation

$$\mathbf{X}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \Lambda_i \mathbf{u}_{t-i}$$

fra tidligere. Vi så i afsnittet om optimale h -trins prediktorer, at forecast error på den optimale h -trins prediktor kunne udtrykkes som

$$\mathbf{X}_{t+h} - \mathbf{X}_t(h) = \sum_{i=0}^{h-1} \Psi_i \mathbf{w}_{t+h-i}.$$

Da $\Lambda_i = \Psi_i P$, og $\mathbf{u}_t = P^{-1} \mathbf{w}_t$, kan forecast error også udtrykkes som

$$\mathbf{X}_{t+h} - \mathbf{X}_t(h) = \sum_{i=0}^{h-1} \Lambda_i \mathbf{u}_{t+h-i}.$$

Vi nøjes nu med at se på den j 'te komponent i dette, nemlig $X_{j,t+h} - X_{j,t}(h)$, som angiver fejlen af h -trins forecastet af $X_{j,t+h}$. Denne bliver

$$\begin{aligned} X_{j,t+h} - X_{j,t}(h) &= \sum_{i=0}^{h-1} (\lambda_{j1,i} u_{1,t+h-i} + \cdots + \lambda_{jK,i} u_{K,t+h-i}) \\ &= \sum_{i=0}^{h-1} \sum_{k=1}^K \lambda_{jk,i} u_{k,t+h-i} \\ &= \sum_{k=1}^K \sum_{i=0}^{h-1} \lambda_{jk,i} u_{k,t+h-i} \\ &= \sum_{k=1}^K (\lambda_{jk,0} u_{k,t+h} + \cdots + \lambda_{jk,h-1} u_{k,t+1}) \end{aligned}$$

hvor $j = 1, \dots, K$, og $\lambda_{jk,i}$ er det jk 'te element i Λ_i . Vi kan se, at fejlen ved at forecaste den j 'te variabel er en linearkombination af alle komponenterne i $\mathbf{u}_{t+1}, \dots, \mathbf{u}_{t+h}$. Da komponenterne i \mathbf{u}_t er ukorrelerede og har varians 1, gælder der, at

$$E(u_{k,t+h-i} u_{j,t+h-i}) = \begin{cases} 1 & \text{for } k = j \\ 0 & \text{ellers.} \end{cases}$$

Dermed bliver MSE af forecast error

$$\begin{aligned}
 \text{MSE}[X_{j,t}(h)] &= E[(X_{j,t+h} - X_{j,t}(h))^2] \\
 &= \sum_{k=1}^K \sum_{i=0}^{h-1} E[(\lambda_{jk,i} u_{k,t+h-i})^2] \\
 &= \sum_{k=1}^K \sum_{i=0}^{h-1} \lambda_{jk,i}^2 \\
 &= \sum_{k=1}^K (\lambda_{jk,0}^2 + \cdots + \lambda_{jk,h-1}^2). \tag{4.83}
 \end{aligned}$$

Her er MSE af $X_{j,t}(h)$ nu blevet splittet op i K bidrag, hvor $(\lambda_{jk,0}^2 + \cdots + \lambda_{jk,h-1}^2)$ er bidraget fra innovationerne hørende til den k 'te variabel, $k = 1, \dots, K$.

Ved at dele hvert af disse bidrag med $\text{MSE}[X_{j,t}(h)]$, fås

$$\omega_{jk,h} = \frac{(\lambda_{jk,0}^2 + \cdots + \lambda_{jk,h-1}^2)}{\text{MSE}[X_{j,t}(h)]}, \quad j, k = 1, \dots, K. \tag{4.84}$$

Her angiver $\omega_{jk,h}$, hvor stor en andel af fejlen på forecastet af den j 'te variabel, der skyldes innovationerne fra den k 'te variabel. Vi finder $\text{MSE}[X_{j,t}(h)]$ som diagonalelementerne i $\Sigma_X(h)$, hvor

$$\Sigma_X(h) = \text{MSE}[\mathbf{X}_t(h)] = \sum_{i=0}^{h-1} \Psi_i \Sigma_u \Psi_i^* = \sum_{i=0}^{h-1} \Lambda_i \Lambda_i^*. \tag{4.85}$$

Det fås fra [Sætning 4.17](#), at hvis $\omega_{jk,h} = 0$ for $h = pK - p + 1$, er $\omega_{jk,h} = 0$ for alle h . Dette ses ved

$$0 = \omega_{jk,pK-p+1} = \frac{(\lambda_{jk,0}^2 + \cdots + \lambda_{jk,p(K-1)}^2)}{\text{MSE}[X_{j,t}(pK - p + 1)]}$$

hvilket betyder, at $\lambda_{jk,i} = 0$ for $i = 0, \dots, p(K-1)$. Da giver [Sætning 4.17](#), at $\lambda_{jk,i} = 0$ for alle i , og dermed er $\omega_{jk,h} = 0$ for alle h . Så hvis den del af fejlen på at forecaste den j 'te variabel h trin frem, der stammer fra den k 'te variabels innovationer, er lig nul for $h = pK - p + 1$, vil den være lig nul for alle trin h .

Umiddelbart kunne man fristes til at tro, at der er en sammenhæng mellem Granger-kausaltet og dekompositionerne af forecast error således, at hvis en gruppe variable ikke er Granger-Kausale for en anden gruppe af variable, så vil fejlene på forecast af de to grupper heller ikke påvirkes af hinanden. Dette er ikke altid tilfældet. Betragt for eksempel en $\text{VAR}(p)$ -proces, som bliver opdelt i to delprocesser som i [\(4.73\)](#). Hvis \mathbf{Z}_t ikke er Granger-kausal for \mathbf{Y}_t , vil $\Psi_{12,i} = 0$ for $i = 1, 2, \dots$. Dekompositionen af forecast error udtrykkes ved elementerne i

Λ_i , som det ses i (4.84) og (4.85). Og da $\Lambda_i = \Psi_i P$, hvor P stammer fra $\Sigma_u = PP^*$, kan vi ikke antage, at nogle elementer i Λ_i er lig nul, selvom elementerne med samme index i Ψ_i er lig nul. Dette skyldes, at Σ_u ikke nødvendigvis er en diagonalmatrix. Det vil den dog være, hvis der ingen øjeblikkelig kausabilitet mellem Z_t og Y_t er.

Hvis Z_t ikke er Granger-kausal for Y_t , og Σ_u er en diagonalmatrix, vil der dog gælde, at Z_t ikke bidrager til fejlen ved at forecaste Y_t .

Ligesom med de ortogonale impuls-responser skal man være forsigtig med at overfortolke dekompositionerne af forecast errors. Disse afhænger af Λ_i , som igen afhænger af matricen P fra dekompositionen af Σ_u .

Kapitel 5

Empiriske tidsrækker II

5.1 Flerdimensionale tidsrækker

I dette afsnit skal vi se på, hvordan man i praksis analyserer en K -dimensional empirisk tidsrække $\{\mathbf{x}_t\}_{t=1}^N$, hvor $\mathbf{x}_t = (x_{1t}, \dots, x_{Kt})^*$. Her er $\mathbf{x}_1, \dots, \mathbf{x}_N$ de observerede værdier af tidsrækken $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$.

Analysen af en K -dimensional empirisk tidsrække ved hjælp af en d gange integreret VARMA-model

$$\Phi(B)\nabla^d(\mathbf{X}_t - \boldsymbol{\mu}) = \Theta(B)\mathbf{w}_t, \quad t \in \mathbb{Z}, \quad (5.1)$$

hvor $p, d, q \in \mathbb{N}_0$, $|\Phi(z)| = |I_K - \Phi_1 z - \dots - \Phi_p z^p| \neq 0$ hvis $|z| \leq 1$, og $|\Theta(z)| = |I_K - \Theta_1 z - \dots - \Theta_q z^q| \neq 0$ hvis $|z| \leq 1$, deles naturligt op i fire dele:

- (a) identifikation
- (b) estimation
- (c) modelkontrol
- (d) strukturel analyse og forecasting.

Under identifikationen bestemmes først d og dernæst ordenen (p, q) af en VARMA-model for observationerne $\nabla^d \mathbf{x}_{d+1}, \dots, \nabla^d \mathbf{x}_N$, der fremkommer ved at tage differenser d gange. Når ordenen er bestemt, skal parametrene $\Phi_1, \dots, \Phi_p, \Theta_1, \dots, \Theta_q, \boldsymbol{\mu}$ og Σ_w estimeres. Punkt (a) og (b) overlapper lidt hinanden, idet estimaterne af parametrene kan være med til at bestemme modellens orden. Herefter foretages der en modelkontrol ved at se på residualerne i den estimerede model. Hvis modellen ikke er passende, starter vi ved punkt (a) igen og forsøger at bestemme en bedre model. Når vi har fundet en passende model, foretages der så vidt muligt en strukturel analyse ved at estimere impuls-responserne, der laves en forecast error dekomposition og der testes for Granger-kausaltet. Til slut forecastes fremtidige værdier af tidsrækken.

Identifikation

Antag, at vi er i besiddelse af N målinger $\mathbf{x}_1, \dots, \mathbf{x}_N$ af K -dimensionale vektorer, som vi ønsker at beskrive med en VARMA-model. Det første, der skal undersøges er, hvorvidt tidsrækken er stationær. Dette gøres ved at afbilde hver af de K komponenter af observationerne mod tiden. Hvis tidsrækken er stationær, er middelværdifunktionen konstant, og kovariansen mellem observationer afhænger kun af tidsforskellen mellem dem. Når de K komponenter af observationerne afbildes, skal de derfor variere omkring en konstant værdi uden nogen form for systematiske afvigelser.

Hvis den K -dimensionale tidsrække er stationær, er dens K komponenter også stationære, hvilket kan undersøges ved hjælp af Dickey-Fuller testet som omtalt i [Bemærkning 3.2](#). Desuden skal de empiriske autokorrelationsfunktioner for hver af de K komponenter gå mod 0 for store lags.

Hvis en empirisk tidsrække $\{\mathbf{x}_t\}_{t=1}^N$ ikke kan antages at være stationær, tager vi differenser d gange indtil, at det kan antages, at tidsrækken $\{\nabla^d \mathbf{x}_t\}_{t=d+1}^N$ er stationær, og derefter arbejder videre med denne.

I det følgende antages, at $d = 0$.

I forbindelsen med bestemmelsen af ordenen af VARMA-modellen (5.1) betragter PROC VARMA fire forskellige matricer.

Krydskorrelationsmatricer

Kovariansmatricen til lag h , $\Gamma_X(h)$, som ifølge (4.1) er

$$\Gamma_X(h) = E[(\mathbf{X}_t - \boldsymbol{\mu})(\mathbf{X}_{t-h} - \boldsymbol{\mu})^*],$$

estimeres ved

$$\hat{\Gamma}_X(h) = C(h) = \frac{1}{N} \sum_{t=h+1}^N (\mathbf{x}_t - \bar{\mathbf{x}})(\mathbf{x}_{t-h} - \bar{\mathbf{x}})^*, \quad (5.2)$$

hvor $\bar{\mathbf{x}} = (\bar{x}_1, \dots, \bar{x}_K)^* = \frac{1}{N} \sum_{t=1}^N \mathbf{x}_t$ er den empiriske middelværdivektor. Det ij 'te element i $\hat{\Gamma}_X(h)$ vil blive betegnet $\hat{\gamma}_{ij}(h) = c_{ij}(h)$ og er den empiriske kovarians mellem $X_{i,t}$ og $X_{j,t-h}$. Korrelationen mellem $X_{i,t}$ og $X_{j,t-h}$ estimeres ved

$$\hat{\rho}_{ij}(h) = \frac{c_{ij}(h)}{\sqrt{c_{ii}(0)c_{jj}(0)}}, \quad h = 1, 2, \dots \quad (5.3)$$

Kommandoen `corry` i PROC VARMAX udskriver estimerterne for krydskorrelationsmatricerne $\rho_X(h)$ til lag h , som blev defineret i (4.4) på side 88.

Det ses af formel (4.47), at for en VMA(q)-proces er $\Gamma_X(h) = 0$ for $h > q$. Om dette er tilfældet, kan belyses ved at betragte de empiriske krydskorrelationsmatricer $\hat{\rho}_X(h)$.

Bemærkning 5.1. PROC VARMAX beregner størrelserne

$$\hat{\Gamma}_{SAS}(h) = \frac{1}{N} \sum_{t=1}^{N-h} (\mathbf{x}_t - \bar{\mathbf{x}})(\mathbf{x}_{t+h} - \bar{\mathbf{x}})^*, \quad (5.4)$$

i stedet for $\hat{\Gamma}_X(h)$ i (5.2). Dette er dog ikke af afgørende betydning, idet

$$\hat{\Gamma}_{SAS}(h) = \hat{\Gamma}_X(-h) = \hat{\Gamma}_X(h)^*.$$

■

I PROC VARMAX betragtes 3 forskellige slags matricer til at undersøge, om

$$\Phi_{k,k} = \begin{cases} \Phi_p & \text{for } k = p \\ 0 & \text{for } k > p, \end{cases}$$

hvor $\Phi_{k,k}$ er den partielle autoregressionsmatricer, der blev defineret i forbindelse med formel (4.7) på side 89. Dette er tilfældet, hvis tidsrækken følger en VAR(p)-model, og denne cut-off egenskab kan derfor benyttes, når vi skal bestemme ordenen p for en VAR-proces.

Partielle autoregressionsmatricer

I praksis benyttes de *empiriske partielle autoregressionsmatricer* $\hat{\Phi}_{k,k}$ for den observerede tidsrække $\mathbf{x}_1, \dots, \mathbf{x}_N$. Disse størrelser beregnes for $k = 1, 2, \dots$ ved at erstatte kovariansmatricerne $\Gamma_X(h)$ i formel (4.7) med de empiriske kovariansmatricer $C(h)$ fra (5.2), dvs. estimererne $\hat{\Phi}_{k,k}$ findes som løsningerne til de empiriske Yule-Walker ligninger $C(h) = \sum_{j=1}^k \hat{\Phi}_{j,k} C(h-j)$, hvor $h = 1, 2, \dots, k$.

Kommandoen `parcoef` i PROC VARMAX beregner de empiriske partielle autoregressionsmatricer $\hat{\Phi}_{k,k}$.

Partielle krydskorrelationsmatricer

Teorien i dette afsnit er fra Reinsel (1997). Betragt residualvektoren

$$\mathbf{w}_{m,t} = \mathbf{X}_t - \sum_{j=1}^{m-1} \Phi_{j,m-1} \mathbf{X}_{t-j} \quad (5.5)$$

fra VAR-modellen af orden $m-1$. Koefficientmatricerne $\Phi_{1,m-1}, \dots, \Phi_{m-1,m-1}$ er bestemt som omtalt ved beregningerne af de partielle autoregressionsmatricer. Tilsvarende betragtes residualvektoren fra den forlæns autoregression

$$\tilde{\mathbf{w}}_{m,t-m} = \mathbf{X}_{t-m} - \sum_{j=1}^{m-1} \tilde{\Phi}_{j,m-1} \mathbf{X}_{t-m+j}. \quad (5.6)$$

Her er $\tilde{\Phi}_{1,m-1}, \dots, \tilde{\Phi}_{m-1,m-1}$ koefficientmatricer, der bestemmes på samme måde som matricerne $\Phi_{1,m-1}, \dots, \Phi_{m-1,m-1}$. Definér matricerne

$$\Sigma_{m-1} = \text{Cov}(\mathbf{w}_{m,t}) = \Gamma_X(0) - \sum_{j=1}^{m-1} \Phi_{j,m-1} \Gamma_X(-j),$$

og

$$\tilde{\Sigma}_{m-1} = \text{Cov}(\tilde{\mathbf{w}}_{m,t-m}) = \Gamma_X(0) - \sum_{j=1}^{m-1} \tilde{\Phi}_{m-j,m-1} \Gamma_X(m-j).$$

Lad $V(m)$ og $\tilde{V}(m)$ være $K \times K$ diagonalmatricer, hvis diagonalelementer er diagonalelementerne fra henholdsvis Σ_{m-1} og $\tilde{\Sigma}_{m-1}$. Den *partielle krydskorrelationsmatriks* til lag m defineres som korrelationen mellem de to residualvektorer $\mathbf{w}_{m,t}$ og $\tilde{\mathbf{w}}_{m,t-m}$. Vi betegner den P_m og får, at

$$P_m = \text{Cor}(\tilde{\mathbf{w}}_{m,t-m}, \mathbf{w}_{m,t}) = \tilde{V}(m)^{-1/2} \text{E}(\tilde{\mathbf{w}}_{m,t-m} \mathbf{w}_{m,t}) V(m)^{-1/2}.$$

Denne matriks er netop krydskorrelationen mellem \mathbf{X}_{t-m} og \mathbf{X}_t givet de mellemliggende værdier $\mathbf{X}_{t-m+1}, \dots, \mathbf{X}_{t-1}$. Det (i, j) 'te element i P_m repræsenterer den partielle korrelation mellem $X_{i,t-m}$ og X_{jt} givet $\mathbf{X}_{t-m+1}, \dots, \mathbf{X}_{t-1}$.

SAS benytter en lidt anden definition af de partielle korrelationsmatricer (se [SAS 9.1 User's Guide](#) side 1763). Her defineres den partielle korrelationsmatriks til lag m som

$$P(m) = \tilde{\Sigma}_{m-1}^{-1/2} \Phi_{m,m}^* \Sigma_{m-1}^{-1/2},$$

hvor $\Phi_{m,m}$ er den partielle autoregressionsmatriks til lag m . Disse matricer $P(m)$ har samme cut-off egenskab som de partielle autoregressionsmatricer og kan derfor også bruges til at bestemme p .

PROC VARMAX beregner de empiriske partielle krydskorrelationsmatricer $\hat{P}(m)$ ved hjælp af kommandoen `pcorr`.

Partielle kanoniske korrelationer

Den sidste korrelationsfunktion, vi skal se på, er den *partielle kanoniske korrelationsfunktion* (teorien er fra [SAS 9.1 User's Guide](#) side 1765). De partielle kanoniske korrelationer til lag m betegnes $\rho_i(m)$, $i = 1, 2, \dots, K$, og de opfylder, at

$$1 \geq \rho_1(m) \geq \rho_2(m) \geq \dots \geq \rho_K(m) \geq 0.$$

Her er $\rho_i(m)$ den i te partielle kanoniske korrelation mellem \mathbf{X}_t og \mathbf{X}_{t-m} givet de mellemliggende værdier $\mathbf{X}_{t-m+1}, \dots, \mathbf{X}_{t-1}$. Disse korrelationer er de kanoniske korrelationer mellem residualvektorerne $\mathbf{w}_{m,t}$ og $\tilde{\mathbf{w}}_{m,t-m}$, som blev defineret i

(5.5) og (5.6). Man finder de kvadrerede partielle kanoniske korrelationer $\rho_i^2(m)$ som de ordnede egenverdier af matricen

$$\begin{aligned} & \text{Cov}(\mathbf{w}_{m,t}, \mathbf{w}_{m,t})^{-1} \text{Cov}(\mathbf{w}_{m,t}, \tilde{\mathbf{w}}_{m,t-m}) \text{Cov}(\tilde{\mathbf{w}}_{m,t-m}, \tilde{\mathbf{w}}_{m,t-m})^{-1} \text{Cov}(\tilde{\mathbf{w}}_{m,t-m}, \mathbf{w}_{m,t}) \\ &= \Phi_{m,m}^* \tilde{\Phi}_{m,m}^*. \end{aligned}$$

De empiriske partielle kanoniske korrelationer $\hat{\rho}_i(m)$, $i = 1, 2, \dots, K$, beregnes i PROC VARMAX ved hjælp af kommandoen `pcancorr`.

Da $\Phi_{m,m} = 0$ for $m > p$, hvis processen $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ er en VAR(p)-proces, vil der også gælde, at

$$\rho_1(m) = \rho_2(m) = \dots = \rho_K(m) = 0 \quad \text{for } m > p, \quad (5.7)$$

hvorfor vi kan bruge de partielle kanoniske korrelationer til at bestemme p .

Kommandoen `pcancor` beregner teststørrelsen

$$(T - m) \text{tr}(\hat{\Phi}_{m,m}^* \hat{\Phi}_{m,m}^*) \approx (T - m) \sum_{i=1}^K \hat{\rho}_i^2(m) \quad (5.8)$$

til at teste hypotesen, at $\Phi_{m,m} = 0$ i en VAR(p)-model, hvor $m > p$. Store værdier af teststørrelsen, som er approksimativt χ^2 -fordelt med K^2 frihedsgrader, er kritiske.

Eksempel 5.2. Vi betragter igen de 86 målinger af nedbør og vandmængde i [Eksempel 1.3](#) på [side 3](#). I [Kapitel 3](#) så vi, at de marginale tidsrækker begge kan modelleres ved en hvid støj. Nu ser vi på den todimensionale tidsrække bestående af simultane målinger af nedbør og vandmængde. [SAS-udskrift 5.1](#) viser, at MINIC foreslår en VAR(1)-model.

SAS-udskrift 5.1

Minimum Information Criterion Based on AICC						
Lag	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	18.937097	18.937094	18.958183	18.951832	19.048906	19.082062
AR 1	18.897373	18.910935	18.986169	19.013213	19.085839	19.130316
AR 2	18.916547	18.964253	19.075268	19.120191	19.189005	19.195077
AR 3	18.982909	18.991762	19.114257	19.197119	19.259899	19.282735
AR 4	19.082945	19.078613	19.199940	19.289414	19.334572	19.290006
AR 5	19.164135	19.128983	19.243387	19.360564	19.285197	19.403546

Vi undersøger nu, om de forskellige korrelationsmatricer bekræfter dette valg. I [SAS-udskrift 5.2](#) ses de første tre empiriske krydskorrelationsmatricer og nedenunder ses en skematisk repræsentation $\hat{\rho}_X(k)$, $k = 1, \dots, 22$.

SAS-udskrift 5.2

Cross Correlations of Dependent Series																							
	Lag	Variable	nedboer										vandmaengde										
Variable/ Lag	0	nedboer	1.00000										0.77374										
		vandmaengde	0.77374										1.00000										
	1	nedboer	0.09314										0.22497										
		vandmaengde	-0.09102										0.09108										
	2	nedboer	0.14935										0.13455										
		vandmaengde	0.04655										-0.05132										
Schematic Representation of Cross Correlations																							
Variable/ Lag	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
nedboer	++
vandmaengde	++
+ is > 2*std error, - is < -2*std error, . is between																							

Standard error på estimerne approksimeres med $1/\sqrt{N}$. Hvis et element i en empirisk krydskorrelationsmatrix ligger i det approksimative 95 %-konfidensinterval $[-2*\text{std_error}, 2*\text{std_error}]$, kan vi antage, at det tilsvarende teoretiske element er lig nul. Dette angives med en “.” i SAS-udskriften. Hvis der står et “+” eller et “-” er estimatet henholdsvis større eller mindre end $2*\text{std_error}$, og korrelationen kan dermed ikke antages at være lig nul.

Standard error for elementerne i estimerne af de partielle krydskorrelationsmatrixer i SAS-udskrift 5.3 beregnes på samme måde. Udskriften peger tilsyneladende, som MINIC, på en VAR(1)-model, idet $\hat{P}(m) = 0$ for $m > 1$ (bortset fra at $\hat{P}(9)_{11} > 0$). Imidlertid er $2/\sqrt{N} = 2/\sqrt{86} = 0.21568$, så også $\hat{P}(2)_{12}$ er signifikant forskellig fra 0, hvilket dog ikke ses i den skematiske repræsentation af \hat{P} . SAS-udskrift 5.3 peger altså på en VAR(2)-model.

SAS-udskrift 5.3

Partial Cross Correlations				
	Lag	Variable	nedboer	vandmaengde
1		nedboer	0.10633	0.24033
		vandmaengde	-0.25847	0.09447
2		nedboer	0.09462	0.22164
		vandmaengde	0.18121	-0.09597
3		nedboer	0.08852	0.12918
		vandmaengde	0.11334	-0.08209
Schematic Representation of Partial Cross Correlations				

Variable/ Lag	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
nedboer
vandmaengde

+ is > 2*std error, - is < -2*std error, . is between

Teststørrelsen i (5.8) og de tilsvarende testsandsynligheder ses i de to sidste søjler i SAS-udskrift 5.4. Testet for lag 1 giver signifikans, hvilket næsten også er tilfældet med testet for lag 2.

SAS-udskrift 5.4

Partial Canonical Correlations					
Lag	Correlation1	Correlation2	DF	Chi-Square	Pr > ChiSq
1	0.27968	0.25802	4	12.31	0.0152
2	0.24306	0.20260	4	8.41	0.0776
3	0.15681	0.13971	4	3.66	0.4538
4	0.10509	0.00757	4	0.91	0.9231
5	0.21875	0.10604	4	4.79	0.3099
6	0.13929	0.00673	4	1.56	0.8167
7	0.15499	0.06322	4	2.21	0.6966
8	0.17319	0.01002	4	2.35	0.6721
9	0.38378	0.01406	4	11.36	0.0228
10	0.21200	0.04695	4	3.58	0.4653

Som resultat af identifikationen vil vi estimere i en VAR(2)-model, idet vi i denne kan undersøge, om en reduktion til VAR(1)-modellen er mulig.

Eksempel 5.3. Vi betragter den tredimensionale tidsrække bestående af de kvartalsvise opgørelser af henholdsvis faste investeringer, disponibel indkomst og udgifter til forbrug i Vesttyskland i perioden 1960–1978. SAS-udskrifterne 3.10–3.12 på side 55 antydede, at de marginale rækker af førstedifferenser af logaritmerne kan analyseres med henholdsvis en AR(4)-model, en hvid støj og en AR(2)-model.

For at forenkle notationen på SAS-udskrifterne lader vi $\mathbf{v} = (v_1, v_2, v_3)^*$ betegne vektoren af førstedifferenser af logaritmerne.

SAS-udskrift 5.5

Minimum Information Criterion Based on AICC						
Lag	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	-24.35433	-24.59227	-24.5270	-24.3435	-24.13765	-24.01170
AR 1	-24.39384	-24.43362	-24.29225	-24.05727	-23.88124	-23.71074
AR 2	-24.48842	-24.42008	-24.12775	-23.87676	-23.63288	-23.32711
AR 3	-24.21081	-24.16972	-23.85596	-23.60945	-23.23246	-22.89423
AR 4	-24.02673	-24.08737	-23.80716	-23.38776	-22.99911	-22.55897
AR 5	-23.84811	-23.82078	-23.35730	-22.98851	-22.43610	-22.00246

I [SAS-udskrift 5.5](#) ses det, at MINIC anbefaler en VMA(1)-model, i hvilken der gælder, at $\rho_X(h) = 0$ for $h > 1$. Anbefaligen er noget overraskende, idet [SAS-udskrift 5.6](#) viser, at mindst 3 elementer i $\hat{\rho}_X(2)$ er signifikant større end 0. (Da vi har $N - 1 = 76 - 1 = 75$ førstedifferenser, er $2 * \text{std_error} = 2 / \sqrt{75} = 0.23094$.)

SAS-udskrift 5.6

Cross Correlations of Dependent Series					
Lag	Variable	v1	v2	v3	
0	v1	1.000000	0.09111	0.26622	
	v2	0.09111	1.000000	0.45569	
	v3	0.26622	0.45569	1.000000	
1	v1	-0.19603	0.17104	-0.04882	
	v2	0.11334	0.03050	0.18423	
	v3	0.12961	0.20221	-0.07928	
2	v1	-0.04054	0.12162	0.24578	
	v2	0.10964	0.09388	0.36568	
	v3	0.11531	0.02286	0.27195	

Schematic Representation of Cross Correlations																				
Variable/ Lag	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
v1	+++	...	+..	-..-	-..
v2	..+++	+.-
v3	++++-	-.

+ is > 2*std error, - is < -2*std error, . is between

Den næstmindste værdi i [SAS-udskrift 5.5](#) antages for en VAR(2)-model. Da hverken de empiriske partielle krydskorrelationsmatricer $\hat{P}(m)$ i [SAS-udskrift 5.7](#) eller de empiriske partielle kanoniske korrelationer $\hat{\rho}_i(m)$ i [SAS-udskrift 5.8](#) strider mod antagelsen at $\Phi_{m,m} = 0$ for $m > 2$, vælger vi en VAR(2)-model.

SAS-udskrift 5.7

Partial Cross Correlations				
Lag	Variable	v1	v2	v3
1	v1	-0.20655	0.19198	-0.05176
	v2	0.08519	-0.07632	0.23398
	v3	0.16740	0.22687	-0.23651
2	v1	-0.12106	0.14976	0.18279
	v2	0.06375	-0.08329	0.43637
	v3	0.17196	-0.03229	0.05003
3	v1	0.02108	0.01997	0.07967
	v2	0.02712	0.13664	0.13103
	v3	-0.09934	-0.01374	0.12885

Schematic Representation of Partial Cross Correlations																			
Variable/ Lag	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
v1	+
v2+
v3

+ is > 2*std error, - is < -2*std error, . is between

SAS-udskrift 5.8

Partial Canonical Correlations						
Lag	Correlation1	Correlation2	Correlation3	DF	Chi-Square	Pr > ChiSq
1	0.43103	0.28357	0.13878	9	21.12	0.0121
2	0.47784	0.26018	0.09492	9	22.27	0.0081
3	0.22504	0.13568	0.03705	9	5.07	0.8281
4	0.38831	0.12923	0.02978	9	11.95	0.2159
5	0.23097	0.03856	0.00238	9	3.84	0.9217
6	0.39169	0.10968	0.00668	9	11.42	0.2481
7	0.17091	0.10499	0.02958	9	2.80	0.9719
8	0.39045	0.10304	0.04563	9	11.07	0.2713
9	0.22075	0.15016	0.09990	9	5.36	0.8016
10	0.38020	0.13086	0.02211	9	10.54	0.3085

Estimation

Det antages nu, at vi har bestemt ordenen p og q af den empiriske tidsrække $\mathbf{x}_1, \dots, \mathbf{x}_N$. Næste skridt bliver at estimere parametrene i modellen. Det drejer sig om koefficientmatricerne Φ_1, \dots, Φ_p fra VAR-delen, $\Theta_1, \dots, \Theta_q$ fra VMA-delen, middelværdien μ og kovariansmatricen for den hvide støj, Σ_w .

I [Opgave 14](#) på [side 162](#) vises, at en K -dimensional VARMA(p, q)-proces kan betragtes som en $K(p + q)$ -dimensional VAR(1)-proces. Derfor vil her kun se på, hvordan man estimerer koefficientmatricerne i en VAR-model.

Vi betragter VAR(p)-processen

$$\mathbf{X}_t = \boldsymbol{\nu} + \Phi_1 \mathbf{X}_{t-1} + \dots + \Phi_p \mathbf{X}_{t-p} + \mathbf{w}_t.$$

SAS estimerer elementerne i koefficientmatricerne Φ_1, \dots, Φ_p og konstantleddet $\boldsymbol{\nu}$ ved *conditional least-squares* (CLS). Det antages, at vi har observeret de N værdier af tidsrækken $\mathbf{x}_1, \dots, \mathbf{x}_N$. Desuden antages det, at vi kender værdien af de p foregående variable fra før tid $t = 1$, nemlig $\mathbf{x}_{-p+1}, \dots, \mathbf{x}_0$.

Hvis A er en $k \times m$ matrix, betegner $\text{vec}(A)$ i det følgende en km -dimensional vektor bestående af de m søjler i A stablet oven på hinanden.

Med denne notation kan VAR(p)-modellen skrives som en flerdimensional multipel regressionsmodel

$$\mathbf{X} = \mathbf{B}\mathbf{T} + \mathbf{W} \quad \text{eller} \quad \mathbf{x} = (\mathbf{T}^* \otimes \mathbf{I}_K)\boldsymbol{\beta} + \mathbf{w},$$

hvor

$$\begin{aligned} \mathbf{X} &= (\mathbf{x}_1, \dots, \mathbf{x}_N), \\ \mathbf{B} &= (\boldsymbol{\nu}, \Phi_1, \dots, \Phi_p), \\ \mathbf{T}_t &= (1, \mathbf{x}_t^*, \dots, \mathbf{x}_{t-p+1}^*)^*, \\ \mathbf{T} &= (\mathbf{T}_0, \dots, \mathbf{T}_{N-1}) \\ \mathbf{W} &= (\mathbf{w}_1, \dots, \mathbf{w}_N), \\ \mathbf{x} &= \text{vec}(\mathbf{X}), \\ \boldsymbol{\beta} &= \text{vec}(\mathbf{B}), \end{aligned}$$

og

$$\mathbf{w} = \text{vec}(\mathbf{W}).$$

Vi kan se, at \mathbf{T}_0 kommer til at bestå af de p variable fra før tid $t = 1$. Matricen \mathbf{B} og dermed også vektoren $\boldsymbol{\beta}$ består af de parametre, som vi ønsker at estimere.

CLS-estimatet $\hat{\boldsymbol{\beta}}$ for $\boldsymbol{\beta}$, er

$$\hat{\boldsymbol{\beta}} = ((\mathbf{T}\mathbf{T}^*)^{-1} \mathbf{X} \otimes \mathbf{I}_K)\mathbf{x},$$

og udtrykt ved $\hat{\mathbf{B}}$ er CLS-estimatet

$$\hat{\mathbf{B}} = \mathbf{X}\mathbf{T}^*(\mathbf{T}\mathbf{T}^*)^{-1}.$$

Estimatet for Σ_w er

$$\begin{aligned} \hat{\Sigma}_w &= (N - (Kp + 1))^{-1} \sum_{t=1}^N \hat{\mathbf{w}}_t \hat{\mathbf{w}}_t^* \\ &= (N - (Kp + 1))^{-1} \hat{\mathbf{W}} \hat{\mathbf{W}}^* \\ &= (N - (Kp + 1))^{-1} (\mathbf{X} - \hat{\mathbf{B}}\mathbf{T})(\mathbf{X} - \hat{\mathbf{B}}\mathbf{T})^*, \end{aligned}$$

hvor $\hat{\mathbf{w}}_t = \mathbf{x} - (\mathbf{T}^* \otimes \mathbf{I}_K)\hat{\boldsymbol{\beta}}$ er residualerne til tid t .

Vi kan estimere middelværdien $\boldsymbol{\mu}$ ved gennemsnittet af observationerne, dvs.

$$\hat{\boldsymbol{\mu}} \leftarrow \frac{1}{N} \sum_{t=1}^N \mathbf{x}_t.$$

Da $\boldsymbol{\mu} = (\mathbf{I}_K - \Phi_1 - \dots - \Phi_p)^{-1} \boldsymbol{\nu}$, kan vi også finde estimatet for $\boldsymbol{\mu}$ ud fra $\hat{\Phi}_1, \dots, \hat{\Phi}_p$ og $\hat{\boldsymbol{\nu}}$, som

$$\hat{\boldsymbol{\mu}} = (\mathbf{I}_K - \hat{\Phi}_1 - \dots - \hat{\Phi}_p)^{-1} \hat{\boldsymbol{\nu}}.$$

Dette var estimationen i en K -dimensional $\text{VAR}(p)$ -model. Som nævnt ovenfor kan en $\text{VARMA}(p, q)$ -model opfattes som en $K(p+q)$ -dimensional $\text{VAR}(1)$ -model, så ovenstående dækker principielt også estimationen i en $\text{VARMA}(p, q)$ -model.

Når vi har beregnet estimerne $\hat{\Phi}_1, \dots, \hat{\Phi}_p, \hat{\Theta}_1, \dots, \hat{\Theta}_q$ kan vi beregne estimerne $\hat{\Psi}_i$ for koefficientmatricerne Ψ_i i $\text{VARMA}(p, q)$ -processens VMA-repræsentation

$$\mathbf{X}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \Psi_i \mathbf{w}_{t-i}.$$

Dette gøres rekursivt med formel (4.52) på side 104 ved at benytte estimerne for Φ_1, \dots, Φ_p og estimerne for $\Theta_1, \dots, \Theta_q$ i stedet for de sande værdier. På samme måde beregnes estimerne $\hat{\Pi}_i$ for koefficientmatricerne Π_i i $\text{VARMA}(p, q)$ -processens VAR-repræsentation

$$\mathbf{X}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \Pi_i \mathbf{X}_{t-i} + \mathbf{w}_t.$$

Dette gøres med rekursionen i formel (4.54).

For en K -dimensional $\text{VAR}(p)$ -model gælder (Lütkepohl, 2007, Proposition 3.1) der, at hvis $\mathbf{T} \mathbf{T}^* / N \xrightarrow[N \rightarrow \infty]{d} \Gamma_p$, Γ_p er ikke-singulær, og den hvide støj er uafhængig, så er

$$\sqrt{N}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \sqrt{N} \text{vec}(\hat{\mathbf{B}} - \mathbf{B}) \xrightarrow[N \rightarrow \infty]{d} N_{K(Kp+1)}(0, \Gamma_p^{-1} \otimes \Sigma_u). \quad (5.9)$$

Igen gælder et lignende resultat for en $\text{VARMA}(p, q)$ -model.

Hvis estimerne er asymptotisk normalfordelte, har vi mulighed for at teste hypoteserne

$$\Phi_{jk,i} = 0 \quad \text{for } i = 1, \dots, p, \quad j, k = 1, \dots, K, \quad (5.10)$$

$$\Theta_{jk,i} = 0 \quad \text{for } i = 1, \dots, q, \quad j, k = 1, \dots, K, \quad (5.11)$$

og

$$\nu_i = 0 \quad \text{for } i = 1, \dots, K \quad (5.12)$$

med t -tests.

Eksempel 5.4. Resultatet af identifikation på side 128 var en $\text{VAR}(2)$ -model for de simultane målinger af nedbør og vandmængde. SAS-udskrift 5.9 viser resultatet af estimationen i denne model.

SAS-udskrift 5.9

The VARMAX Procedure

Type of Model	VAR(2)
Estimation Method	Least Squares Estimation

Constant Estimates

Variable	Constant
ned	342.86372
vand	1099.39324

AR Coefficient Estimates

Lag	Variable	ned	vand
1	ned	0.44055	-0.18129
	vand	0.96864	-0.24644
2	ned	0.31389	-0.04040
	vand	1.09428	-0.30913

Schematic Representation of Parameter Estimates

Variable/ Lag	C	AR1	AR2
ned	+	+ -	..
vand	+	+	+

```
+ is > 2*std error,  -
is < -2*std error,  .
is between,  * is N/A
```

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > t	Variable
ned	CONST1	342.86372	64.39886	5.32	0.0001	1
	AR1_1_1	0.44055	0.17699	2.49	0.0149	ned(t-1)
	AR1_1_2	-0.18129	0.06395	-2.83	0.0058	vand(t-1)
	AR2_1_1	0.31389	0.17733	1.77	0.0806	ned(t-2)
	AR2_1_2	-0.04040	0.06435	-0.63	0.5320	vand(t-2)
vand	CONST2	1099.39324	173.60842	6.33	0.0001	1
	AR1_2_1	0.96864	0.47715	2.03	0.0457	ned(t-1)
	AR1_2_2	-0.24644	0.17240	-1.43	0.1568	vand(t-1)
	AR2_2_1	1.09428	0.47806	2.29	0.0248	ned(t-2)
	AR2_2_2	-0.30913	0.17347	-1.78	0.0786	vand(t-2)

Covariances of Innovations

Variable	ned	vand
ned	6979.94480	14489.83329
vand	14489.83329	50726.77269

Den estimerede VAR(2)-model er med indlysende forkortelser

$$\begin{pmatrix} \text{ned}_t \\ \text{vand}_t \end{pmatrix} = \begin{pmatrix} 342.86372 \\ 1099.39324 \end{pmatrix} + \begin{pmatrix} 0.44055 & -0.18129 \\ 0.96864 & -0.24644 \end{pmatrix} \begin{pmatrix} \text{ned}_{t-1} \\ \text{vand}_{t-1} \end{pmatrix} + \begin{pmatrix} 0.31389 & -0.04040 \\ 1.09428 & -0.30913 \end{pmatrix} \begin{pmatrix} \text{ned}_{t-2} \\ \text{vand}_{t-2} \end{pmatrix} + \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix}, \quad (5.13)$$

hvor

$$\Sigma_w \leftarrow \begin{pmatrix} 6979.94480 & 14489.83329 \\ 14489.83329 & 50726.77269 \end{pmatrix}.$$

Som det ses af såvel den skematiske repræsentation som tabellen ovenfor, er estimerne \hat{v} , $\hat{\Phi}_{11,1}$, $\hat{\Phi}_{12,1}$, $\hat{\Phi}_{21,1}$ og $\hat{\Phi}_{21,2}$, angivet med rødt i (5.13), signifikant forskellige fra 0. VAR(2)-modellen kan derfor ikke reduceres til den VAR(1)-model, som MINIC foreslog.

Eksempel 5.5. For $v = (v_1, v_2, v_3)^*$, vektoren af førstedifferenser af logaritmerne for henholdsvis faste investeringer, disponibel indkomst og udgifter til forbrug, identificerede vi på side 130 en VAR(2)-model.

SAS-udskrift 5.10

The VARMAX Procedure						
Type of Model			VAR(2)			
Estimation Method			Least Squares Estimation			
Model Parameter Estimates						
Equa- tion	Parameter	Estimate	Standard Error	t Value	Pr > t	Variable
v1	CONST1	-0.01672	0.01723	-0.97	0.3352	1
	AR1_1_1	-0.31963	0.12546	-2.55	0.0132	v1(t-1)
	AR1_1_2	0.14599	0.54567	0.27	0.7899	v2(t-1)
	AR1_1_3	0.96122	0.66431	1.45	0.1526	v3(t-1)
	AR2_1_1	-0.16055	0.12491	-1.29	0.2032	v1(t-2)
	AR2_1_2	0.11460	0.53457	0.21	0.8309	v2(t-2)
	AR2_1_3	0.93439	0.66510	1.40	0.1647	v3(t-2)
v2	CONST2	0.01577	0.00437	3.60	0.0006	1
	AR1_2_1	0.04393	0.03186	1.38	0.1726	v1(t-1)
	AR1_2_2	-0.15273	0.13857	-1.10	0.2744	v2(t-1)
	AR1_2_3	0.28850	0.16870	1.71	0.0919	v3(t-1)
	AR2_2_1	0.05003	0.03172	1.58	0.1195	v1(t-2)
	AR2_2_2	0.01917	0.13575	0.14	0.8882	v2(t-2)
	AR2_2_3	-0.01020	0.16890	-0.06	0.9520	v3(t-2)

v3	CONST3	0.01293	0.00353	3.67	0.0005	1
	AR1_3_1	-0.00242	0.02568	-0.09	0.9251	v1(t-1)
	AR1_3_2	0.22481	0.11168	2.01	0.0482	v2(t-1)
	AR1_3_3	-0.26397	0.13596	-1.94	0.0565	v3(t-1)
	AR2_3_1	0.03388	0.02556	1.33	0.1896	v1(t-2)
	AR2_3_2	0.35491	0.10941	3.24	0.0019	v2(t-2)
	AR2_3_3	-0.02223	0.13612	-0.16	0.8708	v3(t-2)

Schematic Representation of Parameter Estimates				
Variable/ Lag	C	AR1	AR2	
v1	
v2	+	
v3	+	..+	..+	

+ is > 2*std error, -
is < -2*std error, .
is between, * is N/A

Covariances of Innovations			
Variable	v1	v2	v3
v1	0.00213	0.00007	0.00012
v2	0.00007	0.00014	0.00006
v3	0.00012	0.00006	0.00009

SAS-udskrift 5.10 viser, at den estimerede VAR(2)-model er

$$\begin{aligned}
 \begin{pmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{pmatrix} &= \begin{pmatrix} -0.01672 \\ 0.01577 \\ 0.01293 \end{pmatrix} + \begin{pmatrix} -0.31963 & 0.14599 & 0.96122 \\ 0.04393 & -0.15273 & 0.28850 \\ -0.00242 & 0.22481 & -0.26397 \end{pmatrix} \begin{pmatrix} v_{1,t-1} \\ v_{2,t-1} \\ v_{3,t-1} \end{pmatrix} \\
 &+ \begin{pmatrix} -0.16055 & 0.11460 & 0.93439 \\ 0.05003 & 0.01917 & -0.01020 \\ 0.03388 & 0.35491 & -0.02223 \end{pmatrix} \begin{pmatrix} v_{1,t-2} \\ v_{2,t-2} \\ v_{3,t-2} \end{pmatrix} + \begin{pmatrix} w_{1t} \\ w_{2t} \\ w_{3t} \end{pmatrix},
 \end{aligned}
 \tag{5.14}$$

hvor

$$\Sigma_w \leftarrow \begin{pmatrix} 0.00213 & 0.00007 & 0.00012 \\ 0.00007 & 0.00014 & 0.00006 \\ 0.00012 & 0.00006 & 0.00009 \end{pmatrix}.$$

Med rødt er angivet de parametre, der ved t -test afviger signifikant fra 0, hvilket også kan ses af den skematiske repræsentation ovenfor. Estimerterne i (5.14) antyder, at hypotesen $H_0 : \Phi_{12,1} = \Phi_{13,1} = \Phi_{12,2} = \Phi_{13,2} = 0$ ikke forkastes i modellen

$$\mathbf{v}_t = \Phi_1 \mathbf{v}_{t-1} + \Phi_2 \mathbf{v}_{t-2} + \mathbf{w}_t$$

således, at førstekomponenten af \mathbf{v} ikke afhænger af de to andre komponenter. Vi skal senere se på test for H_0 .

Modelkontrol

Når vi på grundlag af observationerne $\mathbf{x}_1, \dots, \mathbf{x}_N$ har estimeret parametrene i VARMA(p, q)-processen $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$, skal vi lave modelkontrol. Vi ønsker at kontrollere, om det er rimeligt at antage, at de observerede værdier stammer fra den fundne VARMA(p, q)-model. Her betragter vi modelkontrol i en VAR(p)-proces, hvilket ikke er nogen begrænsning, idet en K -dimensional VARMA(p, q)-proces kan formuleres som $K(p + q)$ -dimensional VAR(1)-proces i (B.13), se Opgave 14.

Vi starter med at undersøge, hvorvidt residualerne $\hat{\mathbf{w}}_{1+p}, \dots, \hat{\mathbf{w}}_N$ fra VAR(p)-processen, hvor

$$\hat{\mathbf{w}}_t = \mathbf{x}_t - (\hat{\Phi}_1 \mathbf{x}_{t-1} + \hat{\Phi}_2 \mathbf{x}_{t-2} + \dots + \hat{\Phi}_p \mathbf{x}_{t-p}), \quad t = 1 + p, \dots, N,$$

opfører sig som en hvid støj. Vi skal derfor have undersøgt, om det kan antages, at residualerne har middelværdi nul og er ukorreleerede. Dette gøres ved først at plotte $\hat{\mathbf{w}}_{1+p}, \dots, \hat{\mathbf{w}}_N$ mod tiden. På et sådan plot skulle vi gerne kunne se, at (komponenterne af) residualerne varierer omkring den konstante værdi nul og ikke udviser nogen form for systematiske afvigelser herfra.

Dernæst skal det undersøges, om residualerne kan antages at være ukorreleerede. Blandt de test, som SAS udfører for at undersøge, om det kan antages, at residualerne er en hvid støj, er Portmanteau-testet i formel (5.15).

Lad $\Gamma_w(h) = E(\mathbf{w}_t \mathbf{w}_{t-h}^*)$ være krydskovariansmatricen for den hvide støj til lag h . Denne estimeres med

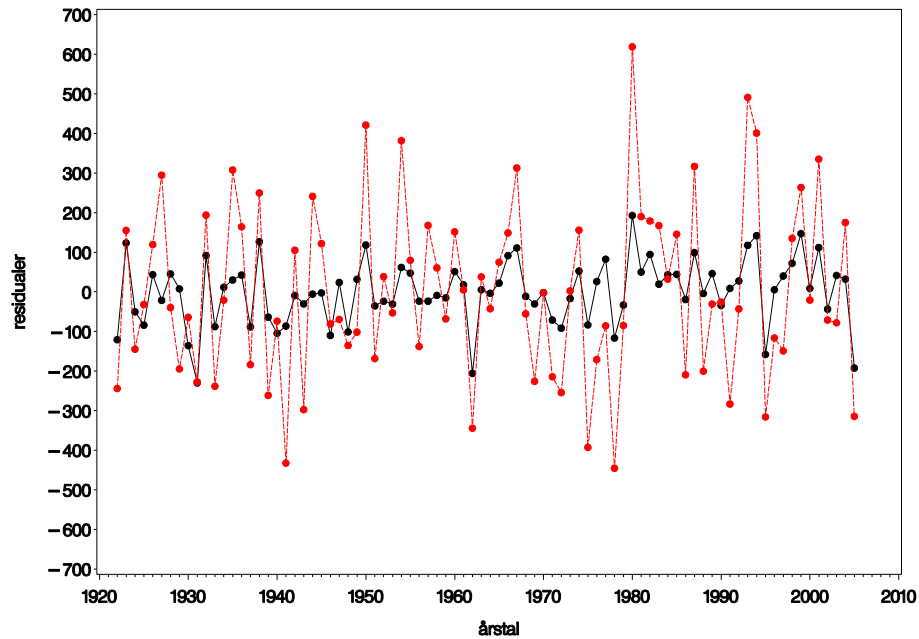
$$C_w(h) = \hat{\Gamma}_w(h) = \frac{1}{N} \sum_{t=h+1}^N \hat{\mathbf{w}}_t \hat{\mathbf{w}}_{t-h}^*,$$

hvor $h = 0, 1, \dots, k < N$. Det ij 'te element i $C_w(h)$ betegnes $c_{ij,h}$ og er den empiriske kovarians mellem $w_{i,t}$ og $w_{j,t-h}$. Lad $R_w(h)$ betegne krydskorrelationsmatricen for den hvide støj til lag h . Denne estimeres med

$$\hat{R}_w(h) = \hat{D}^{-1} C_w(h) \hat{D}^{-1},$$

hvor \hat{D} er en $K \times K$ diagonalmatriks, hvis diagonalelementer er kvadratrødderne af diagonalelementerne i $C_w(0)$. Dermed bliver det ij 'te element i $\hat{R}_w(h)$

$$\hat{r}_{ij,h} = \frac{c_{ij,h}}{\sqrt{c_{ii,0}} \sqrt{c_{jj,0}}},$$



Figur 5.1: Residualerne i VAR(2)-modellen i (5.13) for de 86 simultane målinger af nedbør og vandmængde.

SAS-udskrift 5.12

[illegible]

+ is $> 2 \times \text{std error}$, - is $< -2 \times \text{std error}$, . is between

Partial Canonical Correlations

Lag	Correlation1	Correlation2	DF	Chi-Square	Pr > ChiSq
1	0.04533	0.02500	4	0.22	0.9943
2	0.05933	0.03066	4	0.37	0.9852
3	0.14825	0.07626	4	2.25	0.6897
4	0.17075	0.03374	4	2.42	0.6584
5	0.22095	0.13011	4	5.19	0.2680
6	0.15507	0.00435	4	1.88	0.7583
7	0.16240	0.05964	4	2.30	0.6799
8	0.22131	0.01399	4	3.74	0.4427
9	0.33044	0.03547	4	8.28	0.0817
10	0.17668	0.07262	4	2.70	0.6092
11	0.20970	0.08794	4	3.77	0.4374
12	0.21977	0.05970	4	3.73	0.4432
13	0.11846	0.01028	4	1.00	0.9092
14	0.08519	0.03091	4	0.57	0.9658
15	0.12551	0.10647	4	1.87	0.7598
16	0.31860	0.00725	4	6.91	0.1409

17	0.22627	0.07861	4	3.84	0.4275
18	0.13801	0.07116	4	1.59	0.8104
19	0.13059	0.09501	4	1.70	0.7916
20	0.21082	0.10196	4	3.51	0.4764
21	0.23515	0.00682	4	3.49	0.4799
22	0.14811	0.05337	4	1.54	0.8201

Heller ikke Portmanteau testene i [SAS-udskrift 5.13](#) strider imod, at residualerne kan beskrives ved hjælp af en hvid støj.

SAS-udskrift 5.13

The VARMAX Procedure				
Portmanteau Test for Cross Correlations of Residuals				
Up To Lag	DF	Chi-Square	Pr > ChiSq	
3	4	3.03	0.5520	
4	8	5.60	0.6919	
5	12	10.81	0.5452	
6	16	13.21	0.6571	
7	20	15.90	0.7226	
8	24	19.71	0.7131	
9	28	28.03	0.4630	
10	32	32.24	0.4549	
11	36	37.88	0.3836	
12	40	40.68	0.4403	
13	44	42.93	0.5173	
14	48	43.47	0.6590	
15	52	47.98	0.6327	
16	56	55.98	0.4755	
17	60	59.49	0.4942	
18	64	61.91	0.5508	
19	68	65.07	0.5782	
20	72	67.28	0.6353	
21	76	69.76	0.6796	
22	80	71.34	0.7447	

Som en yderligere kontrol af VAR(2)-modellen i (5.13), analyseres residualerne fra denne med en VAR(1)-model. [SAS-udskrift 5.14](#) viser, at t -testene ikke kan afvise, at alle parametre er 0, og desuden, at et samlet test, for at koefficientmatricen er 0, giver en testsandsynlighed på 0.9940, så residualerne kan beskrives ved hjælp af en hvid støj.

SAS-udskrift 5.14

The VARMAX Procedure		
Type of Model	VAR(1)	
Estimation Method	Least Squares Estimation	

Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > t	Variable
res_n	CONST1	1.53247	8.98566	0.17	0.8650	1
	AR1_1_1	-0.01406	0.17707	-0.08	0.9369	res_n(t-1)
	AR1_1_2	-0.01210	0.06422	-0.19	0.8510	res_v(t-1)
res_v	CONST2	3.05805	24.36801	0.13	0.9004	1
	AR1_2_1	0.03829	0.48019	0.08	0.9366	res_n(t-1)
	AR1_2_2	-0.05511	0.17416	-0.32	0.7525	res_v(t-1)
Testing of the Parameters						
	Test	DF	Chi-Square	Pr > ChiSq		
	1	4	0.23	0.9940		

Eksempel 5.7. Figur 5.2 viser residualerne i VAR(2)-modellen i (5.14) for første-differenser af logaritmerne for henholdsvis faste investeringer, disponibel indkomst og udgifter til forbrug. De tre komponenter af residualerne varierer omkring 0 og forekommer ikke at være korrelerede, hvilket bekræftes af den skematiske repræsentation af de empiriske krydskorrelationer for residualerne i SAS-udskrift 5.15. Heller ikke de empiriske partielle krydskorrelationer eller de empiriske partielle kanoniske korrelationer i SAS-udskrift 5.16 strider mod, at residualerne kan beskrives ved hjælp af en hvid støj. På SAS-udskrift 5.15 ses også Portmanteau-testene i (5.15). Disse giver heller ikke anledning til at betvivle modellen.

	16	126	120.14	0.6304
	17	135	126.64	0.6840
	18	144	131.87	0.7568
	19	153	144.67	0.6724

SAS-udskrift 5.16

Schematic Representation of Partial Cross Correlations																			
Variable/ Lag	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
res_v1	+
res_v2	+
res_v3
+ is > 2*std error, - is < -2*std error, . is between																			
Partial Canonical Correlations																			
Lag	Correlation1	Correlation2	Correlation3	DF	Chi-Square	Pr > ChiSq													
1	0.10134	0.04425	0.01966	9	0.91	0.9996													
2	0.12047	0.03466	0.00596	9	1.12	0.9991													
3	0.29113	0.12069	0.03250	9	7.03	0.6344													
4	0.38440	0.13650	0.00151	9	11.48	0.2441													
5	0.25473	0.11002	0.09662	9	5.87	0.7528													
6	0.24449	0.10080	0.04095	9	4.80	0.8515													
7	0.21260	0.11278	0.03089	9	3.89	0.9188													
8	0.40172	0.08702	0.02485	9	11.02	0.2742													
9	0.24252	0.14495	0.06140	9	5.35	0.8028													
10	0.42768	0.18734	0.00048	9	13.73	0.1321													
11	0.32783	0.20505	0.03281	9	9.34	0.4068													
12	0.33557	0.23213	0.04090	9	10.26	0.3300													
13	0.26780	0.13476	0.07348	9	5.72	0.7679													
14	0.40767	0.20835	0.08705	9	12.81	0.1712													
15	0.34149	0.11596	0.00126	9	7.54	0.5807													
16	0.32771	0.17081	0.02299	9	7.81	0.5529													
17	0.31824	0.14458	0.01880	9	6.86	0.6515													
18	0.36233	0.16616	0.10607	9	9.36	0.4049													
19	0.33899	0.21524	0.01168	9	8.71	0.4640													

Som et sidste check af VAR(2)-modellen i (5.14), er residualerne herfra analyseret med en VAR(1)-model. SAS-udskrift 5.17 herfra viser, at samtlige parametre i denne VAR(1)-model kan antages at være 0, dvs. residualerne kan beskrives ved hjælp af en hvid støj. Af udskriften ses også, at et samlet test, for at de ni elementer i koefficientmatricen i VAR(1)-modellen er 0, ikke forkaster denne hypotese, idet testsandsynligheden er 0.9996.

SAS-udskrift 5.17

The VARMAX Procedure						
Type of Model				VAR(1)		
Estimation Method				Least Squares Estimation		
Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > t	Variable
res_v1	CONST1	-0.00013	0.00536	-0.02	0.9808	1
	AR1_1_1	0.01944	0.12781	0.15	0.8796	res_v1(t-1)
	AR1_1_2	-0.03089	0.57453	-0.05	0.9573	res_v2(t-1)
	AR1_1_3	-0.05860	0.74729	-0.08	0.9377	res_v3(t-1)
res_v2	CONST2	0.00007	0.00136	0.05	0.9617	1
	AR1_2_1	0.00438	0.03235	0.14	0.8927	res_v1(t-1)
	AR1_2_2	0.05185	0.14541	0.36	0.7225	res_v2(t-1)
	AR1_2_3	-0.12947	0.18913	-0.68	0.4960	res_v3(t-1)
res_v3	CONST3	-0.00008	0.00109	-0.07	0.9431	1
	AR1_3_1	0.00135	0.02593	0.05	0.9586	res_v1(t-1)
	AR1_3_2	0.00885	0.11658	0.08	0.9397	res_v2(t-1)
	AR1_3_3	-0.10752	0.15164	-0.71	0.4807	res_v3(t-1)
Testing of the Parameters						
Test		DF	Chi-Square	Pr > ChiSq		
1		9	0.89	0.9996		

Strukturel analyse og forecasting

Når vi har fået fittet en model til data, og modelkontrollen har vist, at denne er passende, er de næste skridt strukturel analyse og/eller forecasting. I den strukturelle analyse undersøges de indbyrdes forhold mellem komponenterne X_{1t}, \dots, X_{Kt} i \mathbf{X}_t . Vi nøjes her med at se på, hvordan SAS udfører tests for Granger-kausalitet.

Det antages i det følgende, at vi under identifikationen og estimationen har fundet frem til, at $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ er en VARMA(p, q)-proces. Lad $\{\mathbf{X}_t\}_{t \in \mathbb{Z}}$ være delt op i de to delprocesser $\{\mathbf{Y}_t\}_{t \in \mathbb{Z}}$ og $\{\mathbf{Z}_t\}_{t \in \mathbb{Z}}$ som i (4.73) på side 112. Vi ønsker nu at undersøge, hvorvidt $\{\mathbf{Z}_t\}_{t \in \mathbb{Z}}$ er Granger-kausal for $\{\mathbf{Y}_t\}_{t \in \mathbb{Z}}$. Vi så i (ii) i Sætning 4.13 på side 113, hvordan vi kunne afgøre dette ved at se på koefficientmatricerne Ψ_i i processens VMA-repræsentation.

Hvis $q = 0$, så vi i Sætning 4.14 på side 115, hvordan vi kunne afgøre det ved at se på koefficientmatricerne fra processens VAR-repræsentation. Under estimationen blev estimerne for Φ_1, \dots, Φ_p fundet sammen med teststørrelser og p -værdier for hypoteserne, at $\Phi_{ij,k} = 0$ for $i, j = 1, \dots, K$ og $k = 1, \dots, p$. Så når vi

skal afgøre, om $\{Z_t\}_{t \in \mathbb{Z}}$ ikke er Granger-kausal for $\{Y_t\}_{t \in \mathbb{Z}}$, skal vi blot aflæse, om det kan antages, at $\Phi_{12,1} = 0, \dots, \Phi_{12,p} = 0$.

SAS kan også lave et *Wald-test* for kausalitet (Se [SAS 9.1 User's Guide](#) side 1770 og [Lütkepohl \(2007\)](#) side 102 og fremefter). Hypotesen om Granger-kausalitet kan formuleres som en hypotese om koefficientmatricerne, der skal antage en bestemt form. Med notationen fra formuleringen af modellen som en multipel regressionsmodel på [side 133](#) kan vi opstille hypotesen som

$$H_0: C\beta = c,$$

hvor C er en $(s \times (K^2 p + K))$ matrix af rang s , og c er en s -dimensional vektor. Under antagelsen, af at resultatet i (5.9) på [side 134](#) gælder, fås Wald-teststørrelsen

$$N(C\hat{\beta} - c)^* [C(\hat{\Gamma}_p^{-1} \otimes \hat{\Sigma}_u)C^*]^{-1} (C\hat{\beta} - c), \quad (5.16)$$

som er asymptotisk χ^2 -fordelt med s frihedsgrader. Når der testes for Granger-kausalitet består matricen C af ettaller og nuller, og vektoren c består af lutter nuller.

Testet i (5.16) har vi faktisk allerede brugt i [SAS-udskrift 5.14](#) på [side 141](#) og i [SAS-udskrift 5.17](#) på foregående side til at vise at $\Phi_1 = 0$ i VAR(1)-modellen for residualerne.

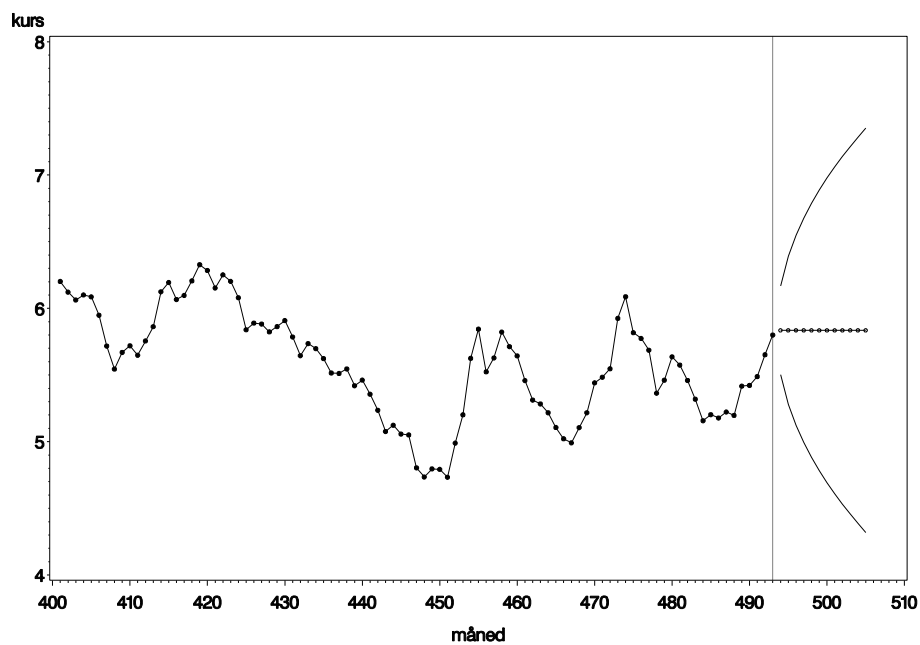
Eksempel 5.8. I [Eksempel 3.19](#) på [side 78](#) viste vi, at en VMA(1)-model kunne beskrive førstedifferenserne for de 493 observationer af kursen på US\$. Denne model giver anledning til det forecast af kursen i de næste 12 måneder, som er vist i [Figur 5.3](#). Forecastet er ikke særlig informativt, idet det er konstant lig med værdien i den næste måned. Forklaringen på dette er, at da den autoregressive orden $p = 0$, giver (4.69) på [side 111](#), at

$$\widehat{\nabla x}_{493}(h) = 0, \quad h > 1,$$

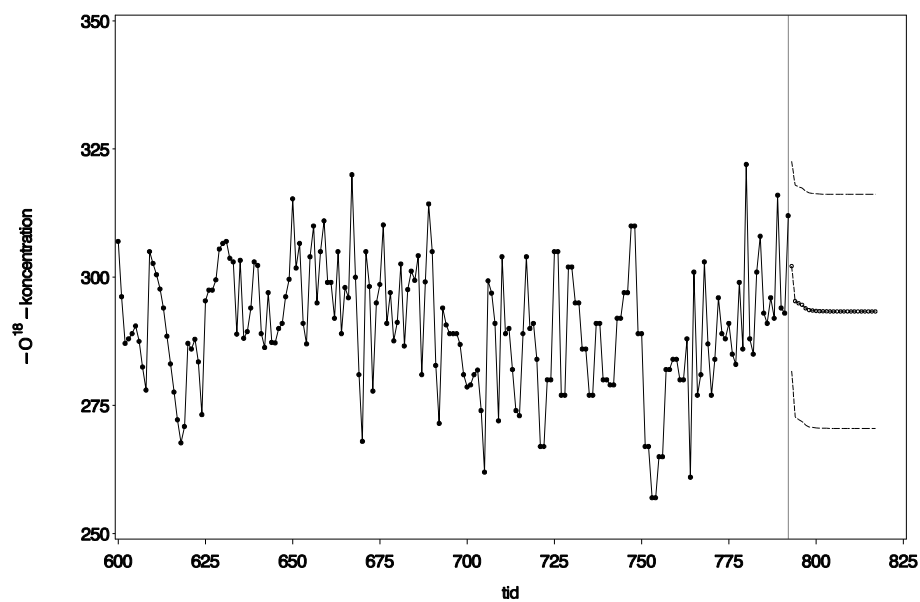
eller

$$\hat{x}_{493}(h) = \hat{x}_{493}(1), \quad h > 1.$$

Eksempel 5.9. I [Figur 5.4](#) ses forecastet af de næste 25 værdier af isotopkurven beregnet i AR(3)-modellen i (3.37) på [side 72](#). Da forecastene opfylder differensligningen $\varphi(B)(\hat{x}_{792}(h) - \mu) = 0$ for $h > 0$, se formel (4.69), nærmer forecastene sig hurtigt processens middelværdi. Modellen er således ikke særlig informativ med hensyn til fremtidige udvikling af isotopkurven, og dermed af den fremtidige udvikling af klimaet.



Figur 5.3: Forecast af og 95 %-konfidensintervaller for kursen på US\$ i de næste 12 måneder beregnet fra januar 2012.



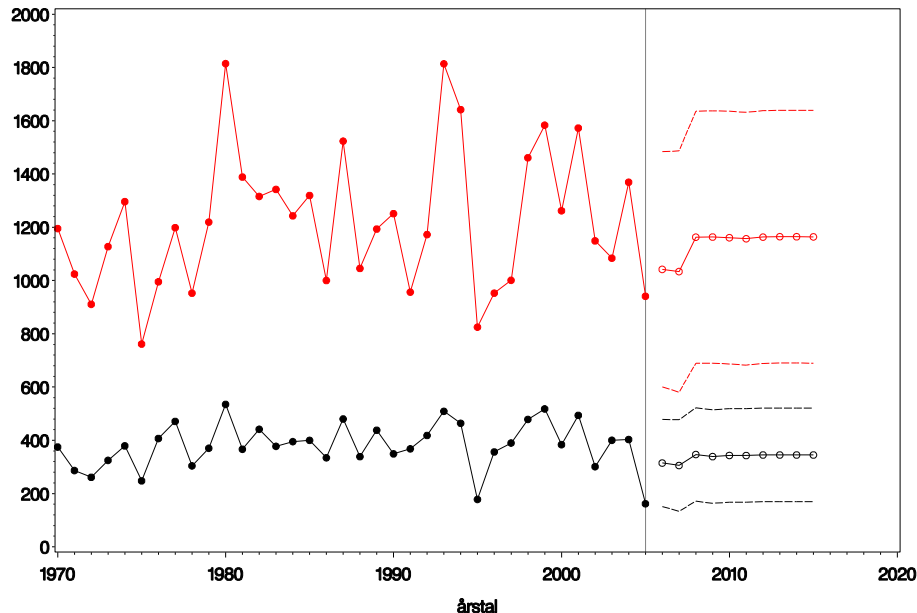
Figur 5.4: Forecast af og 95 %-konfidensintervaller for de næste 25 værdier af isotopkurven beregnet i AR(3)-modellen i (3.37) på side 72.

Eksempel 5.10. Da $\hat{\Phi}_{12,1}$ og $\hat{\Phi}_{21,2}$ i (5.13) er signifikant forskellige fra 0, fås det ved hjælp af [Sætning 4.13](#) på [side 113](#) og [Sætning 4.14](#) på [side 115](#), at nedbøren er Granger-kausal for vandmængden og omvendt. Dette fremgår også af [SAS-udskrift 5.18](#), idet SAS i Granger-Causality Wald Test tester hypotesen, at de variable i GROUP 1 kun er påvirket af sig selv og ingen af de variable i GROUP 2. Hvis hypotesen forkastes, er de variable i GROUP 2 derfor Granger-kausale for de variable i GROUP 1.

SAS-udskrift 5.18

Granger-Causality Wald Test			
Test	DF	Chi-Square	Pr > ChiSq
1	2	8.89	0.0117
2	2	10.60	0.0050
Test 1:	Group 1 Variables:	ned	
	Group 2 Variables:	vand	
Test 2:	Group 1 Variables:	vand	
	Group 2 Variables:	ned	

[Figur 5.5](#) viser forecastet af de næste 10 års værdier af nedbør og vandmængde foretaget i 2005.



Figur 5.5: Forecast af og 95 %-konfidensintervallet for de næste 10 års værdier af nedbør (sort) og vandmængde (rød) foretaget i 2005.

Eksempel 5.11. Som nævnt på [side 137](#) antyder estimerne i VAR(2)-modellen i (5.14) for $\mathbf{v} = (v_1, v_2, v_3)^*$, vektoren af førstedifferenser af logaritmerne for henholdsvis faste investeringer, disponibel indkomst og udgifter til forbrug, at hypotesen

$$H_0 : \Phi_{12,1} = \Phi_{13,1} = \Phi_{12,2} = \Phi_{13,2} = 0 \quad (5.17)$$

ikke forkastes. Ifølge [Sætning 4.13](#) på [side 113](#) og [Sætning 4.14](#) på [side 115](#) betyder dette, at (v_2, v_3) ikke er Granger kausal for v_1 . I [SAS-udskrift 5.19](#) ses det, at hypotesen H_0 i (5.17) ikke forkastes.

SAS-udskrift 5.19

Granger-Causality Wald Test			
Test	DF	Chi-Square	Pr > ChiSq
1	4	6.37	0.1734
Test 1: Group 1 Variables: v1			
Group 2 Variables: v2 v3			

Estimerne i modellen svarende til H_0 ses i

SAS-udskrift 5.20

The VARMAX Procedure						
Type of Model			VAR(2)			
Estimation Method		Least Squares Estimation				
Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > t	Variable
v1	CONST1	0.02353	0.00631	3.73	0.0004	1
	AR1_1_1	-0.21122	0.11726	-1.80	0.0762	v1(t-1)
	AR1_1_2	0.00000	0.00000			v2(t-1)
	AR1_1_3	0.00000	0.00000			v3(t-1)
	AR2_1_1	-0.08230	0.11708	-0.70	0.4846	v1(t-2)
	AR2_1_2	0.00000	0.00000			v2(t-2)
	AR2_1_3	0.00000	0.00000			v3(t-2)
v2	CONST2	0.01712	0.00434	3.94	0.0002	1
	AR1_2_1	0.04758	0.03182	1.49	0.1397	v1(t-1)
	AR1_2_2	-0.15764	0.13735	-1.15	0.2552	v2(t-1)
	AR1_2_3	0.25618	0.16721	1.53	0.1303	v3(t-1)
	AR2_2_1	0.05266	0.03169	1.66	0.1013	v1(t-2)
	AR2_2_2	0.01531	0.13456	0.11	0.9097	v2(t-2)
	AR2_2_3	-0.04163	0.16741	-0.25	0.8044	v3(t-2)
v3	CONST3	0.01526	0.00340	4.49	0.0001	1
	AR1_3_1	0.00385	0.02555	0.15	0.8806	v1(t-1)
	AR1_3_2	0.21636	0.10712	2.02	0.0475	v2(t-1)
	AR1_3_3	-0.31959	0.13041	-2.45	0.0169	v3(t-1)
	AR2_3_1	0.03841	0.02544	1.51	0.1359	v1(t-2)

AR2_3_2	0.34828	0.10494	3.32	0.0015	v2(t-2)
AR2_3_3	-0.07630	0.13057	-0.58	0.5609	v3(t-2)
Covariances of Innovations					
Variable	v1	v2	v3		
v1	0.00234	0.00008	0.00014		
v2	0.00008	0.00014	0.00006		
v3	0.00014	0.00006	0.00009		

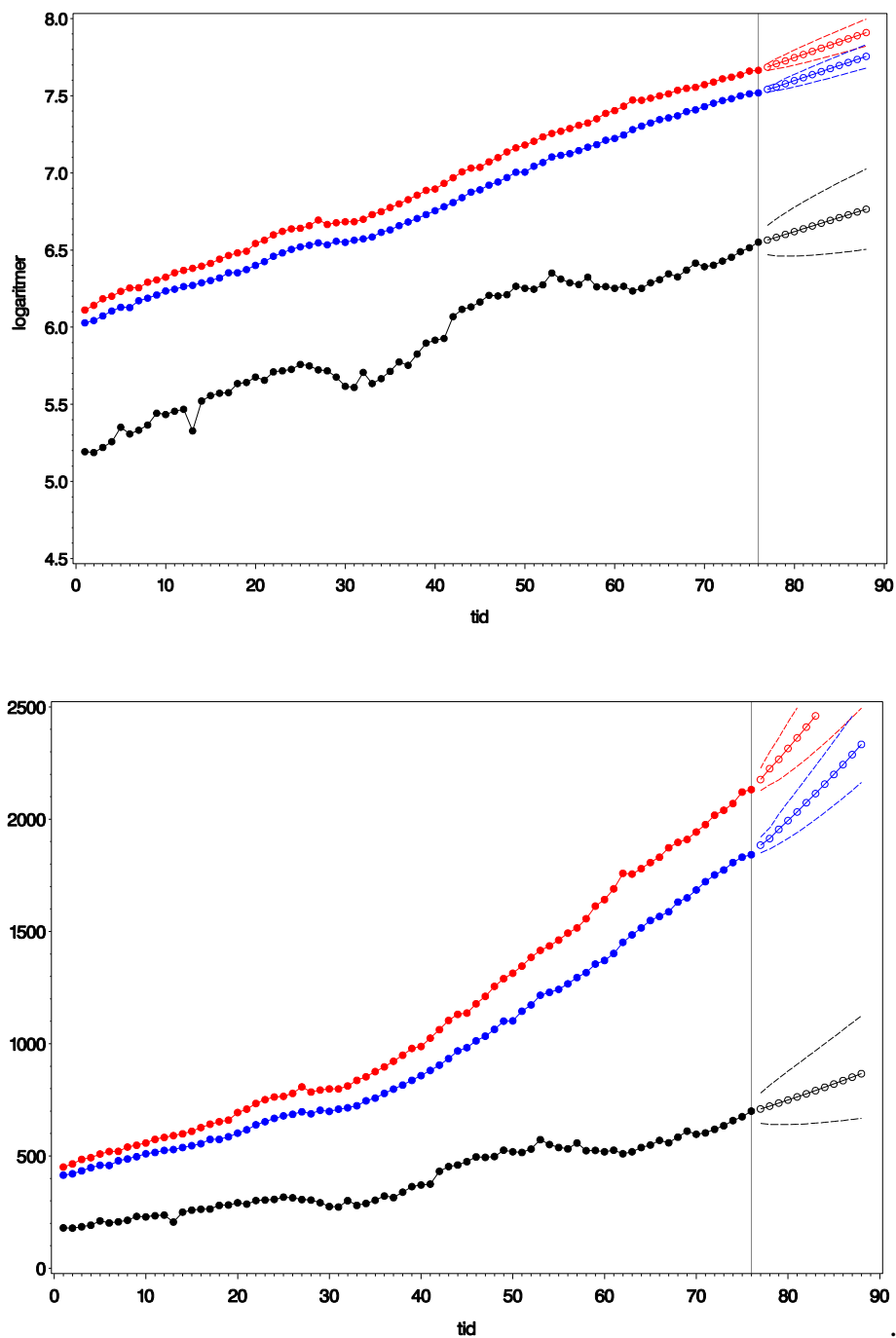
SAS-udskrift 5.10 viser, at den estimerede VAR(2)-model er

$$\begin{aligned}
 \begin{pmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{pmatrix} &= \begin{pmatrix} 0.02353 \\ 0.01712 \\ 0.01526 \end{pmatrix} + \begin{pmatrix} -0.21122 & 0.00000 & 0.00000 \\ 0.04758 & -0.15764 & 0.25618 \\ 0.00385 & 0.21636 & -0.31959 \end{pmatrix} \begin{pmatrix} v_{1,t-1} \\ v_{2,t-1} \\ v_{3,t-1} \end{pmatrix} \\
 &+ \begin{pmatrix} -0.08230 & 0.00000 & 0.00000 \\ 0.05266 & 0.01531 & -0.04163 \\ 0.03481 & 0.34828 & -0.07630 \end{pmatrix} \begin{pmatrix} v_{1,t-2} \\ v_{2,t-2} \\ v_{3,t-2} \end{pmatrix} + \begin{pmatrix} w_{1t} \\ w_{2t} \\ w_{3t} \end{pmatrix}, \\
 &\hspace{20em} (5.18)
 \end{aligned}$$

hvor

$$\Sigma_w \leftarrow \begin{pmatrix} 0.00234 & 0.00008 & 0.00014 \\ 0.00008 & 0.00014 & 0.00006 \\ 0.00014 & 0.00006 & 0.00009 \end{pmatrix}.$$

Forecastet af tidsrækkens næste 12 værdier ses i [Figur 5.6](#).



Figur 5.6: Nederst forecast af og 95 %-konfidensintervaller for de næste 12 kvartalsvise værdier af faste investeringer (sort), disponibel indkomst (rød) og udgifter til forbrug (blå). Øverst logaritmen af de tilsvarende størrelser.

Appendikser

Appendiks A

Nyttige resultater

Følgende teori og resultater er taget direkte fra [Lütkepohl \(2007, se Appendix A og Appendix C\)](#).

Lad A være en $(m \times m)$ matricen

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mm} \end{pmatrix}$$

med egenverdierne $\lambda_1, \lambda_2, \dots, \lambda_n$. Der eksisterer da en ikke-singulær matrix P således, at

$$P^{-1}AP = \begin{bmatrix} \Lambda_1 & & 0 \\ & \ddots & \\ 0 & & \Lambda_n \end{bmatrix} \quad \text{eller} \quad A = P\Lambda P^{-1},$$

hvor vi definerer $\Lambda = P^{-1}AP$, og

$$\Lambda_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & & \ddots & 1 \\ 0 & 0 & \cdots & \cdots & \lambda_i \end{pmatrix}.$$

Denne dekomponering af A kaldes for *Jordan kanoniske form*. En egenverdi λ_i af geometrisk multiplicitet k , optræder k gange i listen af egenverdier, $\lambda_1, \dots, \lambda_n$.

Matricen Λ_i bestemmes ud fra den geometriske og den algebraiske multiplicitet af egenverdien¹ λ_i . Den geometriske multiplicitet af λ_i er dimensionen af egenrummet hørende til egenverdien λ_i .

Antag at alle egenverdierne $\lambda_1, \dots, \lambda_n$ har modulus mindre end 1. Dvs. $|\lambda_i| < 1$ for $i = 1, 2, \dots, n$. Da har vi følgende resultater:

¹For nærmere forklaring, se http://en.wikipedia.org/wiki/Jordan_normal_form

(A.1) $A^j = P\Lambda^j P^{-1} \rightarrow 0$ for $j \rightarrow \infty$.

(A.2) $\sum_{j=0}^{\infty} A^j = (I_m - A)^{-1}$ eksisterer.

(A.3) Følgen A^j , $j = 0, 1, 2, \dots$, er absolut summabel. Dvs.

$$\sum_{j=0}^{\infty} |\alpha_{kl,j}|$$

er endelig for alle $k, l = 1, \dots, m$, hvor $\alpha_{kl,j}$ er det kl 'te element i A^j .

(A.4) Alle egenverdierne for A har modulus mindre end 1, hvis og kun hvis $\det(I_m - Az) \neq 0$ for $|z| \leq 1$. Dvs. polynomiet $\det(I_m - Az)$ har ingen rødder inden for enhedscirklen.

Lad nu A være en vilkårlig $(m \times m)$ matrix, som er delt op i fire delmatricer A_{11} , A_{12} , A_{21} og A_{22} , hvis dimensioner er henholdsvis $(p \times q)$, $(p \times (m - q))$, $((m - p) \times q)$ og $((m - p) \times (m - q))$. Vi skriver da

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

Da gælder følgende:

(A.5)

$$AA^{\text{adj}} = A^{\text{adj}}A = |A|I_m,$$

hvorfor $A^{-1} = |A|^{-1} A^{\text{adj}}$. Her er A^{adj} den adjungerede af A .

(A.6) Hvis A_{22} er kvadratisk og ikke-singulær, så gælder der, at

$$\det(A) = \det(A_{22}) \det(A_{11} - A_{12}A_{22}^{-1}A_{21}).$$

(A.7)

$$A^* = \begin{bmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{bmatrix}$$

(A.8) Hvis $p = q$, og A , A_{11} og A_{22} er ikke-singulære, gælder der, at

$$\begin{aligned} A^{-1} &= \begin{bmatrix} D & -DA_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}D & A_{22}^{-1} + A_{22}^{-1}A_{21}DA_{12}A_{22}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}GA_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}G \\ -GA_{21}A_{11}^{-1} & G \end{bmatrix}, \end{aligned}$$

hvor $D = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}$, og $G = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}$.

Sætning A.1. Antag, at $\{A_i\}$ er en absolut summabel følge af reelle $(K \times K)$ matricer, og $\{z_t\}$ er en følge af K -dimensionale stokastiske vektorer, der opfylder, at

$$E(z_t^* z_t) \leq c$$

for $t = 0, \pm 1, \pm 2, \dots$. Her er c en endelig konstant. Da eksisterer der en følge af K -dimensionale stokastiske variable $\{y_t\}$ således, at

$$\sum_{i=-n}^n A_i z_{t-i} \xrightarrow{n \rightarrow \infty} y_t$$

hvor konvergens sker i kvadratisk middelværdi. Følgen er entydigt bestemt på nær på en nulmængde.

Bemærkning A.2. Sætningen sikrer, at den uendelige VMA-repræsentation af en VAR-proces er veldefineret. Det kan nemlig vises, at koefficientmatricerne Ψ_i fra VMA-repræsentationen for en stabil VAR-proces udgør en absolut summabel følge. ■

Sætning A.3. Antag, at $\{z_t\}$ opfylder betingelserne i [Sætning A.1](#), $\{A_i\}$ og $\{B_i\}$ er absolut summable følger af $(K \times K)$ matricer,

$$y_t = \sum_{i=-\infty}^{\infty} A_i z_{t-i} \quad \text{og} \quad x_t = \sum_{i=-\infty}^{\infty} B_i z_{t-i}.$$

Da gælder der, at

$$E(y_t) = \lim_{n \rightarrow \infty} \sum_{i=-n}^n A_i E(z_{t-i}),$$

og

$$E(y_t x_t^*) = \lim_{n \rightarrow \infty} \sum_{i=-n}^n \sum_{j=-n}^n A_i E(z_{t-i} z_{t-j}^*) B_j^*,$$

hvor grænseværdien af følgen af matricer er lig matricen bestående af grænseværdierne for de enkelte elementers følger.

Appendiks B

Differensligninger

I [Sætning 2.21](#) blev det vist, at autokovariansfunktionen $\{\gamma_k\}$, autokorrelationsfunktionen $\{\rho_k\}$, koefficienterne $\{\psi_k\}$ i rækkeudviklingen af $\theta(z)/\varphi(z)$ og koefficienterne $\{\pi_k\}$ i rækkeudviklingen af $\varphi(z)/\theta(z)$ alle opfylder en differensligning. I dette appendiks præsenteres og illustreres den følgende sætning vedrørende den komplette løsning til en homogen differensligning.

Sætning B.1. *Lad*

$$\varphi(z) = \varphi_0 + \varphi_1 z + \cdots + \varphi_p z^p, \quad z \in \mathbb{C},$$

være et polynomium af grad p og med komplekse koefficienter $\varphi_0, \varphi_1, \dots, \varphi_p$ ($\varphi_p \neq 0$) med alle rødder forskellige fra 0. Lad $G_1^{-1}, \dots, G_r^{-1}$ være de forskellige rødder for $\varphi(z)$ således at G_i^{-1} har multiplicitet p_i , $i = 1, \dots, r$. Da er den fuldstændige løsning til differensligningen

$$\varphi(B)\gamma_k = \varphi_0 \gamma_k + \varphi_1 \gamma_{k-1} + \cdots + \varphi_p \gamma_{k-p} = 0, \quad k > q \quad (\text{B.1})$$

givet ved, at

$$\gamma_k = \begin{cases} \sum_{i=1}^r (A_{i0} + A_{i1}k + \cdots + A_{ip_i-1}k^{p_i-1})G_i^k & \text{hvis } k > q-p \\ \text{vilkaarlig} & \text{hvis } k \leq q-p. \end{cases} \quad (\text{B.2})$$

Her er A_{ij} , $j = 1, \dots, p_i - 1$, $i = 1, \dots, r$, komplekse konstanter.

Hvis $\varphi(z)$ har reelle koefficienter $\varphi_0, \varphi_1, \dots, \varphi_p$ ($\varphi_p \neq 0$) og hvis alle rødder for $\varphi(z)$ er forskellige og ligger uden for enhedscirklen $\{z \in \mathbb{C} : |z| = 1\}$, vil alle reelle løsninger $\{\gamma_k : k \in \mathbb{Z}\}$ for $k > q - p$ være en sum af eksponentielt aftagende led og eksponentielt dæmpede sinussvingninger.

Bevis. Med notationen i sætningen har vi, at

$$\varphi(z) = c \prod_{i=1}^r (1 - G_i z)^{p_i}, \quad z \in \mathbb{C},$$

Da

$$\begin{aligned}(1 - G_i B)(k^j G_i^k) &= k^j G_i^k - G_i(k-1)^j G_i^{k-1} \\ &= (k^j - (k-1)^j) G_i^k\end{aligned}$$

hvis $j \in \mathbb{N}_0$ og $k \in \mathbb{Z}$, og da $k^j - (k-1)^j$ er et polynomium i k af grad $j-1$, har vi for $i = 1, \dots, r$, at

$$\begin{aligned}(1 - G_i B)^{p_i} (A_{i0} + A_{i1}k + \dots + A_{ip_i-1}k^{p_i-1}) G_i^k \\ &= (1 - G_i B)^{p_i-1} (1 - G_i B) (A_{i0} + A_{i1}k + \dots + A_{ip_i-1}k^{p_i-1}) G_i^k \\ &= (1 - G_i B)^{p_i-1} (\tilde{A}_{i0} + \tilde{A}_{i1}k + \dots + \tilde{A}_{ip_i-2}k^{p_i-2}) G_i^k \\ &\vdots \\ &= (1 - G_i B) \check{A}_{i0} G_i^k = 0,\end{aligned}$$

hvor størrelserne \tilde{A}_{ij} og \check{A}_{i0} kan udtrykkes ved hjælp af konstanterne A_{ij} . Heraf ses det, at enhver følge $\{\gamma_k : k \in \mathbb{Z}\}$, der opfylder (B.2), er en løsning til (B.1), da vi for alle $k > q$ har, at

$$\begin{aligned}\varphi(B)\gamma_k &= c \sum_{i=1}^r \prod_{j=1}^r (1 - G_j B)^{p_j} (A_{i0} + A_{i1}k + \dots + A_{ip_i-1}k^{p_i-1}) G_i^k \\ &= 0.\end{aligned}$$

Vi viser nu, at en løsning til (B.1) er af formen (B.2), idet vi dog kun betragter tilfældet $p = 2$. Uden tab af generalitet antager vi, at $q = 0$.

Antag, at

$$(1 - G_1 B)(1 - G_2 B)\gamma_k = 0, \quad k > 0.$$

Sættes

$$y_k = (1 - G_2 B)\gamma_k, \quad k \geq 0, \tag{B.3}$$

fås

$$(1 - G_1 B)y_k = 0, \quad k > 0,$$

eller

$$y_k = G_1^k y_0, \quad k \geq 0.$$

Indsættes dette i (B.3) finder vi, at

$$\gamma_k = G_2 \gamma_{k-1} + G_1^k y_0, \quad k \geq 0.$$

Itereres denne ligning får vi

$$\gamma_k = G_2^{k+1} \gamma_{-1} + y_0 (G_1^k + G_1^{k-1} G_2 + \dots + G_2^k), \quad k \geq -1, \tag{B.4}$$

hvor faktoren til y_0 indeholder $k + 1$ led. Hvis $G_1 \neq G_2$, bliver dette udtryk

$$\begin{aligned}\gamma_k &= G_2^{k+1} \gamma_{-1} + y_0 \frac{G_1^{k+1} - G_2^{k+1}}{G_1 - G_2} \\ &= A_1 G_1^k + A_2 G_2^k, \quad k \geq -1.\end{aligned}$$

Hvis $G_1 = G_2$ fås af (B.4), at

$$\begin{aligned}\gamma_k &= G_2^{k+1} \gamma_{-1} + y_0(k+1)G_2^k \\ &= (A_0 + A_1 k)G_2^k, \quad k \geq -1.\end{aligned}$$

Vi mangler nu kun at vise den sidste påstand i sætningen, og igen betragter vi kun tilfældet $p = 2$. Hvis $\varphi(z)$ har to forskellige reelle rødder medfører (B.2), at en reel løsning til $\{\gamma_k : k \in \mathbb{Z}\}$ til (B.1) for $k \geq q - 2$ er en sum af to eksponentielt aftagende led med reelle koefficienter. Hvis $\varphi(z)$ har to forskellige komplekse rødder G_1^{-1} og G_2^{-1} gælder der, at $G_1 = \bar{G}_2$. For en reel løsning $\{\gamma_k : k \in \mathbb{Z}\}$ til (B.1) gælder, at

$$\gamma_k = A_1 G_1^k + A_2 G_2^k, \quad k > q - 2, \quad (\text{B.5})$$

og derfor

$$\begin{aligned}\gamma_k &= \bar{\gamma}_k = \overline{A_1 G_1^k + A_2 G_2^k} \\ &= \bar{A}_1 G_2^k + \bar{A}_2 G_1^k, \quad k > q - 2,\end{aligned}$$

så

$$0 = (A_1 - \bar{A}_2)G_1^k + (A_2 - \bar{A}_1)G_2^k, \quad k > q - 2.$$

Sættes $G_1 = |G_1|e^{i\chi}$, $\chi \in]-\pi, \pi[$, $\chi \neq 0$, har vi, at $G_2 = |G_1|e^{-i\chi}$ og

$$0 = (A_1 - \bar{A}_2)e^{i\chi k} + (A_2 - \bar{A}_1)e^{-i\chi k}, \quad k > q - 2,$$

eller

$$0 = (A_1 - \bar{A}_2)e^{i2\chi k} + (A_2 - \bar{A}_1), \quad k > q - 2.$$

Heraf ses, at $A_1 = \bar{A}_2$, og sættes $A_1 = |A_1|e^{i\omega}$ fås af (B.5), at

$$\begin{aligned}\gamma_k &= |A_1||G_1|^k (e^{i(\omega+\chi k)} + e^{-i(\omega+\chi k)}) \\ &= 2|A_1||G_1|^k \cos(\omega + \chi k) \\ &= 2|A_1||G_1|^k \sin(\omega + \chi k + \pi/2), \quad k > q - 2.\end{aligned}$$

Hermed er sætningen vist. □

Vi afslutter dette appendiks med et eksempel, hvor autokorrelationsfunktionen for en AR(2)-proces studeres.

Eksempel B.2. Autokorrelationsfunktionen $\{\rho_k\}_{k \in \mathbb{Z}}$ for en AR(2)-proces opfylder differensligningen

$$(1 - \varphi_1 B - \varphi_2 B^2)\rho_k = 0, \quad k > 0, \quad (\text{B.6})$$

where $\varphi_2 \neq 0$. I dette eksempel giver vi et eksplicit udtryk for ρ_k ved hjælp af processens parametre (φ_1, φ_2) .

Løsningen $\{\rho_k\}$ er bestemt ved rødderne i ligningen

$$1 - \varphi_1 z - \varphi_2 z^2 = 0, \quad z \in \mathbb{C}. \quad (\text{B.7})$$

Ifølge (B.2) er ρ_k af formen

$$\rho_k = (A_0 + A_1 k)G^k, \quad k > -2, \quad (\text{B.8})$$

hvis begge rødder i (B.7) er lig med G^{-1} , og af formen

$$\rho_k = A_1 G_1^k + A_2 G_2^k, \quad k > -2, \quad (\text{B.9})$$

hvis rødderne i (B.7) er forskellige og lig med henholdsvis G_1^{-1} og G_2^{-1} . Vi mangler altså blot at bestemme konstanterne i (B.8) and (B.9).

Først bemærker vi, at da $\rho_0 = 1$ og $\rho_{-1} = \rho_1$, får vi fra (B.6) at

$$\rho_1 = \frac{\varphi_1}{1 - \varphi_2}.$$

Lad

$$D = \varphi_1^2 + 4\varphi_2.$$

Hvis $D = 0$, eller ækvivalent hermed, hvis $\varphi_2 = -\varphi_1^2/4$, så er den fælles rod i (B.7) givet ved

$$G^{-1} = 2\varphi_1^{-1}.$$

Fra (B.8) finder vi – ved at sætte $k = 0$ – at

$$A_0 = 1,$$

og sættes $k = 1$ fås

$$\frac{\varphi_1}{1 - \varphi_2} = \rho_1 = (1 + A_1)G = (1 + A_1)\frac{\varphi_1}{2},$$

hvilket medfører, at

$$A_1 = \frac{1 + \varphi_2}{1 - \varphi_2} \left(= \frac{4 - \varphi_1^2}{4 + \varphi_1^2} \right).$$

Hvis $D \neq 0$, har ligningen (B.7) de to rødder

$$G_1^{-1} = (\sqrt{D} - \varphi_1)/(2\varphi_2) \quad \text{og} \quad G_2^{-1} = -(\sqrt{D} + \varphi_1)/(2\varphi_2), \quad (\text{B.10})$$

hvoraf det fås, at

$$G_1 = (\sqrt{D} + \varphi_1)/2 \quad \text{and} \quad G_2 = -(\sqrt{D} - \varphi_1)/2. \quad (\text{B.11})$$

Fra (B.9) med henholdsvis $k = 0$ og $k = 1$ finder vi

$$A_1 + A_2 = 1 \quad \Leftrightarrow \quad A_2 = 1 - A_1$$

og

$$\begin{aligned} \rho_1 &= \frac{\varphi_1}{1 - \varphi_2} = A_1 G_1 + (1 - A_1) G_2 \\ &= A_1 (G_1 - G_2) + G_2 \\ &= A_1 \sqrt{D} - (\sqrt{D} - \varphi_1)/2 \end{aligned}$$

hvilket medfører, at

$$A_1 = \frac{1}{2} + \frac{\varphi_1(1 + \varphi_2)}{2\sqrt{D}(1 - \varphi_2)}$$

og

$$A_2 = \frac{1}{2} - \frac{\varphi_1(1 + \varphi_2)}{2\sqrt{D}(1 - \varphi_2)}. \quad (\text{B.12})$$

Opgaver

Opgave 1. Vis, at en hvid støj har spektraltætheden i (2.11).

Opgave 2. Vis, at det glidende gennemsnit

$$X_t = \sum_{k=-\infty}^{\infty} \psi_k W_{t-k}, \quad t \in \mathbb{Z},$$

er en svagt stationær proces.

Opgave 3. Vis for at for $MA(q)$ -processen

$$X_t = W_t - \theta_1 W_{t-1} - \cdots - \theta_q W_{t-q}$$

gælder der, at $\gamma_k = 0$ for $k > q$.

Opgave 4. Vis Sætning 2.15 på side 13.

Opgave 5. Vis, at mængden

$$\Phi_2 = \{\varphi = (\varphi_1, \varphi_2) \in \mathbb{R}^2 : \varphi(z) = 1 - \varphi_1 z - \varphi_2 z^2 \neq 0 \text{ hvis } |z| \leq 1\},$$

er det indre af trekanten T bestemt af punkterne $(\varphi_1, \varphi_2) = (0, 1)$, $(\varphi_1, \varphi_2) = (2, -1)$ og $(\varphi_1, \varphi_2) = (-2, -1)$.

Opgave 6. Vis formlerne (2.34)–(2.40).

Opgave 7. Vis formlen (2.42).

Opgave 8. Vis formlen (2.44).

Opgave 9. Vis omskrivningen i (3.11).

Opgave 10. Vis formel (4.3).

Opgave 11. Vis formel (4.19).

Opgave 12. Vis formel (4.47).

Opgave 13. Vis formlerne (4.52) og (4.54).

Opgave 14. Opgaven går ud på at vise, at en K -dimensional VARMA(p, q)-proces kan udtrykkes som en $K(p + q)$ -dimensional VAR(1)-proces.

Betragt den K -dimensionale VARMA(p, q)-proces i (4.39), og antag uden tab af generalitet, at den har middelværdi $\mu = 0$.

Lad \mathbf{Z}_t og \mathbf{W}_t være $K(p + q)$ -dimensionale vektorer givet ved

$$\mathbf{Z}_t = \begin{bmatrix} \mathbf{X}_t \\ \mathbf{X}_{t-1} \\ \vdots \\ \mathbf{X}_{t-p+1} \\ \mathbf{w}_t \\ \mathbf{w}_{t-1} \\ \vdots \\ \mathbf{w}_{t-q+1} \end{bmatrix} \quad \text{og} \quad \mathbf{W}_t = \begin{bmatrix} \mathbf{w}_t \\ 0 \\ \vdots \\ 0 \\ \mathbf{w}_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Betragt desuden $K(p + q) \times K(p + q)$ -matricen

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix},$$

hvor $\Phi_{11}(Kp \times Kp)$, $\Phi_{12}(Kp \times Kq)$, $\Phi_{21}(Kq \times Kp)$ og $\Phi_{22}(Kq \times Kq)$ er givet ved

$$\Phi_{11} = \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{p-1} & \Phi_p \\ I_K & 0 & \cdots & 0 & 0 \\ 0 & I_K & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_K & 0 \end{bmatrix},$$

$$\Phi_{12} = \begin{bmatrix} -\Theta_1 & -\Theta_2 & \cdots & -\Theta_{q-1} & -\Theta_q \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

$$\Phi_{21} = 0,$$

og

$$\Phi_{22} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ I_K & 0 & \cdots & 0 & 0 \\ 0 & I_K & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_K & 0 \end{bmatrix}.$$

- (1) Vis, at $\{\mathbf{W}_t\}_{t \in \mathbb{Z}}$ er en hvid støj.
- (2) Vis, at den K -dimensionale VARMA(p, q)-proces kan formuleres som den $K(p + q)$ -dimensionale VAR(1)-proces

$$\mathbf{Z}_t = \mathbf{\Phi} \mathbf{Z}_{t-1} + \mathbf{w}_t. \quad (\text{B.13})$$

VAR(1)-processen (B.13) er som bekendt stabil, hvis

$$\det(I_{K(p+q)} - \mathbf{\Phi}z) \neq 0 \quad \text{for } |z| \leq 1.$$

- (3) Vis ved hjælp af Regneregler (A.6) i Appendiks A, at

$$\begin{aligned} \det(I_{K(p+q)} - \mathbf{\Phi}z) &= \det \begin{bmatrix} I_{Kp} - \mathbf{\Phi}_{11}z & -\mathbf{\Phi}_{12}z \\ 0 & I_{Kq} - \mathbf{\Phi}_{22}z \end{bmatrix} \\ &= \det(I_{Kp} - \mathbf{\Phi}_{11}z) \det(I_{Kq} - \mathbf{\Phi}_{22}z) \\ &= \det(I_{Kp} - \Phi_1 z - \dots - \Phi_p z^p). \end{aligned}$$

Der gælder altså, at VAR(1)-processen i (B.13) er stabil, hvis og kun hvis VARMA(p, q)-processen (4.39) er stabil.

Opgave 15. Opgaven vedrører opdatering af forecasts i en ARMA(p, q)-model. Antag, at vi har observeret processen op til tid t og har beregnet alle forecasts $\hat{x}_t(h)$, $h > 0$. Vi observerer nu den næste værdi x_{t+1} af processen \mathbf{X} og ønsker at finde forecasts $\hat{x}_{t+1}(h)$ baseret på observationerne $\mathbf{x}_{\leq t+1}$ op til og med tid $t + 1$.

- (1) Vis ved hjælp af (4.71), at

$$\begin{aligned} \hat{x}_{t+1}(h) &= \sum_{k=l}^{\infty} \psi_k w_{t+1+h-k} \\ &= \psi_h w_{t+1} + \hat{x}_t(h+1). \end{aligned}$$

specielt for $h = 0$, at

$$x_{t+1} = \hat{x}_{t+1}(0) = \psi_0 w_{t+1} + \hat{x}_t(1) = w_{t+1} + \hat{x}_t(1).$$

- (2) Vis opdateringsformlen

$$\hat{x}_{t+1}(l) = \hat{x}_t(l+1) + \psi_l(x_{t+1} - \hat{x}_t(1)), \quad (\text{B.14})$$

Opgave 16. Vis Sætning 4.14.

Indeks

A

Andersen og Blæsild (2003), 20, 67, 84

absolut summable følger, 154

konvergens af kovarianser, 155

konvergens af middelværdier, 155

konvergens i kvadratisk middel, 155

ac, 11

AIC, 47

Akaike information criterion (AIC), 47

anden ordens proces, 9

AR(1)-processen

asymptotisk varians, 68

egenskaber ved, 26

simulerede udfaldsfunktioner, 27

teoretiske korrelationsfunktioner, 27

AR(2)-processen

asymptotisk kovarians, 68

egenskaber ved, 30

simulerede udfaldsfunktioner, 32, 33

teoretiske korrelationsfunktioner, 32, 33

ARIMA-proces, 43

ARMA-modeller

estimation, 61

identifikation, 44

log likelihood funktion, 62

modelkontrol, 75

profil log likelihood funktion, 62

ARMA(1,1)-processen

asymptotisk kovarians, 68

egenskaber ved, 37

simulerede udfaldsfunktioner, 38, 39

teoretiske korrelationsfunktioner, 38, 39

ARMA(p, q)-proces

AR-repræsentation, 20

autokorrelationsfunktion, 22

autokorrelationsfunktion, differens-
ligning, 22

autokovariansfunktion, 21

autokovariansfunktion, differens-
ligning, 22

bestemmelse af orden, 47

koefficienter i AR-repræsentation,
differensligning, 22

koefficienter i MA-repræsentation,
differensligning, 22

MA-repræsentation, 19

spektraltæthed, 23

\mathcal{A}

mængden af ARMA-processer, 21

\mathcal{A}_0

mængden af ARMA-processer med
AR- og MA-repræsentation, 21

AR(p)-proces, 16

ARMA(0,0)-proces, 20

Yule-Walker ligninger, 23

asymptotisk teori, 66

autokorrelationsfunktion, 11, 88
 empirisk, 42

autokovariansfunktion, 11, 87

autoregression, 16
 af orden p , 16

B

B operator, 18

Bartlett (1955), 42

Bayesian information criterion (BIC),
 48

bestemmelse af orden, 47

betinget middelværdi
 minimum mean squared predik-
 tor, 107

BIC, 48

Box og Jenkins (1970), 14, 41, 68, 77
 β^Z , 7

C

Cholesky dekomposition, 62, 119

D

Δ_Z , 7

Dickey-Fuller test, 45

differensligning

 ARMA(p, q)-proces
 autokorrelationsfunktion, 22
 autokovariansfunktion, 22
 koefficienter i AR-repræsentation,
 22
 koefficienter i MA-repræsentation,
 22

 eksempel, 159

 forecast, 111

 fuldstændig løsning, 156

 homogen, 156

E

Eksempel 1.1

 valutakursen på US\$, 1

 estimation, 69

 identifikation, 48

 modelkontrol, 78

 forecasting, 146

Eksempel 1.2

 Camp Century isotopkurven, 2

 estimation, 70, 72

 identifikation, 51

 modelkontrol, 78

 forecasting, 146

Eksempel 1.3

 nedbørs- og vandløbsmålinger, 3

 endimensionale analyser

 estimation, 72

 identifikation, 52

 flerdimensional analyse

 estimation, 134

 forecasting, 148

 identifikation, 128

 modelkontrol, 139

 strukturel analyse, 148

Eksempel 1.4

 investeringer, indkomst og forbrug,
 4

 endimensionale analyser

 estimation, 73

 identifikation, 54

 modelkontrol, 85

 flerdimensional analyse

 estimation, 136

 forecasting, 148

 identifikation, 130

 modelkontrol, 142

 strukturel analyse, 148

Eksempel 1.5

 renten på amerikanske obligatio-
 ner, 4

 endimensionale analyser

 estimation, 74

 identifikation, 58

 modelkontrol, 86

empirisk autokorrelation, 42

- fordeling af, 42
 - MA, konfidensinterval, 46
 - residualer, 76
 - fordeling, 77
 - empirisk krydskorrelationsmatriks, 125
 - residualer, 138
 - empirisk partiel autokorrelation, 44
 - AR, konfidensinterval, 46
 - residualer, 76
 - empirisk partiel autoregressionsma-
triks, 126
 - empirisk partiel krydskorrelationsma-
triks, 127
 - empirisk tidsrække, 40
 - autokorrelation, 42
 - estimation, 43, 61, 124, 132
 - flerdimensional, 124
 - identifikation, 43, 124
 - modelkontrol, 43, 124
 - partiel autokorrelation, 44
 - strukturel analyse og forecasting, 124
 - svagt stationær, 41
 - test af stationaritet, 45
 - estimation, 43, 61, 124, 132
 - AR, Yule-Walker, 69
 - asymptotisk fordeling af cls esti-
matoren, 134
 - asymptotisk fordeling af maksi-
mum likelihood estimatoren, 68
 - betinget mindste kvadrater, 65
 - cls, 65, 132
 - Eksempel 3.10, 69
 - Eksempel 3.11, 70
 - Eksempel 3.13, 72
 - Eksempel 3.14, 72
 - Eksempel 3.15, 73
 - Eksempel 3.16, 74
 - Eksempel 5.4, 134
 - Eksempel 5.5, 136
 - fordeling af maksimum likelihood
estimator, 66
 - forventet Fisher information, 67
 - maksimum likelihood, 62
 - ARMA-modeller, 61
 - observeret Fisher information, 66
 - partielt afledede af l , 66
 - ubetinget mindste kvadrater, 63
 - uls, 63
- F**
- familien af endelig dimensionale for-
delinger, 8
 - Fisher information
 - forventet, 67
 - observeret, 66
 - fordeling
 - cls estimatorer, 134
 - empirisk autokorrelation, 42
 - residualer, 77
 - empirisk krydskorrelation
 - residualer, 139
 - maksimum likelihood estimatorer, 66
 - forecast
 - differensligning, 111
 - forecast error, 108
 - kovariansmatriks, 109
 - middelværdi, 109
 - opsplitning af varians, 121
 - forecasting, 106, 124, 145
 - Eksempel 5.8, 146
 - Eksempel 5.9, 146
 - Eksempel 5.10, 148
 - Eksempel 5.11, 148
 - VARMA(p, q)-proces, 110
 - VMA-proces, 107
 - forventet Fisher information, 67
- G**
- Gaussisk proces, 10

glidende gennemsnit, 16
 baglæns, 16
 baglæns, af orden q , 16
Granger kausal, 111
Granger-kausaltet, 111
 betingelser, 113
 test, 145
VAR(p)-proces, 115
VMA-proces, 112
Wald test, 146

H

homogen differensligning, 156
hvid støj, 14, 89
 ARMA(0,0)-proces, 20
 ortogonal, 120
 innovationer, 120
spektraltæthed, 15

I

identifikation, 43, 44, 124, 125
 Eksempel 3.4, 48
 Eksempel 3.5, 51
 Eksempel 3.6, 52
 Eksempel 3.7, 54
 Eksempel 3.8, 58
 Eksempel 5.2, 128
 Eksempel 5.3, 130
ikke-negativ definit funktion, 10
impuls-respons analyse, 118
impuls-responser, 118
 VAR(p)-proces, 119
indeksmængde, 6
innovationer, 120
invertibilitet, 95
 VARMA(p, q)-proces, 105

J

Jordan kanonisk form, 153

K

karakteristisk polynomium, 92

omvendt, 92

kausaltet

 Granger, 111

 øjeblikkelig, 115

Kolmogorov (1933), 8

konsistent familie af fordelingsfunktioner, 8

konstruktiv modelkontrol, 77

konvergens i kvadratisk middel, 155

kovariansfunktion, 9

kovariansmatriks

 forecast error, 109

 VARMA(p, q)-proces, 105

 VAR(p)-proces, 96

 VMA(q)-proces, 103

krydskorrelationsfunktion, 88

krydskorrelationsmatriks

 empirisk, 125

 residualer

 empirisk, 138

krydskovariansfunktion, 87

L

Ljung-Box test, 76

Lütkepohl (2007), 4, 117, 119, 120, 134, 139, 153

løsning til system af ligninger, 19

 AR-repræsentation, 20

 eksistens og entydighed, 19

 MA-repræsentation, 19

M

MA(1)-processen

 asymptotisk varians, 68

 egenskaber ved, 28

 simulerede udfaldsfunktioner, 29

 teoretiske korrelationsfunktioner, 29

MA(2)-processen

 asymptotisk kovarians, 68

 egenskaber ved, 34

- simulerede udfaldsfunktioner, 35, 36
 - teoretiske korrelationsfunktioner, 35, 36
- MA(q)-proces, 16
 - ARMA(0, q)-proces, 20
 - flerdimensional, se VMA(q), 101
- middelværdifunktion, 9
- MINIC metoden, 47
- minimum mean squared prediktor, 107
 - betinget middelværdi, 107
- modelkontrol, 43, 75, 124, 138
 - Eksempel 3.19, 78
 - Eksempel 3.20, 78
 - Eksempel 3.21, 85
 - Eksempel 3.22, 86
 - Eksempel 5.6, 139
 - Eksempel 5.7, 142
 - konstruktiv, 77
 - Ljung-Box test, 76
- N
- normal proces, 10
- O
- observeret Fisher information, 66
- omvendt karakteristisk polynomium, 92
- operatoren B , 18
- Opgaver, 161
- opsplitning af varians på forecast error, 121
- ortogonal hvid støj, 120
- ortogonal VMA-proces, 119
- ortogonale impuls-responser
 - VAR(p)-proces, 120
- P
- pac, 12
- partiel autokorrelation, 12
 - empirisk, 44
 - rekursionsformel, 13
- partiel autoregressionsmatriks, 89
 - empirisk, 126
- partiel krydskorrelationsmatriks, 127
 - empirisk, 127
- Portmanteau-test, 138
- prediktion, 106
- prediktionsfejl, 108
- prediktor, 106
 - h -trins, 106
 - minimum mean squared, 107
- PROC ARIMA, 40
- PROC VARMAX, 40
- Q
- Quenouille (1949), 46
- R
- Reinsel (1997), 126
- residualer
 - cls, 65, 75, 138
 - empirisk autokorrelation, 76
 - empirisk partiel autokorrelation, 76
 - krydskorrelationsmatriks, 138
 - maksimum likelihood, 63, 75
 - uls, 63, 75
- \mathbb{R}^Z , 7
- S
- SAS
 - PROC ARIMA, 40
 - MINIC, 46
 - PROC VARMAX, 40
- spektralfordeling, 11
- spektralmål, 11
- spektraltæthed, 11
- stabilitet
 - VAR(p)-proces, 92
 - VARMA(p, q)-proces, 104
 - VAR(1)-proces, 91
- stationaritet
 - streng, 8

svag, 9
 empirisk, 41
 test af, 45
 stokastisk proces, 6
 anden ordens, 9
 $\Delta_{\mathbb{Z}}$, 7
 diskret tid, 6
 familien af endelig dimensionale
 fordelinger, 8
 Gaussisk, 10
 hvid støj, 14, 89
 $I(d)$ -proces, 43
 indeksmængde, 6
 integreret af orden d , 43
 kontinuert tid, 6
 kovariansfunktion, 9
 middelværdifunktion, 9
 normal, 10
 $\mathbb{R}^{\mathbb{Z}}$, 7
 strengt stationær, 8
 svagt stationær, 9
 tidsrække, 6
 tilstandsrum, 6
 udfaldsfunktion, 7
 strukturel analyse, 124, 145
 Eksempel 5.10, 148
 Eksempel 5.11, 148
 svagt stationær proces, 9
 ac, 11
 autokorrelationsfunktion, 11, 88
 autokovariansfunktion, 11, 87
 autoregression, AR, 16
 autoregression, VAR, 89
 flerdimensional, 87
 glidende gennemsnit, MA, 16
 glidende gennemsnit, VMA, 91
 kovariansfunktion, 11
 krydskorrelationsfunktion, 88
 krydskovariansfunktion, 87
 pac, 12

partiel autokorrelationsfunktion,
 12
 partiel autoregressionsmatriks, 89
 spektralfordeling, 11
 spektralmål, 11
 spektraltæthed, 11
 varians, 11

T

test

Dickey-Fuller, 45
 Granger-kausaltet, 145
 hypotesen $\Phi_{m,m} = 0$, 128
 Ljung-Box, 76
 Portmanteau, 138

tid

diskret, 6
 kontinuert, 6
 tidsrække, 6
 empirisk, 40
 flerdimensional, 87

tilstandsrum, 6

diskret, 6
 kontinuert, 6

U

udfaldsfunktion, 7

V

VAR(1)-proces

betingelse for VMA-repræsentation,
 91
 eksistens, 89
 stabil, 91

VARMA-modeller

estimation, 132
 forecasting, 145
 identifikation, 125
 modelkontrol, 138
 strukturel analyse, 145

VARMA(0, q)-proces

VMA(q), 101

- VARMA($p, 0$)-proces
 - VAR(p)-proces, 101
- VARMA(p, q)-proces, 100
 - forecast
 - i matricer i VAR-repræsentationen, 111
 - i matricer i VMA-repræsentationen, 111
 - forecasting, 110
 - invertibel, 105
 - koefficienter i VMA-repræsentation, 104
 - koefficienter i VAR-repræsentation, 105
 - kovariansmatriks, 105
 - stabil, 104
- VAR(p)-proces, 89
 - betingelse for VMA-repræsentation, 95
 - impuls-responser, 119
 - invertibel, 95
 - koefficienter i VMA-repræsentation, 96
 - kovariansmatriks, 96
 - ortogonale impuls-responser, 120
 - som højere dimensional VAR(1)-proces, 91
 - stabil, 92
 - Yule-Walker ligninger, 98
- VMA-proces
 - forecasting, 107
 - Granger-kausaltet, 112
 - opsplitning af varians på forecast error, 121
 - ortogonal, 119
 - øjeblikkelig kausalitet, 116
 - betingelser for, 117
- VMA(q)-proces, 100
 - koefficienter i VAR-repræsentation, 103
 - kovariansmatriks, 103
- W
 - Wold's Decomposition Theorem, 99
- Y
 - Yule-Walker estimation, 69
 - Yule-Walker ligninger, 23, 98
- Ø
 - øjeblikkelig kausalitet, 115
 - VMA-proces, 116
 - betingelser for, 117

Litteratur

- A. H. Andersen og P. Blæsild, *Time Series Analysis, Volume 1+2*, Forelæsningsnoter, Insitut for Mateamtiske Fag, Aarhus Universitet, 2003.
- M. S. Bartlett, *An introduction to stochastic processes, with special reference to methods and applications*, Cambridge, at the University Press, 1955.
- G. E. P. Box og G. M. Jenkins, *Times series analysis. Forecasting and control*, Holden-Day, San Francisco, Calif., 1970. Se desuden [Box et al. \(1994\)](#).
- G. E. P. Box, G. M. Jenkins og G. C. Reinsel, *Time series analysis*, Prentice Hall Inc., Englewood Cliffs, NJ, 1994, 3. udgave. Forecasting and control.
- G. E. P. Box og D. A. Pierce, *Distribution of residual autocorrelations in autoregressive-integrated moving average time series models*. *J. Amer. Statist. Assoc.*, 65, 1509–1526, 1970, ISSN 0162-1459.
- Lotte Würtz Brix, *Flerdimensionale Tidsrækker – med en anvendelse i hydrologi*, Speciale, Intitut for Matematiske Fag, Aarhus Universitet, 2010.
- SAS Insitute Inc., *SAS/ETS© 9.1 User's Guide*, SAS Institute Inc., Cary, NC, 2004.
- A. N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer-Verlag, 1933.
- H. Lütkepohl, *New introduction to multiple time series analysis*, Springer-Verlag, Berlin, 2007, 2. udgave.
- M. H. Quenouille, *Approximate tests of correlation in time-series*. *J. Roy. Statist. Soc. Ser. B.*, 11, 68–84, 1949, ISSN 0035-9246.
- G. C. Reinsel, *Elements of multivariate time series analysis*, Springer Series in Statistics, Springer-Verlag, New York, 1997, 2. udgave.
- F. Risager, *En analyse af klimaet på Grønland*, Statistiske Interna 29, Depth. Theor. Statist., Aarhus Universitet, 1976.