



Bokmål

Faglig kontakt under eksamen:

John Tyssedal 73593534/41645376

EKSAMEN I TMA4285 TIDSREKKEMODELLER

Fredag 7. desember 2012

Tid: 09:00 – 13:00

Hjelpemidler: Kalkulator HP30S, CITIZEN SR-270X eller CITIZEN SR-270X College.
Statistiske tabeller og formler, Tapir forlag.
K. Rottman: Matematisk formelsamling.
Et gult ark (A5 med stempel) med egne håndskrevne formler og notater.

Sensur senest ferdig mandag 7. januar 2013.

Merk: Alle svar skal begrunnes!

Merk: Kalmanfilterligningene er oppgitt på siste side av oppgavesettet.

Oppgave 1

Betrakt en ARMA(1,2)-modell med forventning lik null,

$$(1 - \varphi_1 B)z_t = (1 - \theta_1 B - \theta_2 B^2)a_t,$$

der φ_1 , θ_1 og θ_2 er parametre, B betegner *backshift*-operatoren, dvs. $B^k z_t = z_{t-k}$, og $\{a_t\}$ er hvit støy med forventning lik null og $\text{Var}(a_t) = \sigma_a^2$.

a) Hva betyr det at $\{z_t\}$ er invertibel?

Er $\{z_t\}$ invertibel når $\varphi_1 = \frac{1}{2}$, $\theta_1 = \frac{1}{2}$ og $\theta_2 = -\frac{1}{4}$?

Hva betyr det at $\{z_t\}$ er kovarians-stasjonær?

Er $\{z_t\}$ kovarians-stasjonær når $\varphi_1 = \frac{1}{2}$, $\theta_1 = \frac{1}{2}$ og $\theta_2 = -\frac{1}{4}$?

Vi antar i resten av denne oppgaven at $\{z_t\}$ er både invertibel og kovarians-stasjonær. Som kjent betyr dette spesielt at $\{z_t\}$ har en MA(∞)-representasjon, $z_t = \psi(B)a_t$.

- b) Utled formler for koeffisientene $\{\psi_j\}_{j=0}^{\infty}$ og vis at de kan skrives på formen

$$\psi_0 = 1, \psi_1 = \theta_1 - \varphi_1 \quad \text{og} \quad \psi_j = \varphi_1^{j-2}((\theta_1 - \varphi_1)\varphi_1 + \theta_2) \quad \text{for } j = 2, 3, \dots$$

- c) Fra MA(∞)-representasjonen av $\{z_t\}$ gitt over, vis at variansen til z_t er gitt ved

$$\gamma_0 = \sigma_a^2 \left[1 + (\theta_1 - \varphi_1)^2 + \frac{((\theta_1 - \varphi_1)\varphi_1 + \theta_2)^2}{1 - \varphi_1^2} \right].$$

Fra den samme MA(∞)-representasjonen utled også en tilsvarende formel for γ_1 .

Oppgave 2

Et plott av en observert tidsrekke z_t er vist i øverste venstre hjørne av figur 1. Samme figur viser også estimert autokorrelasjonsfunksjon og partiell autokorrelasjonsfunksjon for z_t , samt den observerte tidsrekken etter en gangs differensiering og estimert autokorrelasjonsfunksjon og partiell autokorrelasjonsfunksjon for den differensierte tidsrekken. Videre i oppgaven skal vi forutsette at vi ønsker å tilpasse en ARIMA(p,d,q)-modell til z_t .

- a) Ut fra plottene i figur 1 diskuter hvilke (tentative) verdier av p , d og q du finner naturlig å benytte.

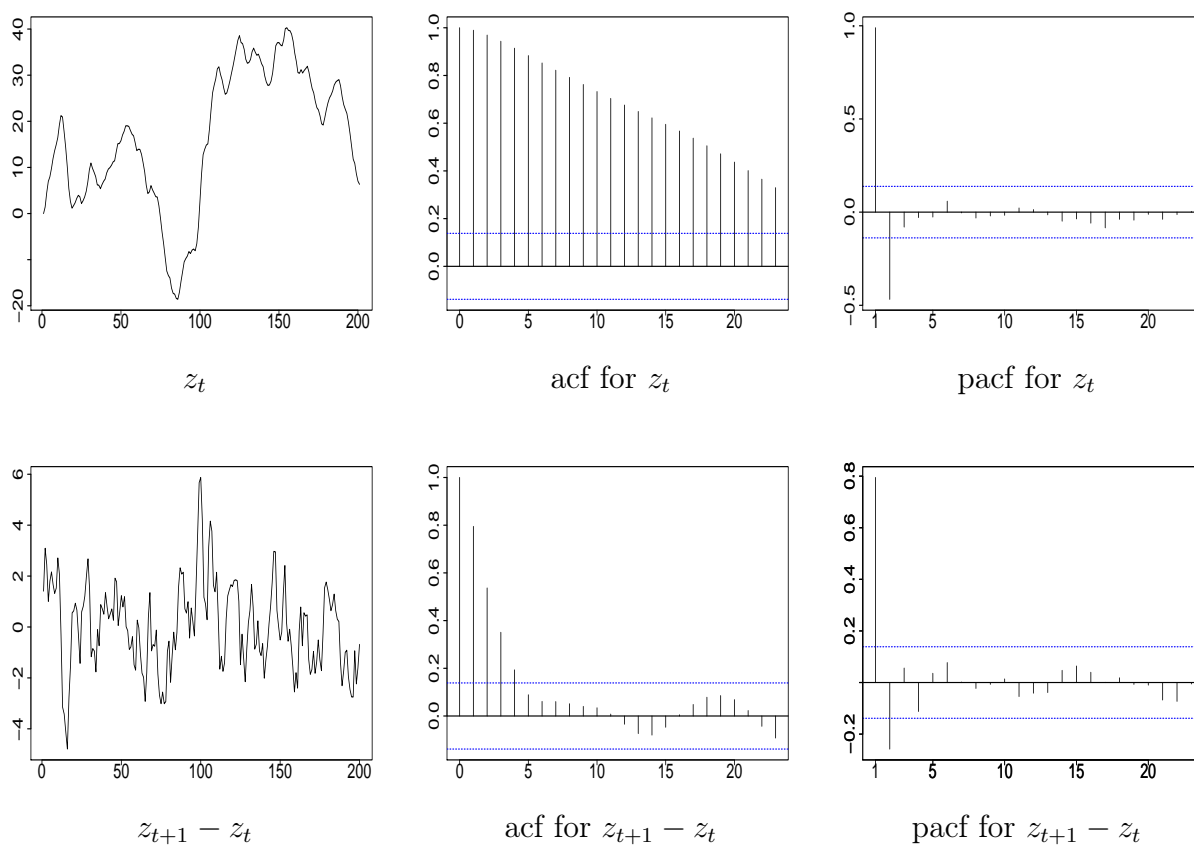
Dersom du konkluderer med at du vil forsøke med en modell hvor $d > 0$, diskuter ut fra plottene i figur 1 også hvorvidt du vil inkludere en deterministisk trendparameter, θ_0 , i modellen.

Husk å begrunne dine valg!

Uavhengig av ditt svar i spørsmålet over, anta at man velger å tilpasse hele seks modeller: ARIMA(1,1,0), ARIMA(2,1,0), ARIMA(3,1,0), ARIMA(1,1,1), ARIMA(1,1,2) og ARIMA(1,1,3), alle uten deterministisk trendparameter. Estimerte parameterverdier med tilhørende standardavvik for de seks modellene er gitt i tabell 1 og plott av estimerte residualer og estimert autokorrelasjonsfunksjon for de estimerte residualene er vist i figur 2.

- b) Basert på resultatene i tabell 1 og figur 2, diskuter kort for hver av de seks ARIMA-modellene hvorvidt den tilhørende estimerte modellen er en rimelig modell for den observerte tidsrekken.

Hvilken av de seks ARIMA-modellene vil du velge som modell for den observerte tidsrekken?

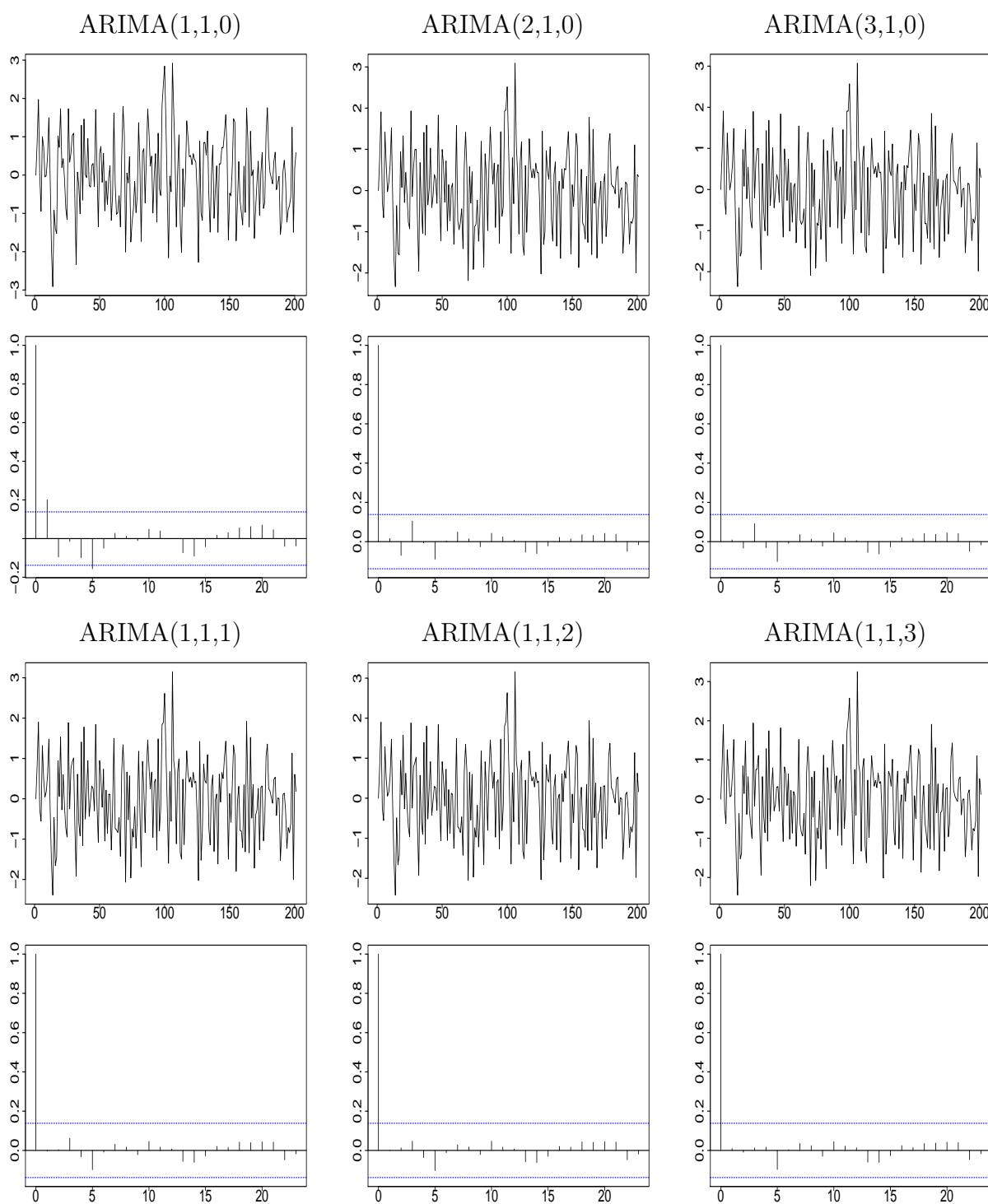


Figur 1: Øvre rad: Observert tidsrekke z_t , samt tilhørende estimert autokorrelasjonsfunksjon og partiell autokorrelasjonsfunksjon. Nedre rad: Tidsrekken z_t etter differensiering, samt tilhørende estimert autokorrelasjonsfunksjon og partiell autokorrelasjonsfunksjon.

	ARIMA(1,1,0)		ARIMA(2,1,0)		ARIMA(3,1,0)	
	verdi	st. avvik	verdi	st. avvik	verdi	st. avvik
φ_1	0.794	0.042	0.993	0.068	1.007	0.071
φ_2			-0.251	0.069	-0.304	0.099
φ_3					0.054	0.071
σ_a^2	1.084		1.016		1.013	
log-like	-292.36		-285.94		-285.66	
AIC	588.71		577.89		579.32	

	ARIMA(1,1,1)		ARIMA(1,1,2)		ARIMA(1,1,3)	
	verdi	st. avvik	verdi	st. avvik	verdi	st. avvik
φ_1	0.673	0.064	0.690	0.084	0.609	0.123
θ_1	0.350	0.084	0.330	0.107	0.398	0.130
θ_2			-0.0267	0.093	0.078	0.136
θ_3					0.105	0.101
σ_a^2	1.008		1.007		1.002	
log-like	-285.14		-285.10		-284.60	
AIC	576.29		578.20		579.21	

Tabell 1: Estimerte parameterverdier med tilhørende standardavvik, samt optimal log-likelihood og AIC-verdier for de seks modellene: ARIMA(1,1,0), ARIMA(2,1,0), ARIMA(3,1,0), ARIMA(1,1,1), ARIMA(1,1,2) og ARIMA(1,1,3).



Figur 2: Estimerte residualer, med tilhørende estimerte autokorrelasjonsfunksjon for de seks modellene: ARIMA(1,1,0), ARIMA(2,1,0), ARIMA(3,1,0), ARIMA(1,1,1), ARIMA(1,1,2) og ARIMA(1,1,3).

Oppgave 3

I denne oppgaven skal vi betrakte tilstandsmodeller (state-space models) på formen

$$x_{t+1} = \Phi x_t + w_{t+1} \quad \text{og} \quad y_t = Ax_t + v_t \quad \text{for } t = 0, 1, \dots, \quad (1)$$

der $x_0 \sim N(\mu_0, \Sigma_0)$, $w_t \sim N(0, Q)$ og $v_t \sim N(0, R)$, samt at $x_0, w_t, t = 1, 2, \dots$ og $v_t, t = 0, 1, \dots$ alle antas uavhengige av hverandre.

- a) Reformuler følgende prosesser til tilstandsmodeller gitt på formen i ligning (1), dvs. spesifiser vektorer og matriser x_t, y_t, Φ, A, Q og R for hver av de to modellene,

$$1) \quad z_t = 0.9z_{t-1} + a_t, \quad u_t = 0.5(z_t + z_{t-1}) + b_t,$$

$$2) \quad z_t = 0.9z_{t-1} + 0.5z_{t-2} + a_t, \quad r_t = 0.5(r_{t-1} + z_{t-1}) + b_t, \quad u_t = 0.5(r_t + z_t) + c_t.$$

Her er z_t, a_t, u_t, b_t, r_t og c_t alle skalare størrelser, $\{a_t\}, \{b_t\}$ og $\{c_t\}$ er uavhengige normalfordelte hvit-støy-prosesser med forventning null og varians henholdsvis σ_a^2, σ_b^2 og σ_c^2 , og i begge tilfeller observerer man kun prosessen $\{u_t\}$.

Videre i oppgaven skal vi anta at vi har en **skalar** tilstandsmodell på formen gitt i ligning (1), slik at $x_t, y_t, \Phi, w_t, Q, A, v_t, R, \mu_0$ og Σ_0 alle er skalare.

- b) Anta i dette punktet at $\Phi = \frac{1}{2}$, $A = 2$ og $R = 1$, og at $P = \lim_{t \rightarrow \infty} P_t^t$ eksisterer, der $P_t^t = \text{Var}[x_t | y_0, \dots, y_t]$.

Vis ved å ta utgangspunkt i kalmanfilterligningene gitt på siste side av dette oppgavesettet at

$$P = - \left(2Q + \frac{3}{8} \right) + \sqrt{Q + \left(2Q + \frac{3}{8} \right)^2}.$$

Skisser P som funksjon av Q . Med utgangspunkt i tilstandsmodellen som her benyttes, gi en intuitiv forklaring på den kvalitative oppførselen av P som funksjon av Q . Forklar spesielt hvorfor verdien P får når $Q = 0$ er rimelig.

- c) For en skalar tilstandsmodell på formen gitt i ligning (1), benytt egenskaper til multi-normalfordelingen til å forklare hvorfor den betingede fordelingen for x_{t+1} gitt y_0, \dots, y_t er en normalfordeling. Forklar tilsvarende at den betingede fordelingen til x_{t+1} gitt y_0, \dots, y_t, y_{t+1} også er en normalfordeling.

Benytt dette til å utlede formler for

$$x_{t+1}^{t+1} = E[x_{t+1}|y_0, \dots, y_{t+1}] \quad \text{og} \quad P_{t+1}^{t+1} = \text{Var}[x_{t+1}|y_0, \dots, y_{t+1}]$$

som funksjon av

$$x_{t+1}^t = E[x_{t+1}|y_0, \dots, y_t], \quad P_{t+1}^t = \text{Var}[x_{t+1}|y_0, \dots, y_t]$$

og størrelsene som definerer tilstandsmodellen. [Merk at du i dette punktet **ikke** skal benytte deg av de generelle kalmanfilterligningene oppgitt på siste side av oppgavesettet.]

Vedlegg:

Kalmanfilterligningene (i notasjonen benyttet på forelesningene):

$$\begin{aligned}x_{t+1}^t &= \Phi x_t^t \\P_{t+1}^t &= \Phi P_t^t \phi^T + Q \\x_{t+1}^{t+1} &= x_{t+1}^t + K_{t+1}(y_{t+1} - A_{t+1}x_{t+1}^t) \\P_{t+1}^{t+1} &= [I - K_{t+1}A_{t+1}]P_{t+1}^t \\K_{t+1} &= P_{t+1}^t A_{t+1}^T [A_{t+1}P_{t+1}^t A_{t+1}^T + R]^{-1}\end{aligned}$$