

Eksamen i STK4060/STK9060 – Tidsrekker, våren 2006

Besvarelsen av oppgavene nedenfor vil utgjøre det vesentlige grunnlaget for karaktergivingen, og utgangspunktet for den muntlige eksaminasjonen. Det er meningen dere skal arbeide selvstendig med oppgavene, men det er selvfølgelig lov å snakke sammen. Fristen for innlevering er **mandag 29. mai kl. 17**. Besvarelsen sendes fortrinnsvis som vedlegg (i .pdf format) til en e-post til eivind.damsleth@hydro.com, eller legges i min posthulle på venstre side i rommet ved siden av Ekspedisjonen i 7. etg. N.H.Abels hus. Håndskrevne besvarelser er OK, såfremt de er greit leselige. Legg bare ved de delene av utskrifter som er nødvendig for å besvare oppgavene, og pass på at det går tydelig frem hvilke deler av utskriftene du refererer til.

Lykke til!

Eivind Damsleth

- 1 -

Filen kpi.txt inneholder 243 månedlige observasjoner av konsumprisindeksen (KPI) fra Statistisk Sentralbyrå (SSB), f.o.m. jan. 1986 t.o.m. mars 2006. Dataene er også gitt i vedlegg 1.

a) Identifiser en passende SARIMA modell for disse dataene, og estimer modellens parametere. Argumenter for de valgene som gjøres (Box-Cox transformering, valg av differensiering, forslag til AR og MA orden for de ordinære og sesongrelaterte parametere).

b) Bruk modellen du har funnet til å lage prognoser for KPI for perioden jan. 2005-mars 2006 (obs. 229-243) med utgangspunkt i dataene t.o.m. desember 2004 (obs. nr. 228). Finn også de tilhørende ca. 95 % konfidensintervallene for hver enkelt prognoseverdi. Sammenlign med de faktiske verdiene fra samme periode. Konklusjon?

Denne sammenligningen gir muligens et noe optimistisk bilde av modellens prognoseegenskaper. Hvorfor?

c) KPI-verdiene er gitt med kun en desimal. Du kan anta at SSB benytter vanlige avrundingsregler. Anta at avrundingsfeilen er uniformt fordelt i intervallet $<-0.05, 0.05]$. Hva blir da en nedre grense for standardavviket til 1-skritts prognosefeilen (= residualt standardavvik hvis du ikke har benyttet noen Box-Cox transformasjon).

d) En mulig definisjon av den årlige inflasjonen i Norge (i %) er: $Y_t = \frac{X_t - X_{t-12}}{X_{t-12}} \cdot 100\%$, hvor

X_t er KPI verdien for måned t og Y_t er den prosentvise endring over de siste 12 måneder.

Bemerk at såfremt den årlige inflasjonen er nær 0 gir en 1. ordens Taylor utvikling av

$$\ln\left(\frac{X_t}{X_{t-12}}\right) = \ln(1 + Y_t / 100) \text{ at } Y_t \approx \ln\left(\frac{X_t}{X_{t-12}}\right) \cdot 100\% = (1 - B^{12}) \ln(X_t) * 100\% . \text{ Vi kan betrakte}$$

tilnærmelsen som likhet i denne sammenheng. Dette kan være nyttig i det følgende.

Norges Bank har satt som sitt inflasjonsmål at den årlige inflasjonen skal være omlag 2.5%. I den senere tid har den vært betydelig lavere. Finn en passende SARIMA modell for $\{Y_t\}$, enten ved å utlede den fra modellen du fant i deloppgave b) eller ved å identifisere/estimere en egen modell for $\{Y_t\}$. Bruk dataene t.o.m. mars 2006 til å lage en prognose for den årlige inflasjonen ved utgangen av desember 2006 og ved utgangen av desember 2007, med tilhørende ca. 95% konfidensintervall. Hva er (ca.) sannsynligheten for at Norges Bank når sitt inflasjonsmål, dvs. at inflasjonen blir $> 2.5\%$ på disse to tidspunktene?

I en industriell prosess blandes et antall forskjellige tørrstoffer etter en gitt vekttoppskrift i en blandetrommel. Etter blanding tømmes trommelen, og resultatet veies. Tømmeprosessen er ikke perfekt, slik at det blir sittende noe tørrstoff igjen i blandetrommelen etter tømningen. Denne resten blir så med i den neste fyllingen. Mengden av denne resten varierer fra fylling til fylling.

Veingen av inputmateriale er beheftet med målefeil. Måleusikkerheten (standardavviket til målefeilen) for veieutstyret er kjent. Det er også målefeil i veingen av sluttproduktet (resultatet av blandeprosessen), også med kjent standardavvik

Bruk følgende notasjon:

I_t = Total faktisk (korrekt) vekt innveid til fylling nr. t (ikke observerbar)

μ_I = Ønsket total innveid vekt iflg. oppskriften, slik at $E I_t = \mu_I$ (kjent)

Z_{I_t} = Total veiefeil ved innveing, slik at $Z_{I_t} = I_t - \mu_I$

σ_I = Standardavvik for Z_{I_t} . (Kjent, beregnet som roten av sum av varianser for de enkelte veieprosessene)

R_t = Rest i blandetrommelen etter tømning av batch nr. t (ikke observerbar)

R_t antas å følge en stasjonær AR(1) prosess: $(R_t - \mu_R) = \varphi(R_{t-1} - \mu_R) + Z_{R_t}$,

hvor $Z_{R_t} \sim \text{WN}(0, \sigma_R^2)$. μ_R antas kjent.

O_t = Faktisk vekt ut fra fylling nr. t (ikke observerbar)

Y_t = Målt vekt ut fra fylling nr. t (Observerbar). Måleprosessen er forventningsrett, slik at $E(Y_t | O_t) = O_t$

Z_{Y_t} = Målefeil ved resultatmålingen, slik at $Z_{Y_t} = Y_t - O_t$

σ_Y = Standardavvik for Z_{Y_t} : $\sigma_Y = \text{Stdv}(Y_t | O_t)$ (kjent)

Dynamikken i denne industrielle prosessen kan da skrives:

$$\begin{aligned} I_{t+1} &= \mu_I + Z_{I_{t+1}} \\ O_{t+1} &= I_{t+1} + R_t + R_{t+1} \\ R_{t+1} - \mu_R &= \varphi(R_t - \mu_R) + Z_{R_{t+1}} \\ Y_{t+1} &= O_{t+1} + Z_{Y_{t+1}} \end{aligned} \quad (1)$$

Denne industrielle prosessen kan formuleres som en *state-space* modell

$$\begin{aligned} \mathbf{X}_{t+1} &= \mathbf{F}\mathbf{X}_t + \mathbf{V}_{t+1} \\ Y_t &= \mathbf{G}\mathbf{X}_{t+1} + W_t \end{aligned} \quad (2)$$

med tilstandsvektor gitt ved $\mathbf{X}_t = (1, I_t, O_t, R_t)'$.

a) Vis at $\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \mu_I & 0 & 0 & 0 \\ \mu_I - (1-\varphi)\mu_R & 0 & 0 & (1-\varphi) \\ (1-\varphi)\mu_R & 0 & 0 & \varphi \end{bmatrix}$, $\mathbf{G} = (0, 0, 1, 0)$, $\mathbf{V}_t = \begin{bmatrix} 0 \\ Z_{I_t} \\ Z_{I_t} - Z_{R_t} \\ Z_{R_t} \end{bmatrix}$ og $W_t = Z_{Y_t}$,

og at (2) oppfyller kravene til en *state-space* representasjon.

b) Hvilke betingelser må parametrene oppfylle for at denne *state-space* representasjonen skal være stabil (stable).

c) Forsøk å gi en uformell "fysisk" tolkning av parameteren φ . Hva innebærer verdier nær +1? Nær -1? Og nær 0?

d) Den (uobserverbare) serien med faktisk output fra blandeprosessen, $\{O_t\}$, kan beskrives ved en ARMA prosess. Finn denne prosessen, og sett opp sammenhengen mellom parametrene i denne ARMA prosessen og parametrene i *state-space* representasjonen. Hva skjer når $\varphi \rightarrow 1$? Og $\varphi \rightarrow 0$?

e) Tilsvarende kan serien med de observerte veiingene av resultatet, $\{Y_t\}$, beskrives ved en ARMA prosess. Hvilken? Og hva er nå sammenhengen mellom parametrene i denne ARMA prosessen og parametrene i *state-space* representasjonen.

f) Filen batch.txt (innholdet er også gitt i vedlegg 2) inneholder 200 sekvensielle observasjoner fra $\{Y_t\}$.

For disse dataene er: $\mu_I = 1000$, $\sigma_I = 20$, $\mu_R = 20$ og $\sigma_Y = 20$.

Identifiser en ARMA modell for disse dataene, og estimer parameterne. I hvilken grad er modellen konsistent med det du fant i deloppgave e)? Bruk resultatene fra e) til å finne estimater for parametrene φ og σ_R .

Vedlegg 1: Kpi.txt – Konsumprisindeksen månedlig jan. 1986 – mars 2006

	1986-89	1990-93	1994-97	1998-2001	2002-2005	2006
Jan.	63.8	81.9	90.7	98.9	109.0	115.6
Feb.	64.0	82.3	91.0	99.3	109.3	116.6
March	64.5	83.2	91.5	99.8	109.7	116.9
April	64.9	83.2	91.6	99.9	109.7	
May	65.0	83.4	91.7	99.7	110.0	
June	66.1	83.7	91.9	100.0	110.1	
July	66.8	83.7	92.1	100.1	109.9	
Aug.	67.0	83.7	92.1	99.8	109.6	
Sep.	68.0	84.4	92.5	100.5	110.2	
Oct.	68.3	85.1	92.6	100.6	110.6	
Nov.	68.5	85.1	92.6	100.7	111.0	
Dec.	68.8	84.9	92.6	100.8	113.1	
Jan.	69.8	85.3	93.1	101.2	113.3	
Feb.	70.5	85.5	93.4	101.4	114.6	
March	71.2	86.1	94.0	102.1	113.8	
April	71.4	86.4	94.0	102.4	112.9	
May	71.6	86.6	94.1	102.2	112.3	
June	71.9	86.6	94.4	102.3	112.0	
July	72.2	86.7	94.3	102.0	111.6	
Aug.	72.3	86.6	94.1	101.7	111.9	
Sep.	73.3	87.3	94.7	102.6	112.5	
Oct.	73.4	87.3	94.6	103.1	112.4	
Nov.	73.6	87.3	94.6	103.5	112.6	
Dec.	73.9	87.3	94.6	103.6	112.6	
Jan.	74.8	87.3	94.1	104.1	112.4	
Feb.	75.2	87.5	94.2	104.6	112.6	
March	76.3	88.3	94.6	104.7	113.1	
April	76.5	88.5	95.0	105.1	113.3	
May	76.7	88.6	95.1	105.1	113.4	
June	77.1	88.8	95.2	105.7	113.4	
July	77.1	88.9	95.6	105.4	113.3	
Aug.	77.0	88.7	95.5	105.3	113.0	
Sep.	78.0	89.1	95.9	106.2	113.7	
Oct.	78.1	89.2	96.3	106.3	114.0	
Nov.	78.1	89.3	96.3	106.8	114.0	
Dec.	78.0	89.2	96.2	106.7	113.8	
Jan.	78.6	89.5	97.0	107.6	113.6	
Feb.	78.9	89.8	97.3	108.4	113.7	
March	79.6	90.6	97.5	108.6	114.2	
April	80.0	90.8	97.4	109.1	114.8	
May	80.3	90.8	97.7	109.6	115.2	
June	80.7	90.9	97.9	109.7	115.3	
July	80.8	90.8	97.7	108.9	114.9	
Aug.	80.6	90.6	97.7	108.1	115.1	
Sep.	81.2	91.0	98.1	108.7	116.0	
Oct.	81.4	91.0	98.3	108.6	116.0	
Nov.	81.4	91.0	98.4	108.7	116.0	
Dec.	81.3	90.9	98.5	108.9	115.9	

Vedlegg 2: Batch.txt – 200 (syntetiske) data fra en industriell blandeopprosess

1	941.1	51	980.0	101	1013.6	151	996.4
2	1006.9	52	970.9	102	1002.6	152	1054.2
3	1002.0	53	1030.0	103	975.3	153	975.1
4	1032.8	54	1021.8	104	1053.1	154	1017.0
5	965.0	55	952.8	105	1031.6	155	989.0
6	988.4	56	1040.9	106	956.2	156	1044.3
7	954.3	57	944.7	107	998.2	157	1002.1
8	1046.2	58	1040.1	108	1052.5	158	988.1
9	983.8	59	1024.4	109	970.7	159	1031.8
10	1007.6	60	957.7	110	1061.2	160	931.8
11	1005.9	61	1019.2	111	1021.0	161	1026.9
12	969.0	62	1036.3	112	1010.9	162	978.7
13	993.1	63	980.1	113	1013.2	163	1028.2
14	1004.2	64	962.0	114	967.4	164	933.1
15	1037.6	65	1025.4	115	1019.5	165	1058.3
16	1025.5	66	946.8	116	928.9	166	1058.7
17	912.5	67	1008.8	117	1060.1	167	890.4
18	1045.2	68	972.3	118	926.1	168	1043.6
19	961.7	69	1008.5	119	978.6	169	960.3
20	989.9	70	1055.5	120	1029.1	170	1045.8
21	1011.6	71	997.1	121	1032.0	171	960.3
22	934.2	72	1070.5	122	1020.2	172	1000.6
23	1033.5	73	921.8	123	1009.6	173	1008.8
24	970.2	74	1064.9	124	960.9	174	962.8
25	973.3	75	1000.2	125	1041.7	175	1042.8
26	1027.3	76	1005.3	126	932.8	176	1022.5
27	950.7	77	974.9	127	1072.8	177	1009.7
28	1002.4	78	984.3	128	968.1	178	980.3
29	934.8	79	956.8	129	1010.2	179	1036.3
30	1017.5	80	1066.1	130	989.6	180	990.3
31	1001.7	81	971.5	131	1045.7	181	946.3
32	967.1	82	1015.8	132	929.3	182	1035.7
33	1058.7	83	1003.8	133	1061.2	183	990.9
34	969.5	84	991.2	134	924.0	184	986.4
35	1033.6	85	974.5	135	993.6	185	1057.3
36	959.9	86	1049.0	136	1019.9	186	974.4
37	1007.1	87	975.9	137	1017.1	187	1019.8
38	928.8	88	1034.7	138	965.5	188	1002.5
39	1006.0	89	982.1	139	982.8	189	1087.1
40	932.9	90	986.2	140	1066.1	190	987.3
41	1021.8	91	1021.6	141	970.4	191	1036.5
42	992.8	92	950.9	142	1074.7	192	1003.0
43	1014.8	93	1039.1	143	1031.3	193	1054.3
44	956.4	94	965.6	144	959.5	194	990.4
45	1033.8	95	992.7	145	1037.3	195	997.6
46	980.2	96	1004.0	146	1014.4	196	1026.9
47	1000.4	97	995.5	147	1045.3	197	966.8
48	957.7	98	1002.0	148	981.6	198	1027.0
49	1093.3	99	1004.4	149	1048.7	199	959.6
50	975.8	100	1021.3	150	992.2	200	1057.1