UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i	INF 3440 / INF 4440 — Signalbehandling
Eksamensdag:	27. oktober 2003 – 10. november 2003
Tid for eksamen:	12.00 – 12.00

Oppgavesettet er på 7 sider.

Vedlegg:	Ingen
Tillatte hjelpemidler:	Alle trykte og skrevne

Kontroller at oppgavesettet er komplett før
du begynner å besvare spørsmålene.

Eksamensoppgaven er et individuelt arbeid, slik at hver student skal levere egen besvarelse.

Resultatene skal presenteres i en skriftlig rapport der teori, utskrifter av program-kode og figurer vises. Kun program-kode skrevet i Matlab vil bli rettet! Husk at både regneoppgaver og programmeringskode skal kunne leses og forstås av sensor, uten for mye besvær. Det vil derfor bli lagt vekt på at regneoppgavene er klart besvart, og at de har en logisk oppbygning og at det er lett å finne hva som er endelig svar. Alle svar basert på Matlab-utregninger skal dokumenteres med relevant program-kode. Program-koden skal være oversiktlig og kommentert.

Oppgave 1

Ryssdalutvalget foreslår en styrking av heltidsstudenten, og uten ekstrajobb ser det ut til at INF3440-studentene kommer til å få dårlig råd. Heldigvis har foreleser en idé til hvordan studentene muligens kan tjene noen ekstra penger uten altfor store anstrengelser.

En aksjemegler i et ikke-navngitt meglerhus i Oslo har tipset om bevegelser i markedet de siste månedene som tyder på at det kommer en ny oppgangsperiode på børsen. Det ryktes om at flere sentrale indeksforvaltere allerede har gått inn med store summer på OBX-indeksen på Oslo Børs.

Til tross for store tap på slutten av 90-tallet har foreleser ennå tro på at aksjeinvesteringer kan lønne seg. Dessverre mangler han den nødvendige egenkapitalen. Han har derfor sagt seg villig til å investere et større beløp fra hver av studentene[1]. Men istedet for å satse blindt denne gangen også har han utviklet noen investeringsstrategier. En idé går ut på å designe et FIR-filter som kan brukes til å predikere aksjemarkedet.

[1] Mot et mindre vederlag.

(Fortsettes på side 2.)

Prediksjonsfilteret av orden p er gitt ved

$$\hat{x}[n] = -\sum_{k=1}^{p} a_k x[n-k], \tag{1}$$

hvor $\hat{x}[n]$ er den predikerte verdien av $x[n]$. Problemet vårt er å bestemme koeffisientene $\{a_k\}_{k=1}^{p}$ som utfører prediksjonen best mulig. Et vanlig mål er å minimere feilen i minste kvadraters forstand (antar $N \geq p$):

$$E = \sum_{n=1}^{N} |e[n]|^2 = \sum_{n=1}^{N} |x[n] - \hat{x}[n]|^2.$$

Lineær prediksjonsproblemet kan også skrives på matriseform som

$$-\underbrace{\begin{pmatrix} x[1] & \dots & x[p] \\ \vdots & & \vdots \\ x[N-p] & \dots & x[N-1] \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} \tilde{a}_1 \\ \vdots \\ \tilde{a}_p \end{pmatrix}}_{\tilde{a}} + \underbrace{\begin{pmatrix} e[p+1] \\ \vdots \\ e[N] \end{pmatrix}}_{e} = \underbrace{\begin{pmatrix} x[p+1] \\ \vdots \\ x[N] \end{pmatrix}}_{x},$$

eller som $-X\tilde{a} + e = x$. Dermed kan feilen skrives som $e = X\tilde{a} + x$.

I Figur 1 er OBX-indeksen plottet i tidsrommet 2.1.1997 til 3.10.2003. Dataene som er brukt er hentet fra Oslo Børs, og ligger tilgjengelig på:

http://www.ifi.uio.no/∼inf3440/OBX.txt

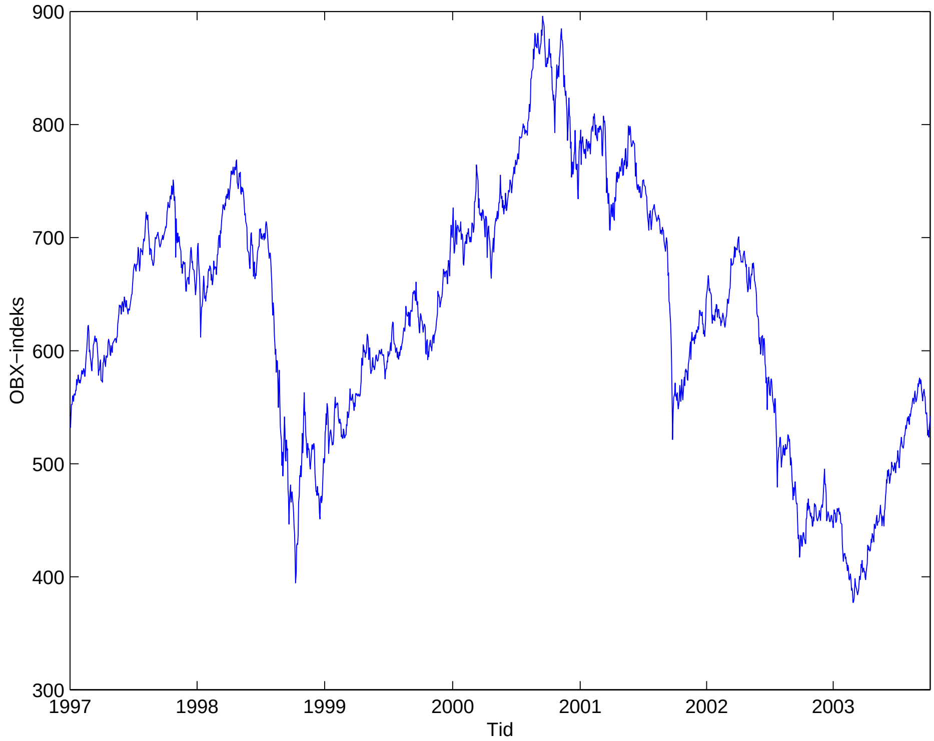
Filen består av fire kolonner. Kolonne én er datoer på formen år–måned–dato, mens kolonne to inneholder indeks-verdier for tilhørende datoer. De to resterende kolonnene skal vi ikke bruke, men inneholder henholdsvis OBX-indeksens høyeste og laveste verdi den dagen.

For å lese inn dataene til Matlab kan dere benytte dere av kodesnutten vist under, som legger OBX-indeksen i variabelen `OBX` og tilhørende datoer i variabelen `date`. Variabelen `date` består av åtte siffer. De fire første er år, de to neste er måned, mens de to siste er dag.

```
OBX = []; date = [];
fid = fopen(['OBX.txt']);
line = fgetl(fid); line = fgetl(fid); line = fgetl(fid);
while isstr(line)
  OBX = [OBX; str2num(line(12 : 17))];
  date = [date; floor(str2num(line([1 : 4, 6, 7, 9, 10])))];
  line = fgetl(fid);
end
clear fid, line;
```

Vi skal gjøre noen antagelser for å forenkle oppgaven litt. Siden børsen bare er åpen på vanlige arbeidsdager skal vi anta at rente-året er 250 dager langt, og at året kun består av disse rente-dagene (dvs. året i denne oppgaven er definert til å vare i 250 dager istedet for 365 dager).

(Fortsettes på side 3.)



Figur 1: OBX-indeksen i tidsrommet 2. januar 1997 til 3. oktober 2003.

1a

Hva blir utbyttet dersom vi setter inn 10 000 kr. i en bankkonto med 5 % rente første virkedag i 1999 og lar de stå frem til 3. oktober 2003?

Hva blir utbyttet dersom vi investerer samme beløpet i OBX-indeksen over samme periode?

1b

Lag en Matlab-funksjon `predfilt` som finner x og X ut fra data gitt i en innparameter, og som returnerer vektoren $\tilde{a}$ som minimerer $e^T e$. For å finne vektoren $\tilde{a}$ gitt x og X kan man i Matlab for eksempel bruke operatoren \, som utfører venstredivisjon av matriser.

1c

Vi skal nå bruke en tredje ordens modell ($p = 3$). Lag vektoren med predikerte verdier fra hele 1999 ved $\hat{x}_1 = -X\tilde{a}$. Lag også tilsvarende vektor $\hat{x}_2$ ved å finne koeffisientene $\{a_k\}_{k=1}^{p}$, og sette disse inn i (1). Du kan gjerne bruke `filter`-funksjonen i Matlab til å gjøre dette. Dokumenter med programkode!

Regn ut den totale kvadrerte feilen $e^T e$ mellom x og $\hat{x}_1$ eller $\hat{x}_2$, og plott differensen $|x[n] - \hat{x}[n]|^2$. Plott den totale kvadrerte feilen for $p = 1, \ldots, 25$.

(Fortsettes på side 4.)

Hva er en rimelig verdi av p, slik at ikke feilen avtar betydelig for større verdier?

1d

Vi skal nå prøve å benytte oss av følgende investeringsstrategi ved å bruke lineærpredikatoren med samme verdi for p som vi fant i Oppgave 1c. Lineærpredikatoren skal brukes til å finne morgendagens verdi av OBX-indeksen. Dersom denne vokser mer enn den risikofrie renten investeres alle pengene i OBX-indeksen. Ellers settes alle pengene i banken.

I likhet med tidligere satser vi 10 000 kr. Hva blir utfallet dersom vi bruker denne strategien fra begynnelsen av 2000 og frem til idag? Sammenlign med svarene fra Oppgave 1a. Sammenlign også med gevinsten dersom vi alltid hadde sikker viten om morgensdagens verdier av OBX-indeksen.

1e

Parsevals teorem sier at for to sekvenser $g[n]$ og $h[n]$ er

$$\sum_{n=-\infty}^{\infty} g[n]h^*[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} G(e^{j\omega})H^*(e^{j\omega})\mathrm{d}\omega.$$

Bruk dette resultatet til å vise analytisk at koeffisientene a_k er valgt slik at de minimerer uttrykket

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{X(e^{j\omega})}{\hat{X}(e^{j\omega})}\right|^2 \mathrm{d}\omega,$$

hvor

$$\hat{X}(e^{j\omega}) = \frac{1}{1+\sum_{k=1}^{p} a_k e^{-j\omega k}}.$$

Plott $X(e^{j\omega})$ og $\hat{X}(e^{j\omega})$ i samme plott, basert på samme p-verdi som du fant i Oppgave 1c og data fra hele 1999. Hva ser du? Det kan hende aksene på grafen må skaleres for å få et klart plott av begge funksjonen. Skaler $\hat{X}(e^{j\omega})$ med faktoren

$$E = \sum_{n=1}^{N} |e[n]|^2.$$

Hva ser du nå? Kan du forklare hvorfor denne faktoren ble valgt?

Gjenta samme analyse for $p = 150$, og sammenlign!

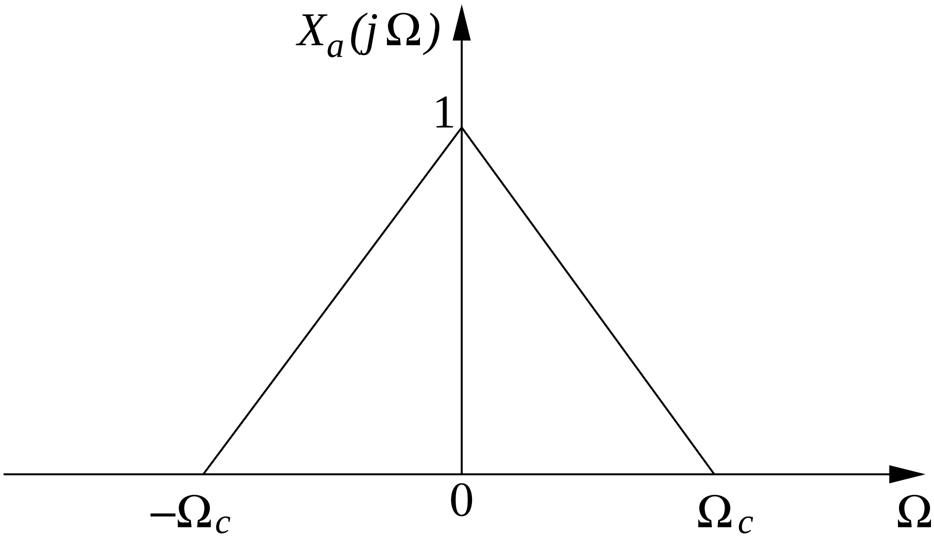
Oppgave 2

La det tids-kontinuerlige signalet $x_a(t)$ være gitt ved

$$X_a(j\Omega) = \begin{cases} 1 - \frac{|\Omega|}{\Omega_c}, & |\Omega| \le \Omega_c \\ 0, & \text{ellers} \end{cases}$$

i frekvensdomenet, slik som vist i Figur 2.

(Fortsettes på side 5.)



Figur 2: $x_a(t)$ vist i frekvensdomenet.

2a

Utled et analytisk uttrykk for $x_a(t)$ ved hjelp av den inverse tids-kontinuerlige Fourier-transformen, for et tids-kontinuerlig signal $g_a(t)$ gitt ved

$$g_a(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G_a(j\Omega)e^{j\Omega t}\mathrm{d}\Omega.$$

Skriv så enkelt som mulig!

2b

La nå $x[n] = x_a(nT)$, hvor samplingsraten $T = \frac{\pi}{40\Omega_c}$. La $X(e^{j\omega})$ være den tids-diskrete Fourier-transformen til $x[n]$. Skisser $X(e^{j\omega})$ for $0 \leq \omega < 2\pi$.

2c

For å regne ut den inverse DFT numerisk kan Matlab-funksjonen `ifft` være nyttig.

Definer nå

$$X[k] = X(e^{j\omega})\big|_{\omega=\frac{2\pi k}{N}}, \quad 0 \leq k \leq N-1,$$

og la $\hat{x}[n]$ være den inverse DFT til $X[k]$ $(0 \leq n \leq N-1)$.

Finn $\hat{x}[n]$ for $\Omega_c = 10\pi$ og $N = 800$ ved hjelp av Matlab. Plott $\hat{x}[n]$ for $n \in \{0, \ldots, 200\}$. Er forskjellen stor fra tilsvarende plott for $x[n]$, og i så tilfelle hvorfor? Argumenter med relevant teori fra læreboken!

(Fortsettes på side 6.)

Oppgave 3

Definerer LTI-systemet

$$h[n] = \begin{cases} a^n, & n \geq 0, \quad 0 < a < 1 \\ 0, & n < 0. \end{cases}$$

3a

La inngangen til systemet være

$$x[n] = \begin{cases} 1, & 0 \leq n \leq N-1 \\ 0, & \text{ellers.} \end{cases}$$

Finn et analytisk uttrykk for utgangen $y[n]$ ved å bruke definisjonen på konvolusjon.

3b

Bruk $\mathcal{Z}$-transformen (og den inverse) til å finne et analytisk uttrykk for utgangen $y[n]$, med samme inngang som i Oppgave 3a.

3c

Finn $\mathcal{Z}$-transformen til sekvensen

$$y[n] = \sum_{k=-\infty}^{n} x[k]$$

uttrykt ved $\mathcal{Z}$-transformen til $x[n]$. Kan du si noe om ROC til $Y(z)$?

Oppgave 4

4a

Det kan vises at periodske sampler til et signal bestående av en sum av komplekse eksponensialer tilfredsstiller

$$x[n+p] + a_{p-1}x[n+p-1] + \cdots + a_0 x[n] = 0, \tag{2}$$

for en gitt orden p og sett med koeffisienter $a_0, \ldots, a_{p-1}$.

Et spesialtilfelle av (2) kan brukes til å lage en effektiv rekursiv formel for å generere sampler av signalet

$$x[n] = \sin(\omega n + \phi). \tag{3}$$

Den rekursive formelen er gitt ved

$$x[n+2] = \alpha x[n+1] - x[n]. \tag{4}$$

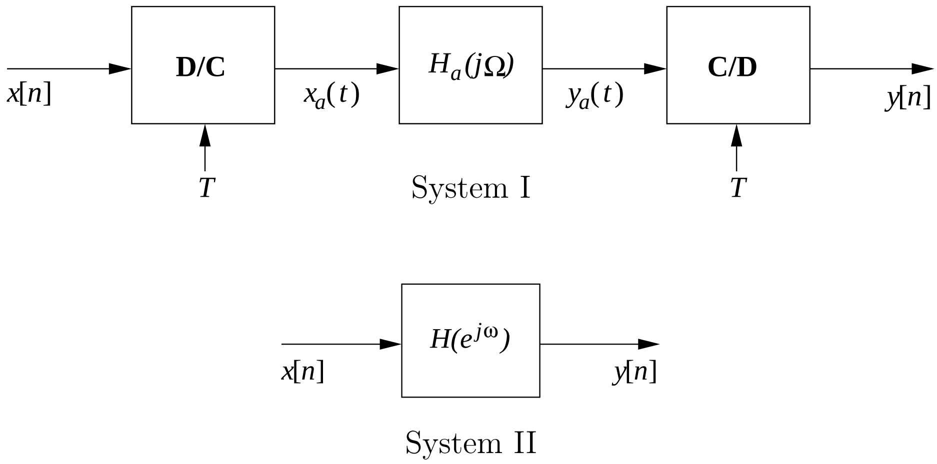
Hvilke krav må α tilfredsstille for at (3) skal være en løsning av (4)? Hva må initialbetingelsene være?

(Fortsettes på side 7.)

4b

Figur 3 viser to LTI-systemer, henholdsvis System I og System II. System II har frekvensresponsen

$$H(e^{j\omega}) = e^{-j\omega/2}, \quad |\omega| \leq \pi.$$



Figur 3: To LTI-systemer.

Bestem T og impulsresponsen $h_a(t)$ til $H_a(j\Omega)$ slik at System I og System II er ekvivalente.

4c

Regn ut hva utgangen $y[n]$ blir dersom inngangen er som gitt i (3), dvs.

$$x[n] = \sin(\omega n + \phi).$$