T-61.3040 Statistical Modeling of Signals

Final Exam 2.9.2008

In the exam you are allowed to have a calculator (non-programmable or memory emptied) and basic mathematical tables (no tables containing material directly associated with the course). The results of the exam will be announced eventually through the Noppa system, and not anymore through the newsgroup opinnot.tik.informaatiotekniikka.

1. (max 6p)

Explain *briefly* the following topics without unnecessary detail:
i) Wide-sense stationarity (WSS) (2p)
ii) Wiener filter (2p)
iii) Wold decomposition (2p)

2. (max 6p)

You have observed the real-valued zero-mean WSS process $x(n)$ and you know the following values of the autocorrelation: $r_x(0) = 1$, $r_x(1) = 2/3$, $r_x(2) = 0$, $r_x(3) = 1/3$, $r_x(4) = 0$.

i) Form an optimal linear predictor having two weights, where you predict the value $x(n)$ using the observations $x(n-1)$ and $x(n-2)$. (2p)
ii) What is the mean-square prediction error of your predictor? How much smaller is the prediction error compared to the variance $\mathrm{Var}(x(n))$ of the process? (2p)
iii) Can you improve the prediction accuracy using a linear predictor with two weights, where you predict the value $x(n)$ using observations $x(n-2)$,$x(n-3)$ or observations $x(n-3)$,$x(n-4)$? (2p)

3. (max 6p)

Answer the following propositions either "true" or "false": you may also leave any of them unanswered. A correct answer gives 1 point, a wrong answer -1 points, and a missing answer zero points. However, the total number of points you receive from this problem cannot become negative; the total number of points is at least zero. No need to justify your answers.

a) If a process consists of two sinusoids in white noise, and the sinusoids have different frequencies, then using the periodogram we can resolve the frequencies of the sinusoids, if there are enough observations of the process.
b) For both the power spectrum and the pseudospectrum, the integral over some frequency band tells the power of the process on that band.
c) The conditional variance $\mathrm{var}(x(n+1)|x(n), x(n-1), \ldots)$ of a normally distributed AR(p) process does not depend on the parameters $a(1), \ldots, a(p)$.
d) The resolution of the periodogram is worse than the resolution of Bartlett's method and Welch's method.
e) An MA(q) process has infinitely many nonzero autocorrelations $r_x(k)$.
f) When applying Wiener filtering to the process $x(n) = d(n) + v(n)$, if the noise $v(n)$ is zero-mean white noise that is uncorrelated with the desired signal $d(n)$, then the Wiener filter can be solved using only the autocorrelation $r_v(k)$ of the noise and the autocorrelation $r_d(k)$ of the desired signal.

4. (max 6p)

You have measured the following observations from the real-valued zero-mean WSS process $x(n)$:
$x(0) = 1$, $x(1) = 0$, $x(2) = -1/2$, $x(3) = 0$, $x(4) = 1/4$, $x(5) = 0$
You want to estimate a linear predictor with two weights, which at time n predicts the value $x(n)$ based on the values $x(n-1)$,$x(n-2)$. You want to use the LMS algorithm to estimate the predictor.

i) You want the LMS algorithm to converge in the mean. Choose a step size so that this happens; estimate the required information and make any necessary assumptions. (2p)
ii) Run the LMS algorithm for two update steps, starting from time $n = 2$ and multipliers $\mathbf{w} = [1, 1]^T$. (2p)
iii) Assume that we know the mean-square prediction error (MSE) of an AR(2) model optimal for this process; the value of the MSE is $s_{min} = 0.1$.

   Assume that we have run the LMS algorithm twice: in run 1 we used a step size that is one thousandth of the step size you chose in part i); in run 2 we used one tenth thousandth of the step size in part i). Assume that in both runs the algorithm was run for a very long time, and the algorithm had converged in the mean.

   In both runs, after the convergence in the mean, the MSE of the LMS algorithm is still on average larger than s_{min}; explain why this is so. Denote the average MSE achieved in run 1 by s_1 and the average MSE achieved in run 2 by s_2; estimate (approximate) the ratio $(s_2 - s_{min})/(s_1 - s_{min})$. (2p)

T-61.3040 Signaalien tilastollinen mallinnus

Tentti 2.9.2008

Tentissä saa olla mukana laskin (ei ohjelmoitava tai muisti tyhjä) ja matematiikan perustaulukot (ei taulukoita joissa on kurssin aiheisiin suoraan liittyvää materiaalia). Tentin tulokset ilmoitetaan aikanaan Noppa-järjestelmän kautta, ei siis enää uutisryhmässä opinnot.tik.informaatiotekniikka.

1. (max 6p)

Selitä *lyhyesti* seuraavat asiat menemättä tarpeettomasti yksityiskohtiin:
i) Väljässä mielessä stationäärisyys (WSS) (2p)
ii) Wiener-suodin (2p)
iii) Woldin hajotelma (2p)

2. (max 6p)

Olet havainnut reaaliarvoisen nollakeskiarvoisen WSS-prosessin $x(n)$ ja tiedät seuraavat autokorrelaation arvot: $r_x(0) = 1$, $r_x(1) = 2/3$, $r_x(2) = 0$, $r_x(3) = 1/3$, $r_x(4) = 0$.

i) Muodosta optimaalinen kaksikertoiminen lineaarinen ennustin, missä ennustat arvoa $x(n)$ havaintojen $x(n-1)$ ja $x(n-2)$ avulla. (2p)

ii) Mikä on ennustimesi keskimääräinen neliöllinen ennustusvirhe? Kuinka paljon ennustusvirhe on prosessin varianssia $\mathrm{Var}(x(n))$ pienempi? (2p)

iii) Voitko parantaa ennustustarkkuutta sellaisen kaksikertoimisen lineaarisen ennustimen avulla, missä ennustat arvoa $x(n)$ havaintojen $x(n-2)$,$x(n-3)$ tai havaintojen $x(n-3)$,$x(n-4)$ avulla? (2p)

3. (max 6p)

Vastaa seuraaviin väitteisiin joko "tosi" tai "epätosi" tai jätä vastaamatta. Oikea vastaus antaa yhden pisteen, väärä -1 pistettä ja vastaamatta jättäminen nolla pistettä. Tästä tehtävästä saamasi kokonaispistemäärä ei kuitenkaan voi laskea negatiiviseksi; kokonaispistemäärä on vähintään nolla. Vastauksia ei tarvitse perustella.

a) Jos prosessi koostuu kahdesta sinisignaalista valkoisessa kohinassa, ja sinisignaaleilla on eri taajuudet, niin periodogrammin avulla voidaan erottaa sinisignaalien taajuudet, mikäli prosessista on riittävästi havaintoja.

b) Sekä tehospektrin että pseudospektrin integraali jonkin taajuuskaistan yli kertoo prosessin tehon tällä kaistalla.

c) Normaalijakautuneen AR(p)-prosessin ehdollinen varianssi $\mathrm{var}(x(n+1)|x(n), x(n-1), \ldots)$ ei riipu parametreista $a(1), \ldots, a(p)$.

d) Periodogrammin resoluutio on huonompi kuin Bartlettin menetelmän ja Welchin menetelmän resoluutio.

e) MA(q)-prosessilla on äärettömän monta nollasta poikkeavaa autokorrelaatiota $r_x(k)$.

f) Jos prosessin $x(n) = d(n) + v(n)$ Wiener-suodatuksessa kohina $v(n)$ on nollakeskiarvoista valkoista kohinaa joka ei korreloi halutun signaalin $d(n)$ kanssa, niin Wiener-suodin voidaan ratkaista pelkästään kohinan autokorrelaation $r_v(k)$ ja halutun signaalin autokorrelaation $r_d(k)$ avulla.

4. (max 6p)

Olet mitannut seuraavat havainnot reaaliarvoisesta nollakeskiarvoisesta WSS-prosessista $x(n)$:
$x(0) = 1$, $x(1) = 0$, $x(2) = -1/2$, $x(3) = 0$, $x(4) = 1/4$, $x(5) = 0$
Haluat estimoida kaksikertoimisen lineaarisen ennustimen, joka hetkellä n ennustaa arvoa $x(n)$ arvoista $x(n-1)$,$x(n-2)$. Haluat käyttää LMS-algoritmia estimoimaan ennustimen.

i) Haluat, että LMS-algoritmi suppenee odotusarvon suhteen. Valitse askelpituus siten, että näin käy; estimoi tarvittavat tiedot ja tee mahdolliset tarvittavat oletukset. (2p)

ii) Aja LMS-algorithmia kahden päivitysaskeleen verran, lähtien hetkestä $n = 2$ ja kertoimista $\mathbf{w} = [1, 1]^T$. (2p)

iii) Oletetaan, että tunnetaan tälle prosessille optimaalisen AR(2)-mallin keskimääräinen neliöllinen ennustusvirhe (MSE); sen arvo on $s_{min} = 0.1$.

Oletetaan että LMS-algoritmia on ajettu kaksi kertaa: kerralla 1 käytettiin askelpituutta, joka on tuhannesosa i)-kohdassa valitsemastasi askelpituudesta; kerralla 2 käytettiin kymmenestuhannesosaa i)-kohdan askelpituudesta. Oletetaan, että kummallakin kerralla algoritmia ajettiin hyvin pitkään, ja se oli supennut odotusarvon suhteen.

Kummallakin ajokerralla, kun algoritmi oli supennut odotusarvon suhteen, LMS-algoritmin MSE on silti edelleen keskimäärin suurempi kuin s_{min}; selitä, mistä tämä johtuu. Merkitään kerralla 1 saavutettua keskimääräistä MSE:tä s_1 ja kerralla 2 saavutettua keskimääräistä MSE:tä s_2; estimoi (approksimoi) suhde $(s_2 - s_{min})/(s_1 - s_{min})$. (2p)