



Hemuppgift 3, 2009, SF1851 Optimeringslära för E.

Examinator: Per Enqvist, penqvist@math.kth.se.

Hemtalsassistent: Maja Karasalo, karasalo@math.kth.se.

I denna hemuppgift är det tillåtet att samarbeta i grupper om högst två personer. Begränsat samarbete med andra grupper är också tillåtet men måste nämnas i rapporten. En rapport per grupp ska lämnas in, och ni ska berätta med egna ord hur ni har löst problemet. Det är inte tillåtet att kopiera delar av annans grupps rapport! Svaren till frågorna i uppgiften ska finnas i den skriftliga huvudrapporten. Relevanta utskrifter av matlabkörningar och plottar ska bifogas till rapporten. Fullständig programkod ska finnas tillgänglig på begäran. Skriv era namn, personnummer och emailadress på framsidan av rapporten.

Rapporten ska lämnas in (utskriven på papper) innan måndagen den 12:e klockan 17.00. Notera att denna deadline är strikt.

I denna uppgift ska vi betrakta signaler som kan modelleras som följer av korrelerade stokastiska variabler, nämligen stationära stokastiska processer. En stationär stokastisk process kan även betraktas som en summa av cosinusfunktioner med olika frekvenser och slumpmässig amplitud och fas. Inom spektralestimering så analyserar man frekvensinnehållet för sådana signaler. Som i talbehandling (speech processing) är det vanligt att använda stationära stokastiska processer och spektralestimering för att, t.ex. skatta fundamentalfrekvensen för den ljudvåg som skapas av stämbanden.

Givet en dataföljd som kan approximeras som en stationär stokastisk process, så ges ett sätt att skatta frekvensinnehållet i signalen av att använda ett s.k. periodogram. Om signalen ges av $y = (y_1, y_2, \dots, y_N)$, då kan vi definiera Fourier transformen $Y(\theta) = \sum_{k=0}^N y_k \cos(k\theta)$. En skattning av spektraltätheten ges då av $\Phi(\theta) = 1/N \cdot |Y(\theta)|^2$ och denna beskriver energiinnehållet i signalen y vid frekvensen θ .

I Matlab kan en skattning erhållas genom

```
>> Phi = abs(fft(y)).^2;
```

Denna skattning av spektraltätheten bestämmer värden vid Fourier frekvenserna $\theta_k = k\pi/(N/2)$ där $k = 0, \dots, N/2 - 1$. Det är dock mycket mera användbart att ha tillgång till en kontinuerlig funktion som approximerar dessa skattningar och representerar de huvudkaraktäristiska egenskaperna för spektraltätheten över alla frekvenser.

Det är vanligt att representera ljudsignaler i deciBel (dB) skala. Låt $\Psi = 10 \log_{10}(\Phi)$.

Antag nu att en rationell modell ska anpassas till de datavärden som skattats m.h.a. periodogrammet i dB-skalan. Vi väljer att använda en rationell modell, antag att $\hat{\Psi}(\theta) = \frac{S(\theta)}{A(\theta)}$, där

$$S(\theta) = s_0 + s_1 \cos(\theta) + \dots + s_n \cos(n\theta),$$

och

$$A(\theta) = 1 + a_1 \cos(\theta) + \dots + a_n \cos(n\theta).$$

I denna uppgift låter vi $n = 8$. Vad vi skulle vilja göra är att minimera avståndet mellan de skattade värden för Ψ och den rationella funktionen $\hat{\Psi}$, men vilket avstånd ska vi använda? Det mest naturliga valet av avstånd vore att använda den diskretiserade 2-normen av skillnaden mellan Ψ och $\hat{\Psi}$, d.v.s.,

$$\|\Psi - \hat{\Psi}\|_2^2 = \sum_{k=0}^{N/2-1} |\Psi(\theta_k) - \hat{\Psi}(\theta_k)|^2$$

Detta leder till ett icke-linjärt minstakvadratproblem som vi inte ska lösa i denna uppgift. *Kommentar: Matlabkommandot `lsqcurvefit` kan användas för att (försöka) lösa det icke-linjära minstakvadratproblemet att minimera $\|\Psi - \frac{S}{A}\|_2$. Men, det icke-linjära minstakvadratproblemet är inte konvext, har flera lokala minimum och konvergens för algoritmen beror på var man startar. Inga extrapoleringar delas ut för att man löser det problemet.*

Eftersom vi vill formulera problemet som ett kvadratisk optimeringsproblem (linjärt minstakvadratproblem) betraktar vi istället den diskretiserade 2-normen för $A(\Psi - \hat{\Psi}) = A\Psi - S$, d.v.s. $\|A\Psi - S\|_2^2$. Notera att detta är en kvadratisk funktion i koefficienterna för A och S .

Signalen som vi ska använda är en ljudsignal. Vi ska använda signalen i Matlab som kallas för `handel`, vilket är en snutt av en kör som sjunger Händels Messiah. Utför följande kommandon:

```
>> load handel
>> N=1024;
>> Psi = 10*log10(1/N*abs(fft(y(10001:10000+N)))).^2);
>> Psi = Psi(1:N/2);
```

(Du kan lyssna på den genom att skriva `sound(y,Fs)`)

Detta ger dig en vektor med värden för $\Psi(\theta_k)$ där $\theta_k = k\pi/(N/2)$ och $k = 0, \dots, N/2 - 1$.

a) Betrakta optimeringsproblemet: $\min_{A,S} \|A\Psi - S\|_2^2$.

$$\|\Psi(\theta)A(\theta) - S(\theta)\|_2^2 = \sum_{k=0}^{N/2-1} |\Psi(\theta_k)A(\theta_k) - S(\theta_k)|^2$$

där

$$\theta_k = k\pi/(N/2), \quad k = 0, \dots, N/2 - 1.$$

Formulera optimeringsproblemet som ett kvadratisk programmeringsproblem på formen

$$\min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{b}\|^2,$$

där variabelvektorn \mathbf{x} , bivillkorsmatrisen \mathbf{A} och högerledsvektorn \mathbf{b} ska bestämmas.

Ledning: Ett bra sätt att börja är att först bestämma din vektor \mathbf{x} med variabler och i vilken ordning dessa ska stå. Därefter bestämma ett uttryck för varje term i summan i målfunktionen. Slutligen, att uttrycka hela summan med hjälp av matrisalgebra.

Kom ihåg att hålla reda på vad som är era okända variabler och vad som är givna värden.

Notera att om $\mathbf{z} = (z_1, z_2, \dots, z_M)^T$, i så fall gäller $\sum_{k=1}^M |z_k|^2 = \mathbf{z}^T \mathbf{z}$

Förklara och motivera eran formulering, i.e., beskriv variabler, bivillkor och målfunktion som ni använder och förklara varför detta QP löser det givna problemet.

Till sist, lös det kvadratiske optimeringsproblemet med Matlab.

- b) Analysera lösningen som ni har fått. Plotta $\Psi(\theta_k)A(\theta_k)$ och $S(\theta_k)$ i samma graf med $\theta_k, k = 0, 1, \dots, N/2 - 1$ på x-axeln, och kommentera angående anpassningen, är det som förväntat? Plotta även $\Psi(\theta_k)$ och $S(\theta_k)/A(\theta_k)$ i en annan graf och kommentera anpassningen, är det som förväntat?
- c) Nu lägger vi till extra bivillkor till optimeringsproblemet. Antag att vi vill ha en lösning där

$$\sum_{k=1}^n a_k = 0, \quad \text{och} \quad \sum_{k=0}^n s_k = 0.$$

Vi har lärt oss två metoder för att lösa kvadratiske problem med linjära bivillkor, nu ska hälften av er använda nollrumsmetoden och den andra hälften ska använda Lagrange metoden.

- (i) Om den äldsta gruppmedlemmen har födelsedag på en udda dag så ska ni använda nollrumsmetoden för att lösa detta problem. Det numeriskt bästa sättet att bestämma ett nollrum använder singularvärdesfaktorisering, och denna finns implementerad i Matlab kommandot `null`.
- (ii) Om den äldsta gruppmedlemmen har födelsedag på en jämn dag, ska ni använda Lagrange metoden istället.
- d) Analysera lösningen ni har fått till problemet med bivillkor. Plotta $\Psi(\theta_k)A(\theta_k)$ och $S(\theta_k)$ i samma graf och kommentera angående anpassningen, är det som förväntat? Plotta även $\Psi(\theta_k)$ och $S(\theta_k)/A(\theta_k)$ i en annan graf och kommentera angående anpassningen, är det som förväntat?

Kommentar: Ett liknande angreppssätt används inom talbehandling för att karaktärisera korta talsegment. Tanken är att, om du jämför två inspelningar av en person som säger samma ord två gånger, så är datan ganska olika i tidsdomänen. Däremot, så i frekvensplanet så är de betydligt mera lika men har fortfarande för många parametrar. Genom att anpassa en rationell modell kan man reducera antalet parametrar kraftigt, nämligen till parametrarna $s_0, \dots, s_n, a_1, \dots, a_n$. Dessa kan sedan användas för taligenkänning, talariigenkänning, talkomprimering m.m.

Lycka till!