Signaler och System

TENTAMEN I MÖNSTERIGENKÄNNING, 3P

13 JANUARI 1999, 8.00-14.00

Examinator: Björn Hansson, tel 018 - 471 3071
Hjälpmedel: Mathematical Handbook, Mathematics Handbook (BETA), räknare, utdelad formelsamling

Observera att alla svar måste motiveras! Svar utan motivering ger poängavdrag.
Skriv bara på ena sidan av papperet!

Lycka till!

1. (3p)
   En examinator i en viss kurs har på tentamensinstruktionerna bett studenterna att endast skriva på ena sidan av papperet. Om vi antar att det finns en sannolikhet p att denna uppmaning inte läses (samt att alla lästa uppmaningar följs) blir antalet studenter som skriver på båda sidorna en binomialfördelad stokastisk variabel X. Med totala antalet studenter n får vi då

   $$P(X = i|p) = \begin{pmatrix} n \\ i \end{pmatrix} p^i(1-p)^{n-i}, \quad i = 0, 1, \ldots, n, \quad E[X|p] = np$$

   a) Givet N st tentamina och utfallen $\mathcal{N} = \{n_1, n_2, \ldots, n_N\}$ samt $\mathcal{X} = \{x_1, x_2, \ldots, x_N\}$, härled och beräkna en ML-skattning av p under antagandet att $x_k \in \mathcal{X}$ är oberoende av varandra.

   b) På en tidigare tentamina visade det sig att det bland 49 skrivande studenter var två som inte upptäckte uppmaningen. Vad blir ML-skattningen av p i detta fall?

2. (5p)
   Civ och Civerth skall åka på semester men kan inte komma överens om hur länge de skall stanna på varje resmål. De har därför gjort varsin tabell med värden på hur gärna de vill stanna på respektive ställe. De har bakat ihop allt till en kostnadsfunktion och kommer att uppnå den optimala kompromissen om du kan hjälpa dem med att minimera funktionen

   $$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x},$$

   där A är en symmetrisk matris.

   a) Ange hur $\mathbf{x}_n$ skall uppdateras med gradientmetoden. Använd $\mathbf{x}_0$ som startvärde.

   b) Vad krävs av steglängden η och $\mathbf{A}$ för att metoden skall konvergera?

   c) Utför ett steg med gradientmetoden. Använd $\eta = \frac{1}{16}$, $\mathbf{x}_0 = (3 \;\; 4)^T$ och

   $$\mathbf{A} = \begin{pmatrix} 0 & 2 \\ 2 & 3 \end{pmatrix}$$

   Får vi konvergens i detta fall? Motivera!

3. (5p)
I statistisk mönsterigenkänning kan man ibland behöva tillämpa icke lärarledda metoder, där klustring med 'c-means'-algoritmen kan ge bra resultat. Antag att $\mathbf{x}_i$ är normalfördelade med känd varians σ^2 och att en träningsmängd $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ finns tillgänglig.

   a) Vad är det för skillnad på träningsmängderna som används av en lärarledd resp icke lärarledd träningsmetod?

   b) Beskriv i detalj (gärna med pseudo-kod) hur 'c-means'-algoritmen ser ut i detta fall. För klassificeringen har du tillgång till en "avståndsfunktion" $d(\mathbf{x}_i, \mu_j)$ som anger "avståndet" mellan träningsvektor $\mathbf{x}_i$ och klass c_j med väntevärde μ_j.

   c) Vilka startvärden behövs för algoritmen? Nämn något om betydelsen för hur dessa väljs.

4. (6p)
Du skall konstruera en klassificerare som kan skilja på objekt som är lika vanliga och härrör från två olika klasser (c_1 och c_2). För detta ändamål har du tillgång till sensorer som mäter två värden per objekt och du vet att mätsignalerna är normalfördelade.

   a) Ange a priori-sannolikheten för respektive klass.

   b) Beräkna diskrimantfunktionen för respektive klass.

   c) Vilken ekvation ges den optimala beslutsytan av? Skriv på formen

$$\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = 0$$

   d) Använd följande väntevärden och kovariansmatriser

$$\mu_1 = \begin{pmatrix} 3 \\ 3 \end{pmatrix}, \quad \mu_2 = \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \quad \Sigma = \Sigma_1 = \Sigma_2 = \frac{1}{2} \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$$

   Beräkna den optimala beslutsytan och ange den explicit i de ingående variablerna (x_1 och x_2 eller x och y). Vilken geometrisk form har den? Rita noggrannt en figur där bl a nivåytor och beslutsytor ingår med korrekta skalor. Kovariansmatrisen Σ har följande egenvärden och egenvektorer

$$\lambda_1 = 4, \quad \lambda_2 = 1, \quad \mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

5. (6p)
Du skall implementera en XOR-grind med ett neuronnät bestående av neuroner med två ingångar. Olinjäriteten är en 'hard limit' och ges av

$$\varphi(v, \theta) = \begin{cases} 0 & , \quad v < \theta \\ 1 & , \quad v \geq \theta \end{cases}$$

   a) Definiera problemet och markera de olika utsignalerna i mönsterrummet.

   b) Motivera varför problemet ej går att lösa med endast en neuron.

   c) Konstruera ett nät samt ange viktvektorer $\mathbf{w}_i$ och 'bias':er θ_i som löser problemet.

   d) Visa hur θ_i kan inkorporeras i en utökad viktvektor $\tilde{\mathbf{w}}_i$.

6. (4p)
Du skall träna ett litet neuronnät med fyra ingångar, två gömda noder och tre utgångar. Alla neuroner är olinjära (sigmoid, $\varphi(v) = \frac{1}{1+e^{-v}}$) med utsignaler i intervallet $[0, 1]$. Du har tillgång till träningsdata enligt följande

$$\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}, \quad \mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_N\}, \quad \mathbf{x}_k = \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \\ x_4^{(k)} \end{pmatrix}, \quad \mathbf{d}_k = \begin{pmatrix} d_1^{(k)} \\ d_2^{(k)} \\ d_3^{(k)} \end{pmatrix}$$

a) Rita nätet, numrera neuronerna och markera var respektive vikt kommer in.

b) Givet $\Delta w_{ji}^{(k)}$ för träningsexempel k och nodpar $i \to j$, ange matematiskt hur vikterna skall uppdateras med 'pattern'- resp 'epoch learning'.

c) Ange hur samtliga $\Delta w_{ji}^{(k)}$ skall beräknas.

7. (6p)
En enorm asteorid är på väg mot jorden och Civerth har tänkt sig undkomma katastrofen genom att utrusta sin husbil med en fotosensor och en jetmotor. Han har nämligen räknat ut att asteoriden har en mycket karakteristisk ljusstyrka och genom att mäta denna kan han upptäcka om asteoriden är ovanför husbilen. Om så är fallet trycker han gasen i botten för att undkomma.
Enda problemet är att Civerth inte riktigt törs lita på sina kunskaper i mönsterigenkänning och ber därför dig om hjälp. Han berättar att de objekt han mäter ljusstyrkan på kan delas in i två klasser: *asteorid* (c_1) samt *övrigt* (c_2), där den senare är uppskattningsvis hundra gånger vanligare. Han påstår vidare att mätvärdena är normalfördelade med följande väntevärden och varianser

$$\mu_1 = 2, \quad \sigma_1^2 = 0.04, \quad \mu_2 = 0, \quad \sigma_2^2 = 1$$

Han anser sig behöva tre handlingar som klassificeringsmetoden skall välja mellan: *gasa* (α_1), *stanna* (α_2) samt *fråga ESA* (α_3). De objekt som hamnar i den senare gruppen klassas av ESA mot en smärre avgift. Efter att noga övervägt situationen har Civerth kommit fram till följande kostnadsfunktioner

$$\begin{array}{lcl@{\qquad}lcl} \lambda_{11} & = & 0 & \lambda_{12} & = & 5 \\ \lambda_{21} & = & 100 & \lambda_{22} & = & 0 \\ \lambda_{31} & = & 10 & \lambda_{32} & = & 10 \end{array}$$

a) Visa $P(c_1|x)f(x)$ och $P(c_2|x)f(x)$ grafiskt. Vilket resultat skulle Bayesisk klassificering ge?

b) Härled de förväntade kostnaderna för respektive handling givet en mätning x.

c) Efter vilken regel väljs den optimala (minst förväntad kostnad) handlingen ut givet en mätning x? Beräkna och markera de optimala beslutsgränserna i figuren ovan.

d) Vad blir sannolikheten att Civerth blir mosad av asteoriden resp sannolikheten att han förbrukar bränsle i onödan? Markera dessa i figuren.