Uppsala universitet
Signaler and System
MG 018 - 471 3229

Tentamen i mönsterigenkänning

Teknisk fysik, 4E, 4Sy Magistern, fredag 12 december, 1997, 8.00-14.00

Hjälpmedel: Mathematical Handbook, Mathematics Handbook (BETA), räknare formelsamling utdelad på lektion.

Obs! Svar utan motivering ger noll poäng!

Problem #1 (1+1+1+1=4p)
a) Förklara med hjälp av en figur varifrån namnet "error backpropagation" kommer.
b) På vilket sätt är stokastisk approximation och "pattern learning" relaterade?
c) Två mönster $\mathbf{x} = [1 - 1 - 1 - 1]$ och $\mathbf{y} = [-1 - 11 - 1]$ lagras i ett Hopfieldnät. Hur utvecklas tillståndsvektorn då nätet startar i tillståndet $\mathbf{x}$?
d) Härled LMS-regeln (deltaregeln) för en neuron med olineariteten $tanh(x)$.

Problem #2 (3p)
Härled optimal (minsta felsannolikhet) beslutsgräns för två normalfördelade mönster med $\mathbf{m}_1 = [-1, 0]^T, \mathbf{m}_2 = [+1, 0]^T, C_1 = C_2 = I, P(klass1) = P(klass2) = 0.5$.

Problem #3 (2+1=3p)
Antag att mönster $\mathbf{x}$ genereras från en Gaussian mixture,

$$f(\mathbf{x}|\mathbf{w}) = \sum_{j=1}^{J} P(c_j) f(\mathbf{x}|c_j, \mathbf{w}_j)$$

a) Beskriv en algoritm i form av ett antal punkter eller i form av pseudokod som visar hur man genererar exempel från en sådan fördelning. Du kan anta att du har tillgång till en slumpgenerator och till en funktion gauss$(\mathbf{m}, C)$ som genererar ett exempel från en normalfördelning med medelvärde $\mathbf{m}$ och kovariansmatris C.
b) Vid ett experiment där 10 mönster genererades visade det sig att alla exempel genererades från samma normalfördelning $f(\mathbf{x}|c_j, \mathbf{w}_j)$. Hur stor är sannolikheten för denna händelse?

Problem #4 (3p)
Exempel genereras med samma sannolikhet från två normalfördelningar med medelvärden $+1$ och -1, båda har variansen ett. Beräkna felsannoliheten för en klassificerare som har beslutsgräns vid 0.1. Ledning: Utnyttja tabell för att beräkna aktuell integral.

Problem #5 (3p)
Vid ett experiment observeras ett konstant mönster $\mathbf{s}$ i additivt gaussiskt brus $\mathbf{v}$ med medelvärde noll och kovariansmatris C i form av en vektor $\mathbf{x} = \mathbf{s} + \mathbf{v}$. 120 exempel $\{\mathbf{x}_p\}, p = 1, 2, \ldots, 120$ finns tillgängliga. Bestäm maximum likelihoodskattning av mönstret $\mathbf{s}$.

Problem #6 (2+2=4p)
Vid ett experiment genererades följande mönster: $[2,1]^T, [0,1]^T, [1,101]^T, [1,-99]^T$.
a) Utför PCA på dessa data; bestäm medel, egenvärden och egenvektorer.
b) Komprimera och mönstret $[0,25]^T$ till en dimension och rekonstrera sedan. Obs! Glöm inte medelvektorn!

Problem #7 (1+1=2p)
Linjära diskriminantfunktioner för att klassificera data i två klasser kan skapas direkt med hjälp av olika metoder. Grundproblemet är att hitta en vektor $\mathbf{w}$ och en tröskel w_0 sådan att $\mathbf{w}^T\mathbf{x} > w_0$ ska gälla för alla (så många som möjligt) från den ena klassen.
a) Förklara hur problemet kan konverteras till att bestämma en vektor $\mathbf{u}$ sådan att

$$A\mathbf{u} > \mathbf{0}$$

där A är en matris som skapas med hjälp av exemplen. Ledning: Byt tecken på alla exempel i den ena klassen.
b) Förklara hur det konverterade problemet i deluppgift a) kan lösas med hjälp införande av en positiv koefficientvektor $\mathbf{b}$ och pseudoinversen till A.

Problem #8 (3p)
a) Förklara matematiskt varför felytan som beskriver medelkvadratfelet vid träning av en linjär neuron aldrig har lokala minima (lokala minima kan finnas i den meningen att ett helt underrum är en lösning, men inte i övrigt).
b) Förklara med hjälp av figur fördelen med att ha adaptiv steplängd med avseende på konvergenshastighet och lokala minima. Obs! Detaljerad figur krävs - annars noll poäng.

Lycka till!