

TENTAMEN A482: NEURALE NETWERKEN

vrijdag 29 augustus 1997, 09.00 - 12.00 uur

1. Backpropagation training

Bij het backpropagation netwerk gebruikt men veelal de zgn. “mean squared error” als foutmaat.

- a. Wat verstaat men onder de “mean squared error”? Geef for-mule.
- b. Hoe luidt het functie-approximatie theorema voor backpropagation netwerken?
- c. Voornoemde error hangt af van de gewichten in het backpropagation netwerk. Wat verstaat men in dit verband onder het error-oppervlak?
- d. Geef enige eigenschappen van het in punt c. genoemde error-oppervlak voor het backpropagation netwerk.

2. Leerregels

- a. Onderstaand zijn enige verschillende typen leerregels genoemd. Geef van elk een korte tekstuele (d.w.z. zonder mathematische formules) typering, waarin zo goed mogelijk de essentie naar voren komt:
 - Leerregel van Hebb (Coincidence learning)
 - Performance learning
 - Competitive learning.

- b. Performance learning wordt toegepast in de zgn. ADALINE, een eenvoudig type processing element voor de vorming van affiene combinaties:

$$y' = w_0 + w_1 x_1 + \dots + w_n x_n = w^T x$$
$$w = (w_0, w_1, \dots, w_n)^T$$
$$x = (x_0, x_1, \dots, x_n)^T; \quad x_0 = 1$$

We wensen een gewichtsvector $w = w^*$ te bepalen zodanig dat de ADALINE een willekeurig proces $y = f(x)$ zo goed mogelijk benadert, d.w.z. dat de LMS (Least Mean Square) error:

$$F(w) = \sum_k (y_k - y'_k)^2$$

minimaal is (E is de verwachtingsoperator, $y_k = f(x_k)$, y'_k is de ADALINE-output voor input x_k ; x_k ($k = 1, 2, 3, \dots$) wordt uit R_{n+1} getrokken overeenkomstig een waarschijnlijkheidsdichtheid p).

- b 1.** Wat verstaat men onder voornoemde verwachtingsoperator E ?
- b2.** Hoe ziet het oppervlak $F(w)$ in R_{n+2} eruit en welke conclusie volgt hieruit voor het bestaan van w^* (vector w waarvoor $F(w)$ minimaal is)?
- b3.** w^* kan gevonden worden via de gradient (steepest descent) methode:

$$w_{k+1} = w_k + \alpha E[\delta_k x_k]; \delta_k = y_k - y'_k.$$

Leid dit af.

- b4.** Hoe luidt de Widrow/Hoff regel en wat is de relatie met de in b3 genoemde formule?
- b5.** De Widrow/Hoff regel kent twee veel voorkomende varianten: de batch versie en de momentum versie. Wat houden deze varianten in?

3. Neural control

Onderdeel van veel neural control systemen vormt de “**neural identifier**”.

- a.** Wat is een “neural identifier”?
- b.** Verklaar de noodzaak van zgn. TDLs (Tapped Delay Lines) bij gebruik van een backpropagation netwerk als identifier.
- c.** Wat is Iterative Inversion Control? Geef configuratie-schema en verklaar de rol van de “neural identifier” hierin.
- d.** Wat is Forward-Modelled Inverse Control? Geef configuratie-schema en verklaar de rol van de “neural identifier” hierin.

4 . Kohonen netwerk

- a.** Bespreek kort de training van een Kohonen netwerk. Beschouw hiertoe achtereenvolgens:
- de **competitie** tussen de **neuronen** onderling (geef in dit verband ook de **formule** voor Euclidische afstand tussen input-vector en gewichten-vector van een neuron)
 - de **aanpassing** van de gewichten-vector van het winnend neuron en diens omgeving; geef **mathematische formulering**.
- b.** Veronderstel, u doet bij enige duizenden mensen metingen m.b.t. hun bloeddruk en hartslag. Alle gemeten waarden liggen in de “(bloeddruk, hartslag)-ruimte” in een gegeven rechthoek, bepaald door de **minimale** en **maximale** waarden van deze metingen. U wilt de statistische eigenschappen van de metingen bepalen met behulp van een Kohonen netwerk, bestaande uit 21 neuronen in een 1-dimensionele structuur (d.w.z. een “string” van 21 neuronen).
Veronderstel verder dat 1/3 van de mensen die u gemeten heeft een erg lage bloeddruk en erg lage hartslag blijkt te hebben (binnen de grenzen van bovengenoemde rechthoek), 1/3 van de mensen een erg hoge bloeddruk en erg hoge hartslag, terwijl het overige deel van de metingen uniform verdeeld blijkt over het **restant** van de rechthoek.
Hoe kunt u deze verdeling aflezen uit het getrainde Kohonen netwerk? Beschouw achtereenvolgens
- de “feature map” in de **gewichtenruimte** van het netwerk
 - de “feature map” op het **netwerk zelf** (topologische continuïteit).

5. Group Method of Data Handling (GMDH)

Een onbekende functie $y = f(x_1, x_2, x_3, x_4, x_5, x_6)$ moet zo goed mogelijk benaderd worden m.b.v. een GMDH netwerk.

- Hoe ziet de architectuur van een dergelijk netwerk er uit? Hoeveel **neuronen** zijn er in de eerste laag, de **enerlaatste laag** en de **laatste laag**?
- Geef **algemene formule** voor het **uitgangssignaal** van een neuron in het GMDH netwerk als functie van zijn **inganssignalen**.
- De output van het gehele netwerk is een zgn. **Ivakhnenko polynoom** in x_1, x_2, x_3, x_4, x_5 en x_6 . Wat is dit? Indien het aantal hidden layers van het GMDH netwerk N bedraagt, wat is dan de **graad** van het resulterende Ivakhnenko polynoom?
- Beschrijf verbaal in hoofdlijnen het **trainen** van een GMDH netwerk