

Podle vztahu $\delta_i^{h-1} = y_i^{h-1} (1 - y_i^{h-1}) \sum_{k=1}^n w_{ki}^h \delta_k^h$ zpětně šíříme chybu do vrstvy, která je blíže vstupům. Na závěr kroku 4 modifikujeme váhy:

$$w_{ij}^l(t+1) = w_{ij}^l(t) + \Delta w_{ij}^l(t),$$

$$\Theta_i^l(t+1) = \Theta_i^l(t) + \Delta \Theta_i^l(t).$$

Krok 4 opakujeme pro všechny vrstvy sítě tak, že začínáme vrstvou výstupní ($l = o$), pak následují skryté vrstvy ($l = h$). V případě, že zpracováváme skrytou vrstvu, která je nejbližší vstupní vrstvě, nahradíme y_j^{l-1} v rovnici pro výpočet Δw za odpovídající vstupní hodnotu, tedy za x_j .

Krok 5. Ukončení výběru vzorů z trénovací množiny

Jestliže jsme předložili síti všechny vzory z trénovací množiny pokračujeme krokem 6, jinak se vracíme na krok 2.

Pozn: Při sekvenčním výběru je realizace tohoto testu jednoduchá. Vybíráme-li však vzory náhodně, musíme v průběhu učení poznamenávat, které vzory byly použity, a které nikoli.

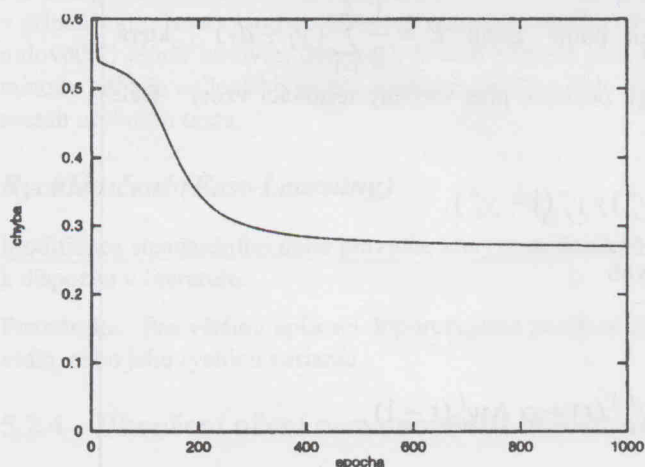
Krok 6. Ukončení procesu učení

Jestliže byla chyba neuronové sítě za poslední epochu menší, než námi zvolené kritérium, nebo jsme vyčerpali počet epoch, pak učení ukončíme. Jinak pokračujeme krokem 2.

Problém uváznutí v lokální minimu je možné řešit:

- ☐ vhodnou volbou parametrů η a α ,
- ☐ vhodnou strategií výběru vzorů z trénovací množiny (např.) náhodně,
- ☐ vhodným počátečním nastavení vah.

Pokud zvolíme větší hodnotu parametru η bude krok tak velký, že malá lokální minima překročíme. Bohužel příliš velký parametr vede k nežádoucím oscilacím. Poněkud účinnější je parametr α , který přidává do postupu váhovým prostorem setrvačnost. Budeme-li mít dostatečnou setrvačnost, udržíme si stejný směr jakým jsme se do lokálního minima dostali. Takto lokální minimum překonáme.



Parametry η a α se volí v rozsahu $<0,1>$.

Nejméně odolná strategie výběru vzorů z hlediska uváznutí v lokálním minimu je sekvenční výběr vzorů z trénovací množiny. Při této strategii může velice pomoci setrvačnost. Nejúčinnější metodou pro překonání lokálních minim se jeví náhodný výběr. Dalším významným faktorem, který značně ovlivňuje výsledek učení, je nastavení vah na počátku učení. Váhy musí být dostatečně malé, náhodně vygenerované v rozsahu menším ± 0.3 .

Obr. 5.12 Průběh chyby při uváznutí v lokálním minimu

Pokud je čtenář zklamán tím, že nedostal konkrétní doporučené hodnoty pro volbu parametrů, či univerzální strategii výběru, musí si uvědomit, že všechny