Home | Startseite | Impressum | Kontakt | Gästebuch

Aufgabe: Bringen Sie ein neuronales Netz mit Hilfe der Teleskopsumme zur Konvergenz.

Lösung: Das neuronale Netz gehorche der Trainingsabbildung

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix}.$$

Ausgehend von einem ersten Näherungsvektor für die Eingangsneuronen

$$\begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix}$$

ergibt sich mit der Trainingsmatrix ein erster Ist-Vektor für die Ausgangsneuronen:

$$\begin{pmatrix} y_1^{(1)} \\ \vdots \\ y_i^{(1)} \\ \vdots \\ y_m^{(1)} \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix}.$$

Die folgende Abbildung entspricht somit der Anfangsbedingung unseres neuronalen Netzwerks:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} - \begin{pmatrix} y_1^{(1)} \\ \vdots \\ y_i^{(1)} \\ \vdots \\ y_m^{(1)} \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \left[\begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} \right].$$

Die zweite Näherung erhalten wir analog:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} - \begin{pmatrix} y_1^{(2)} \\ \vdots \\ y_i^{(2)} \\ \vdots \\ y_m^{(2)} \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \left[\begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} x_1^{(2)} \\ \vdots \\ x_j^{(2)} \\ \vdots \\ x_n^{(2)} \end{pmatrix} \right],$$

wobei wir im folgenden beide Seiten der Gleichung mit Termen erweitern, die sich gegenseitig wegheben. Stellen wir entsprechend um, so folgt

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} + \begin{pmatrix} y_1^{(1)} \\ \vdots \\ y_i^{(1)} \\ \vdots \\ y_m^{(1)} \end{pmatrix} - \begin{pmatrix} y_1^{(2)} \\ \vdots \\ y_i^{(2)} \\ \vdots \\ y_m^{(2)} \end{pmatrix} - \begin{pmatrix} y_1^{(1)} \\ \vdots \\ y_i^{(1)} \\ \vdots \\ y_m^{(1)} \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix}$$
$$\times \left[\begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} - \begin{pmatrix} x_1^{(2)} \\ \vdots \\ x_j^{(2)} \\ \vdots \\ x_n^{(2)} \end{pmatrix} - \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} \right].$$

Mit Hilfe der Substitution

$$\begin{pmatrix} \Delta x_1^{(1)} \\ \vdots \\ \Delta x_j^{(1)} \\ \vdots \\ \Delta x_n^{(1)} \end{pmatrix} = \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} - \begin{pmatrix} x_1^{(2)} \\ \vdots \\ x_j^{(2)} \\ \vdots \\ x_n^{(2)} \end{pmatrix}$$

können wir daraus den Ausdruck

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} - \begin{pmatrix} y_1^{(2)} \\ \vdots \\ y_i^{(2)} \\ \vdots \\ y_m^{(2)} \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \left[\begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \Delta x_1^{(1)} \\ \vdots \\ \Delta x_j^{(1)} \\ \vdots \\ \Delta x_n^{(1)} \end{pmatrix} - \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} \right]$$

ableiten. Im 3. Iterationsschritt leiten wir nach dem gleichen Prinzip den dritten Ist-Vektor her:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} + \begin{pmatrix} y_1^{(1)} \\ \vdots \\ y_i^{(1)} \\ \vdots \\ y_m^{(1)} \end{pmatrix} - \begin{pmatrix} y_1^{(2)} \\ \vdots \\ y_i^{(2)} \\ \vdots \\ y_m^{(2)} \end{pmatrix} + \begin{pmatrix} y_1^{(2)} \\ \vdots \\ y_i^{(2)} \\ \vdots \\ y_m^{(2)} \end{pmatrix} - \begin{pmatrix} y_1^{(3)} \\ \vdots \\ y_i^{(3)} \\ \vdots \\ y_n^{(3)} \end{pmatrix} - \begin{pmatrix} y_1^{(1)} \\ \vdots \\ y_i^{(1)} \\ \vdots \\ y_m^{(1)} \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix}$$

$$\times \left[\begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} - \begin{pmatrix} x_1^{(2)} \\ \vdots \\ x_j^{(2)} \\ \vdots \\ x_n^{(2)} \end{pmatrix} + \begin{pmatrix} x_1^{(2)} \\ \vdots \\ x_j^{(2)} \\ \vdots \\ x_n^{(2)} \end{pmatrix} - \begin{pmatrix} x_1^{(3)} \\ \vdots \\ x_j^{(3)} \\ \vdots \\ x_n^{(3)} \end{pmatrix} - \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} \right].$$

Mit den beiden Differenzen

$$\begin{pmatrix} \Delta x_1^{(1)} \\ \vdots \\ \Delta x_j^{(1)} \\ \vdots \\ \Delta x_n^{(1)} \end{pmatrix} = \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} - \begin{pmatrix} x_1^{(2)} \\ \vdots \\ x_j^{(2)} \\ \vdots \\ x_n^{(2)} \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} \Delta x_1^{(2)} \\ \vdots \\ \Delta x_j^{(2)} \\ \vdots \\ \Delta x_n^{(2)} \end{pmatrix} = \begin{pmatrix} x_1^{(2)} \\ \vdots \\ x_j^{(2)} \\ \vdots \\ x_n^{(2)} \end{pmatrix} - \begin{pmatrix} x_1^{(3)} \\ \vdots \\ x_j^{(3)} \\ \vdots \\ x_n^{(3)} \end{pmatrix}$$

ergibt sich weiter

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} - \begin{pmatrix} y_1^{(3)} \\ \vdots \\ y_i^{(3)} \\ \vdots \\ y_n^{(3)} \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \left[\begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \Delta x_1^{(1)} \\ \vdots \\ \Delta x_j^{(1)} \\ \vdots \\ \Delta x_n^{(1)} \end{pmatrix} + \begin{pmatrix} \Delta x_1^{(2)} \\ \vdots \\ \Delta x_j^{(2)} \\ \vdots \\ \Delta x_n^{(2)} \end{pmatrix} - \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} \right].$$

Für beliebiges k können wir die Summation in gewohnter Weise fortsetzen:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} - \begin{pmatrix} y_1^{(k+1)} \\ \vdots \\ y_i^{(k+1)} \\ \vdots \\ y_n^{(k+1)} \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \left[\begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \Delta x_1^{(1)} \\ \vdots \\ \Delta x_j^{(1)} \\ \vdots \\ \Delta x_n^{(1)} \end{pmatrix} + \ldots + \begin{pmatrix} \Delta x_1^{(k)} \\ \vdots \\ \Delta x_j^{(k)} \\ \vdots \\ \Delta x_n^{(k)} \end{pmatrix} - \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} \right].$$

Fassen wir sämtliche Differenzen zusammen,

$$\sum_{l=1}^{k} \Delta x_j^{(l)} = \sum_{l=1}^{k} \left(x_j^{(l)} - x_j^{(l+1)} \right),$$

so erhalten wir in Summationsschreibweise die allgemeine Netzwerkgleichung

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} - \begin{pmatrix} y_1^{(k+1)} \\ \vdots \\ y_i^{(k+1)} \\ \vdots \\ y_n^{(k+1)} \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \left[\begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} + \sum_{l=1}^{k} \begin{pmatrix} \Delta x_1^{(l)} \\ \vdots \\ \Delta x_j^{(l)} \\ \vdots \\ \Delta x_n^{(l)} \end{pmatrix} - \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} \right].$$

Wenn wir die Definition

$$x_j^{(l)} = \frac{x_j^{(1)}}{l}$$

zugrunde legen und diese in die Teleskopsumme einsetzen, überleben nur der erste und der letzte Term,

$$\sum_{l=1}^{k} \left(x_j^{(l)} - x_j^{(l+1)} \right) = x_j^{(1)} \sum_{l=1}^{k} \left(\frac{1}{l} - \frac{1}{l+1} \right) = x_j^{(1)} \sum_{l=1}^{k} \frac{1}{l(l+1)} = x_j^{(1)} \left(1 - \frac{1}{k+1} \right),$$

während sich alle anderen paarweise wegheben. Substituieren wir diesen Ausdruck in obiger Formel, vereinfacht sich die Matrixabbildung weiter,

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} - \begin{pmatrix} y_1^{(k+1)} \\ \vdots \\ y_i^{(k+1)} \\ \vdots \\ y_n^{(k+1)} \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \left[\begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} - \frac{1}{k+1} \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} \right].$$

Nach Kürzen der Trainingsmatrix erhalten wir schließlich den Iterationsalgorithmus

$$\begin{pmatrix} y_1^{(k+1)} \\ \vdots \\ y_i^{(k+1)} \\ \vdots \\ y_n^{(k+1)} \end{pmatrix} = \frac{1}{k+1} \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix}.$$

Im Grenzfall $k \to \infty$ konvergiert der Ist-Vektor gegen null:

$$\lim_{k\to\infty} \begin{pmatrix} y_1^{(k)} \\ \vdots \\ y_i^{(k)} \\ \vdots \\ y_n^{(k)} \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} \lim_{k\to\infty} \frac{1}{k} = \mathbf{0}.$$

Damit erreicht man nach unendlich vielen Iterationen den Ziel- oder Soll-Vektor

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} = \lim_{k\to\infty}\left[\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} - \begin{pmatrix} y_1^{(k)} \\ \vdots \\ y_i^{(k)} \\ \vdots \\ y_n^{(k)} \end{pmatrix}\right]$$

als Differenz zum Grenzwert des Ist-Vektors. Letzterem entspricht das homogene lineare Gleichungssystem

$$\begin{pmatrix} y_1^{(k)} \\ \vdots \\ y_i^{(k)} \\ \vdots \\ y_n^{(k)} \end{pmatrix} = \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} x_1^{(k)} \\ \vdots \\ x_j^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

welches in Komponentenschreibweise lautet:

$$\begin{aligned} w_{11}x_1^{(k)} + w_{1j}x_j^{(k)} + w_{1n}x_n^{(k)} &= 0, \\ &\vdots \\ w_{i1}x_1^{(k)} + w_{ij}x_j^{(k)} + w_{in}x_n^{(k)} &= 0, \\ &\vdots \\ w_{m1}x_1^{(k)} + w_{mj}x_j^{(k)} + w_{mn}x_n^{(k)} &= 0. \end{aligned}$$

Homogene Gleichungssysteme sind stets lösbar. Als Beispiel seien die Lösungen einer quadratischen $3\times 3-$Matrix angeführt:

$$\begin{aligned} w_{11}x_1^{(k)} + w_{12}x_2^{(k)} + w_{13}x_3^{(k)} &= 0, \\ w_{21}x_1^{(k)} + w_{22}x_2^{(k)} + w_{23}x_3^{(k)} &= 0, \\ w_{31}x_1^{(k)} + w_{32}x_2^{(k)} + w_{33}x_3^{(k)} &= 0. \end{aligned}$$

Zunächst lösen wir die dritte Gleichung nach $x_3^{(k)}$ auf,

$$x_3^{(k)} = -\frac{w_{31}}{w_{33}}x_1^{(k)} - \frac{w_{32}}{w_{33}}x_2^{(k)},$$

und setzen diesen Ausdruck in die beiden anderen Gleichungen ein, womit wir unser ursprüngliches lineares Gleichungssystem auf ein System von zwei Gleichungen mit zwei Unbekannten reduziert haben:

$$\begin{aligned} \left(w_{11} - \frac{w_{13}w_{31}}{w_{33}}\right)x_1^{(k)} + \left(w_{12} - \frac{w_{13}w_{32}}{w_{33}}\right)x_2^{(k)} &= 0, \\ \left(w_{21} - \frac{w_{23}w_{31}}{w_{33}}\right)x_1^{(k)} + \left(w_{22} - \frac{w_{23}w_{32}}{w_{33}}\right)x_2^{(k)} &= 0. \end{aligned}$$

Sodann lösen wir die zweite Gleichung nach $x_2^{(k)}$ auf, d.h.

$$x_2^{(k)} = -\frac{w_{21}w_{33} - w_{23}w_{31}}{w_{22}w_{33} - w_{23}w_{32}} x_1^{(k)},$$

und setzen diese Relation in die dritte ein, woraus sich für $x_3^{(k)}$ der Ausdruck

$$x_3^{(k)} = \frac{w_{21}w_{32} - w_{22}w_{31}}{w_{22}w_{33} - w_{23}w_{32}} x_1^{(k)}$$

ergibt. Mittels der Definition

$$x_1^{(k)} \equiv \frac{x_1^{(1)}}{k},$$

lauten dann die drei Lösungen

$$\begin{aligned}
x_1^{(k)} &= \frac{1}{k} x_1^{(1)}, \\
x_2^{(k)} &= -\frac{1}{k} \frac{w_{21}w_{33} - w_{23}w_{31}}{w_{22}w_{33} - w_{23}w_{32}} x_1^{(1)}, \\
x_3^{(k)} &= \frac{1}{k} \frac{w_{21}w_{32} - w_{22}w_{31}}{w_{22}w_{33} - w_{23}w_{32}} x_1^{(1)}.
\end{aligned}$$

Wir brauchen also nur einen beliebigen Anfangswert der gegebenen Aufgabenstellung herausgreifen, um den vollständigen Satz von Lösungen iterativ berechnen zu können. Es müssen stets alle drei Eingangsneuronen gegen Null konvergieren, damit wir daraus den Ist-Vektor

$$\begin{pmatrix} y_1^{(k)} \\ \vdots \\ y_i^{(k)} \\ \vdots \\ y_n^{(1)} \end{pmatrix}$$

der Ausgangsneuronen erhalten. Dieser muß entsprechend gegen Null konvergieren, um den Zielvektor beliebig genau approximieren zu können. Erst wenn die Abweichungen

$$\begin{pmatrix} x_1^{(k+1)} \\ \vdots \\ x_j^{(k+1)} \\ \vdots \\ x_n^{(k+1)} \end{pmatrix} = \sum_{l=1}^{k} \begin{pmatrix} \Delta x_1^{(l)} \\ \vdots \\ \Delta x_j^{(l)} \\ \vdots \\ \Delta x_n^{(l)} \end{pmatrix} - \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} y_1^{(k+1)} \\ \vdots \\ y_i^{(k+1)} \\ \vdots \\ y_m^{(k+1)} \end{pmatrix} = \sum_{l=1}^{k} \begin{pmatrix} \Delta y_1^{(l)} \\ \vdots \\ \Delta y_i^{(l)} \\ \vdots \\ \Delta y_m^{(l)} \end{pmatrix} - \begin{pmatrix} y_1^{(1)} \\ \vdots \\ y_i^{(1)} \\ \vdots \\ y_m^{(1)} \end{pmatrix}$$

identisch verschwinden, geht das neuronale Netz

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} y_1^{(k+1)} \\ \vdots \\ y_i^{(k+1)} \\ \vdots \\ y_n^{(k+1)} \end{pmatrix} + \begin{pmatrix} w_{11} & \cdots & w_{1j} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ij} & \cdots & w_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mj} & \cdots & w_{mn} \end{pmatrix} \left[\begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} x_1^{(k+1)} \\ \vdots \\ x_j^{(k+1)} \\ \vdots \\ x_n^{(k+1)} \end{pmatrix} \right],$$

in die Trainingsgleichung über. Das Verfahren ist identisch zur linearen Regression. Jeder $x_j^{(k)}$ -Wert erzeugt automatisch einen $y_i^{(k)}$ -Wert, und nach jeder weiteren Iteration werden die Abweichungen zu den Soll-Werten immer geringer. Die nächstfolgenden Terme werden also so lange ins Netz zurückgespeist, bis die Trainingsmatrix hinreichend gut approximiert ist. Dabei verbirgt sich die eigentliche Rekursion des Netzes hinter den Termen der „unsichtbaren" Teleskopsumme:

$$\begin{pmatrix} x_1^{(k+1)} \\ \vdots \\ x_j^{(k+1)} \\ \vdots \\ x_n^{(k+1)} \end{pmatrix} = \begin{pmatrix} x_1^{(k)} \\ \vdots \\ x_j^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix} - \begin{pmatrix} \Delta x_1^{(k)} \\ \vdots \\ \Delta x_j^{(k)} \\ \vdots \\ \Delta x_n^{(k)} \end{pmatrix}.$$

Als Startwert ist festgelegt:

$$\begin{pmatrix} \Delta x_1^{(0)} \\ \vdots \\ \Delta x_j^{(0)} \\ \vdots \\ \Delta x_n^{(0)} \end{pmatrix} = \begin{pmatrix} x_1^{(0)} \\ \vdots \\ x_j^{(0)} \\ \vdots \\ x_n^{(0)} \end{pmatrix} - \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix} \equiv \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_j^{(1)} \\ \vdots \\ x_n^{(1)} \end{pmatrix}.$$

Für genügend großes k geht die Teleskopsumme in die Teleskopreihe über:

$$\sum_{l=1}^{\infty} \Delta x_j^{(l)} = \sum_{l=1}^{\infty} \left(x_j^{(l)} - x_j^{(l+1)} \right) = x_j^{(1)}.$$

Der Anwender entscheidet durch ein Abbruchkriterium selbst, wann sein Ergebnis gut genug ist.

