

Vorgehen

Initiale Werte

Die Lernrate γ wähle man nicht zu groß und nicht zu klein. Ist sie zu klein, so geht das Lernen nur langsam voran; ist sie zu groß, so überspringt ein Lernschritt leicht das Optimum und man springt nur hin und her. Ein guter Anfangswert ist 0,1.

Für die Initialisierung der Gewichte wähle man Werte, die ein Lernen ermöglichen. Haben wir z.B. die Eingabedaten x auf einen Mittelwert von null und eine Varianz von 1 normiert, so sollten bei $n=100$ Eingaben und somit 100 Gewichten jedes Gewicht im Schnitt nur $1/100$ betragen, um eine Gesamtausgabe $z = 1$ zu erreichen und im linearen Bereich einer Quetschfunktion $S(z)$ zu bleiben.

Hinweise zu Aufgabe 1.2

- b. Ein Netz bestehend aus zwei Schichten löst dieses Problem. An dem Ergebnis aus Aufgabe 1b können Sie erkennen, warum Sie auch *mindestens* ein zweischichtiges Netz benötigen.