The Error Surface of the 2-2-1 XOR Network: Stationary Points with infinite Weights†

Ida G. Sprinkhuizen-Kuyper
Egbert J.W. Boers

Abstract

In this paper it is proved that the error surface of the two-layer XOR network with two hidden units has a number of regions with local minima. These regions of local minima occur for combinations of the weights from the inputs to the hidden nodes such that one or both hidden nodes are saturated (give output 0 or 1) for at least two patterns. However, boundary points of these regions of local minima are saddle points. From these results it can be concluded that from each finite point in weight space a strictly decreasing path exists to a point with error zero. Furthermore we give proofs that points with error zero exist, and that points with the output unit saturated are either saddle points or (local) maxima. In [10] it is proved that stationary points with finite weights are either saddle points or absolute minima.

1 Introduction

To investigate the error surfaces of XOR networks thoroughly is important, since Prechelt [5] found in his investigation of articles on learning algorithms in neural networks that *20 articles (18%) employed the "grandfather" of all neural network problems, the XOR or n-bit parity.* So there are many experimental results that can possibly better be explained with more knowledge of the error surface of these networks.

More insight in the error surfaces of a number of concrete problems can give a better insight in the specific behaviour of the learning algorithms under investigation.

In literature we found a number of results concerning the error surface of these networks, but we didn't find a complete investigation of the error

†. Technical Report 96-10, Dept. of Computer Science, Leiden University. Available as ftp://ftp.wi.leidenuniv.nl/pub/CS/TechnicalReports/1996/tr96-10.ps.gz

surfaces of the XOR networks. In [7, 8] we described our results for the error surface of the XOR network with one hidden node and connections directly from the inputs to the output node (see figure 1a). In this paper together with

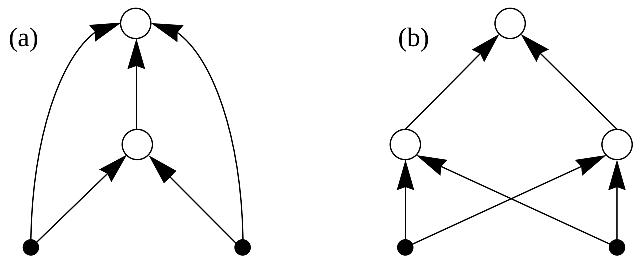

Figure 1. The simplest XOR network (a) and one with two hidden nodes (b).

[10] we give a complete investigation of the error surface of the two-layer XOR network with two hidden nodes (see figure 1b). The transfer function used is the usual sigmoid $f(x) = 1/(1+e^{-x})$. We consider the quadratic error function

$$E = \frac{1}{2}\sum_{\alpha}(O_\alpha - t_\alpha)^2$$

while in literature also the “cross-entropy”

$$L = -\sum_{\alpha}\ln\left[(O_\alpha)^{t_\alpha}(1-O_\alpha)^{1-t_\alpha}\right]$$

is used. The difference is that the terms

$$R_\alpha = (O_\alpha - t_\alpha)f'(I_\alpha)$$

which occur in the partial derivatives of the quadratic error E simplify to

$$R_\alpha' = (O_\alpha - t_\alpha)$$

for the cross-entropy L. So the analysis for the quadratic error E is more complicated than that for the cross entropy L. Especially more stationary points (points where all partial derivatives with respect to the weights are zero, so the gradient of the error is zero in these points) occur for E than for L. We mention the stationary points for which the output node is saturated (I_α is equal to plus or minus infinity) for at least one of the patterns. In subsection 4.2 we will show that these points are saddle points or local maxima. Also some more stationary points occur for finite weights (see [10]).

Since all stationary points of the error surface with the cross-entropy L form a subset of the stationary points for the quadratic error E, it is easily checked that all results obtained here also hold for the cross entropy L. So especially the regions of local minima found for the quadratic error E, which are summarized in tables 2 until 11, are also regions of local minima for the cross-entropy L.

In this paper it is proved that the error surface of the 2-2-1 XOR network has a number of regions consisting of local minima. These regions of local minima occur for combinations of the weights from the inputs to the hidden nodes such that one or both hidden nodes are saturated (give output 0 or 1, since the input is either $+\infty$ or $-\infty$) for at least two patterns. However, boundary points of these regions of local minima are saddle points. From these results it can be concluded that from each finite point in weight space a strictly decreasing path exists to a point with error zero. Furthermore we give proofs that points with error zero exist, that stationary points with finite weights are either saddle points or absolute minima, and that points with the output unit saturated are either saddle points or (local) maxima.

Relation to previous work

Blum [1] investigated the 2-2-1 XOR network with the cross-entropy as error function. He restricted the weights to be symmetrical. For the 5 remaining independent weights he proved that exact solutions exist for the XOR problem. In section 3 we give our proof that the XOR problem can be represented exactly by the 2-2-1 network. Blum's proof is more complicated since the network considered has less degrees of freedom. In the same paper [1] Blum identified a linear manifold of stationary points to be local minima. In this paper it is shown that for the network with 9 independent weights this line does not contain local minima. In [9] we proved that also with symmetric constraints on the weights no local minima exist on the manifold given by Blum. Hamey [2] also found that Blum's proof was incorrect.

Lisboa and Perantonis [4] characterize the stationary points of all two-layer XOR networks with and without connections directly from the inputs to the output unit, considering the cross-entropy as error function. They also give some local minima for the 2-2-1 network and tell that they checked that these points are indeed local minima by considering the second order partial derivatives. However, they do not give details of their proofs. Four of their local minima are found to be numerical equivalent to local minima resulting from our research (see section 5.7). The fifth point is not a local minimum (see [10]).

Hamey [2, 3] also investigated the 2-2-1 XOR network with the cross-entropy as error function. In [2] he shows that for all points with finite weights a finite non-ascending trajectory exists to a point with error zero. He concludes that the 2-2-1 XOR network has no (regional) local minima. He defines a regional local minimum as a local minimum that has to be reached by a non-ascending path from points in the neighbourhood. In [3] Hamey proves that all finite stationary points are saddle points. In this paper we prove that the 2-2-1 XOR network has local minima for infinite values of the weights from the inputs to the hidden nodes. However, this result does not contradict Hamey's results, since our definition of a local minimum in a point (finite or infinite) is that a local minimum is attained in a point $\boldsymbol{w}$ if for all points $\boldsymbol{w}'$ in a neighbourhood of $\boldsymbol{w}$ the inequality $f(\boldsymbol{w}) \leq f(\boldsymbol{w}')$ holds and we accept that points with infinite weights exist. Such an infinite local minimum can trap a learning algorithm, since a decreasing path to a point with error zero will first get closer to the infinite point and will not necessarily reach a neighbourhood where the learning algorithm can escape from the region with the local minimum value.

Contents of the paper

In section 2 the network and its parameters are introduced and also the error function is given. In section 3 it is proved that the network can represent the XOR function, as specified in section 2, exactly. In [7] we introduced the notions of stable stationary points, i.e. points which are stationary points for the error of each individual pattern, and instable stationary points. The latter points are stationary points for the total error, but not for each individual pattern. In section 4 it is proved that stable stationary points are either absolute minima with error zero or saddle points or (local) maxima, but not local minima. Both finite and infinite weights are investigated in this section. Logically the next section would contain the proofs that instable stationary points with finite weights can not be local minima. We decided to publish this part separately in [10]. In section 5 the instable points with infinite weights are treated. Here it is found that local minima only can exist if at most two patterns are learned. The resulting local minima with two patterns learned are summarized in subsection 5.3, while subsection 5.5 contains the local minima with one pattern learned, and section 5.6 contains the local minima with all patterns giving a wrong output. In subsection 5.7 some examples of local minima found in literature are shown to belong to one of the classes found earlier. The paper ends with section 6 containing some conclusions.

2 The network

In this paper we investigate the error surface of the network with two hidden units and without direct connections from the inputs to the output (see figure 2).

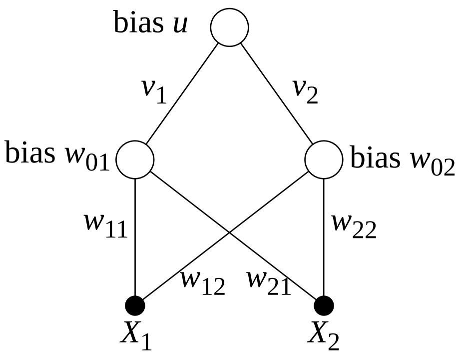

Figure 2. The XOR network with 2 hidden units

For the XOR function we assume that the patterns given in table 1 should be learned/represented by the network.

Table 1: Patterns for the XOR problem

Pattern	$\boldsymbol{X_1}$	$\boldsymbol{X_2}$	**desired output**
P_{00}	0	0	0.1
P_{01}	0	1	0.9
P_{10}	1	0	0.9
P_{11}	1	1	0.1

The input of the output unit is for the four patterns:

$$\begin{aligned} A_{00} &= u + v_1 f(w_{01}) + v_2 f(w_{02}) \\ A_{01} &= u + v_1 f(w_{01} + w_{21}) + v_2 f(w_{02} + w_{22}) \\ A_{10} &= u + v_1 f(w_{01} + w_{11}) + v_2 f(w_{02} + w_{12}) \\ A_{11} &= u + v_1 f(w_{01} + w_{11} + w_{21}) + v_2 f(w_{02} + w_{12} + w_{22}) \end{aligned} \tag{2.1}$$

So the four patterns result in output values equal to $f(A_{00})$, $f(A_{01})$, $f(A_{10})$ and $f(A_{11})$, respectively.

The mean square error is equal to:

$$E = \frac{1}{2}(f(A_{00}) - 0.1)^2 + \frac{1}{2}(f(A_{01}) - 0.9)^2 + \frac{1}{2}(f(A_{10}) - 0.9)^2 + \frac{1}{2}(f(A_{11}) - 0.1)^2 \quad (2.2)$$

The weight space has a number of symmetries for this problem, which we will exploit in order to reduce the number of different cases that have to be investigated. Especially, we will consider the following four transformations of the weight space:

Transformation 2.1: (interchanging the inputs using the symmetry of the training patterns with respect to the inputs)

$w_{11}' = w_{21}$, $w_{21}' = w_{11}$, $w_{12}' = w_{22}$, $w_{22}' = w_{12}$, other weights equal.

Transformation 2.2: (interchanging P_{00} and P_{11}, and P_{01} and P_{10})

$w_{01}' = w_{01} + w_{11} + w_{21}$, $w_{02}' = w_{02} + w_{12} + w_{22}$, $w_{11}' = -w_{11}$, $w_{21}' = -w_{21}$, $w_{12}' = -w_{12}$, $w_{22}' = -w_{22}$, other weights equal.

Transformation 2.3: (using that $f(x) = 1 - f(-x)$, and interchanging patterns with desired output 0.1 and those with desired output 0.9)

$u' = -u - v_1 - v_2$, $w_{01}' = -w_{01} - w_{21}$, $w_{02}' = -w_{02} - w_{22}$, $w_{11}' = -w_{11}$, $w_{12}' = -w_{12}$, other weights equal.

Transformation 2.4: (mirroring the network)

$v_1' = v_2$, $v_2' = v_1$, $w_{i1}' = w_{i2}$, $w_{i2}' = w_{i1}$, $i \in \{0, 1, 2\}$, other weights equal.

3 Representation

In this section we prove that the XOR function can be represented exactly by the network with two hidden units given in figure 2.

The XOR function is exactly represented by the network if the weights u, v_1, v_2, w_{01}, w_{11}, w_{21}, w_{02}, w_{12} and w_{22} are such that the following equations hold:

$$\begin{aligned} f(A_{00}) &= 0.1 \\ f(A_{01}) &= 0.9 \\ f(A_{10}) &= 0.9 \\ f(A_{11}) &= 0.1 \end{aligned} \quad (3.1)$$

Application of the inverse of f on both sides of (3.1), using (2.1), leads to:

$$
\begin{aligned}
& u + v_1 f(w_{01}) + v_2 f(w_{02}) = f^{-1}(0.1) \\
& u + v_1 f(w_{01} + w_{21}) + v_2 f(w_{02} + w_{22}) = f^{-1}(0.9) \\
& u + v_1 f(w_{01} + w_{11}) + v_2 f(w_{02} + w_{12}) = f^{-1}(0.9) \\
& u + v_1 f(w_{01} + w_{11} + w_{21}) + v_2 f(w_{02} + w_{12} + w_{22}) = f^{-1}(0.1)
\end{aligned}
\tag{3.2}
$$

From these equations it follows that:

$$
\begin{aligned}
-u &= v_1 f(w_{01}) + v_2 f(w_{02}) - f^{-1}(0.1) \\
&= v_1 f(w_{01} + w_{21}) + v_2 f(w_{02} + w_{22}) - f^{-1}(0.9) \\
&= v_1 f(w_{01} + w_{11}) + v_2 f(w_{02} + w_{12}) - f^{-1}(0.9) \\
&= v_1 f(w_{01} + w_{11} + w_{21}) + v_2 f(w_{02} + w_{12} + w_{22}) - f^{-1}(0.1)
\end{aligned}
\tag{3.3}
$$

leading to the following three equations for the weights with exception of u:

$$
\begin{aligned}
& v_1 (f(w_{01}) - f(w_{01} + w_{11} + w_{21})) + \\
& \quad + v_2 (f(w_{02}) - f(w_{02} + w_{12} + w_{22})) = 0 \\
& v_1 (f(w_{01} + w_{21}) - f(w_{01} + w_{11})) + \\
& \quad + v_2 (f(w_{02} + w_{22}) - f(w_{02} + w_{12})) = 0 \\
& v_1 (f(w_{01}) - f(w_{01} + w_{21})) + \\
& \quad + v_2 (f(w_{02}) - f(w_{02} + w_{22})) = -2f^{-1}(0.9)
\end{aligned}
\tag{3.4}
$$

The set of equations (3.4) is a set of three linear equations in the two variables v_1 and v_2.

Let us consider points with $w_{11} = w_{21} \neq 0$ and $w_{12} = w_{22} \neq 0$. For these points the second equation of (3.4) is identically zero, and for almost all values of the weights w_{ij} the first and third equation will be linearly independent, so v_1 and v_2 are determined by these equations and u follows from (3.3). Thus at least one region in weight space exists where the patterns P_{00}, P_{01}, P_{10} and P_{11} are represented exactly. The dimension of this region is at least 4 and probably 5, since the condition that the first two equations of (3.4) have to be linearly dependent results in one restriction on the 6 weights w_{ij}.

4 Stable stationary points

Let us introduce

$$
\begin{aligned}
R_{00} &= (f(A_{00}) - 0.1) f'(A_{00}) \\
R_{01} &= (f(A_{01}) - 0.9) f'(A_{01}) \\
R_{10} &= (f(A_{10}) - 0.9) f'(A_{10}) \\
R_{11} &= (f(A_{11}) - 0.1) f'(A_{11})
\end{aligned}
\tag{4.1}
$$

Stable stationary points are obtained when the gradient of the error is zero for each of the four patterns separately, thus if

$$R_{00} = R_{01} = R_{10} = R_{11} = 0 \tag{4.2}$$

The cases with all weights finite and one or more weights infinite are considered here.

4.1 Finite weights

If all weights are finite the only points with all R_{ij}'s equal to zero are the points satisfying equations (3.1) and thus all patterns are learned exactly and the error is zero.

4.2 Output 0 or 1

We have to investigate those points where one or more of the terms A_{ij} are infinite and the other terms result in the desired output.

Let us consider points in weight space in the neighbourhood of such a stable stationary point. We will show that it is not possible that an infinite value of A_{00} corresponds to a local minimum. The other cases A_{01}, A_{10} and/or A_{11} tending to plus or minus infinity are treated by transformations of the weight space.

First let us try to keep A_{01}, A_{10} and A_{11} constant. By (2.1) the effect of small variations of v_1, w_{01}, w_{11} and w_{21} on A_{00}, A_{01}, A_{10} and A_{11} is:

$$\begin{aligned}
\Delta A_{00} &= f(w_{01})\,\Delta v_1 + v_1 f'(w_{01})\,\Delta w_{01} \\
\Delta A_{01} &= f(w_{01}+w_{21})\,\Delta v_1 + v_1 f'(w_{01}+w_{21})\,(\Delta w_{01}+\Delta w_{21}) \\
\Delta A_{10} &= f(w_{01}+w_{11})\,\Delta v_1 + v_1 f'(w_{01}+w_{11})\,(\Delta w_{01}+\Delta w_{11}) \\
\Delta A_{11} &= f(w_{01}+w_{11}+w_{21})\,\Delta v_1 + \\
&\qquad + v_1 f'(w_{01}+w_{11}+w_{21})\,(\Delta w_{01}+\Delta w_{11}+\Delta w_{21})
\end{aligned} \tag{4.3}$$

Solving Δw_{01}, Δw_{10} and Δw_{11} from the equations for $\Delta A_{01} = \Delta A_{10} = \Delta A_{11} = 0$, results in:

$$\begin{aligned}
&\Delta A_{00} = \\
&f'(w_{01})\,\Delta v_1\left(\frac{f(w_{01})}{f'(w_{01})} - \frac{f(w_{01}+w_{21})}{f'(w_{01}+w_{21})} - \frac{f(w_{01}+w_{11})}{f'(w_{01}+w_{11})} + \right. \\
&\qquad \left. + \frac{f(w_{01}+w_{11}+w_{21})}{f'(w_{01}+w_{11}+w_{21})}\right) = \\
&-f'(w_{01})\,\Delta v_1 e^{w_{01}}\left(1-e^{w_{11}}\right)\left(1-e^{w_{21}}\right)
\end{aligned}$$

Thus if $w_{11} \neq 0$ and $w_{21} \neq 0$ then it is possible to vary the weights v_1, w_{01}, w_{11} and w_{21} such that A_{00} becomes closer to the desired value, while A_{01},

A_{10} and A_{11} remain constant. The effect is that the error decreases when the weights are altered in a direction away from the stationary point. So if $w_{11} \neq 0$ and $w_{21} \neq 0$ then $A_{00} \to \pm\infty$ will never result in a local minimum.

Similarly it is proved that for $w_{12} \neq 0$ and $w_{22} \neq 0$ it is possible to vary the weights v_2, w_{02}, w_{12} and w_{22} such that A_{00} becomes closer to the desired value, while A_{01}, A_{10} and A_{11} remain constant, and also in this case $A_{00} \to \pm\infty$ will not yield a local minimum.

The cases that have to be investigated further are the cases where both $w_{11} = 0$ or $w_{21} = 0$, and $w_{12} = 0$ or $w_{22} = 0$. These cases lead to the four cases:

- $w_{11} = 0$ and $w_{12} = 0$,
- $w_{11} = 0$ and $w_{22} = 0$,
- $w_{21} = 0$ and $w_{12} = 0$,
- $w_{21} = 0$ and $w_{22} = 0$.

In the first case ($w_{11} = 0$ and $w_{12} = 0$) equation (2.1) becomes:

$$\begin{aligned} A_{00} &= A_{10} = u + v_1 f(w_{01}) + v_2 f(w_{02}) \\ A_{01} &= A_{11} = u + v_1 f(w_{01} + w_{21}) + v_2 f(w_{02} + w_{22}) \end{aligned} \tag{4.4}$$

Stable stationary points can only be found in this case if all four terms A_{ij} are infinite. If v_1 is varied a little bit such that A_{00} and A_{10} are moving away from infinity and w_{21} is varied correspondingly such that A_{01} and A_{11} remain constant, then the total error is decreased and thus $A_{00} \to \pm\infty$ will not result in a local minimum in this case.

In the second case ($w_{11} = 0$ and $w_{22} = 0$) equation (2.1) becomes:

$$\begin{aligned} A_{00} &= u + v_1 f(w_{01}) + v_2 f(w_{02}) \\ A_{01} &= u + v_1 f(w_{01} + w_{21}) + v_2 f(w_{02}) \\ A_{10} &= u + v_1 f(w_{01}) + v_2 f(w_{02} + w_{12}) \\ A_{11} &= u + v_1 f(w_{01} + w_{21}) + v_2 f(w_{02} + w_{12}) \end{aligned} \tag{4.5}$$

From equations (4.5) it follows that:

$$A_{00} - A_{01} - A_{10} + A_{11} = 0 \tag{4.6}$$

So if $A_{00} \to \pm\infty$ then at least one of the other terms A_{01}, A_{10} or A_{11} will also approach $\pm\infty$. Because of equation (4.6) one of the following possibilities will occur: A_{00} and A_{11} have opposite sign or A_{00} has the same sign as A_{01} (or A_{10}) where the concerning terms are approaching $\pm\infty$.

If A_{00} and A_{11} have opposite sign we can vary v_1, v_2, w_{21} and w_{12} such that A_{01} and A_{10} remain constant, resulting in:

$$
\begin{aligned}
\Delta A_{00} &= f(w_{01})\,\Delta v_1 + f(w_{02})\,\Delta v_2 \\
\Delta A_{01} &= f(w_{01}+w_{21})\,\Delta v_1 + f(w_{02})\,\Delta v_2 + v_1 f'(w_{01}+w_{21})\,\Delta w_{21} \\
\Delta A_{10} &= f(w_{01})\,\Delta v_1 + f(w_{02}+w_{12})\,\Delta v_2 + v_2 f'(w_{02}+w_{12})\,\Delta w_{12} \\
\Delta A_{11} &= f(w_{01}+w_{21})\,\Delta v_1 + f(w_{02}+w_{12})\,\Delta v_2 + \\
&\quad v_1 f'(w_{01}+w_{21})\,\Delta w_{21} + v_2 f'(w_{02}+w_{12})\,\Delta w_{12}
\end{aligned}
\tag{4.7}
$$

Using the equations for $\Delta A_{01} = \Delta A_{10} = 0$ leads to:

$$\Delta A_{11} = -f(w_{01})\,\Delta v_1 - f(w_{02})\,\Delta v_2 = -\Delta A_{00} \tag{4.8}$$

So it is possible to change the values of v_1, v_2, w_{21} and w_{12} such that A_{01} and A_{10} remain constant, and both A_{00} and A_{11} move away from infinity, thus decreasing the error. So this case will not result in a local minimum.

If A_{00} and A_{01} approach infinity with the same sign, we find analogously to the previous case that varying v_1, v_2, w_{21} and w_{12} such that A_{10} and A_{11} remain constant leads to:

$$\Delta A_{01} = f(w_{01})\,\Delta v_1 + f(w_{02})\,\Delta v_2 = \Delta A_{00} \tag{4.9}$$

So both A_{01} and A_{00} can be moved away from infinity, resulting in a decreasing error. So also this case will not result in a local minimum. From symmetry it is clear that also the case where A_{00} and A_{10} approach infinity, with the same sign, will not lead to a local minimum.

The third case ($w_{21} = 0$ and $w_{12} = 0$) and the fourth case ($w_{21} = 0$ and $w_{22} = 0$) are equivalent to the second and the first case, respectively.

So we can conclude that no local minima will be found if A_{00} approaches infinity. From transformations 2.1, 2.2 and 2.3 it can be concluded that also no local minima will be found if one of the other terms A_{01}, A_{10} and/or A_{11} approaches infinity.

In the proofs we did not really use the fact that we were considering stable stationary points. So we can extend these results immediately to unstable stationary points.

Conclusion 4.1 *Stationary points with the output for at least one of the patterns equal to 0 or 1 cannot be local minima.*

Conclusion 4.2 *The only stable stationary points that behave like a minimum are the points with all four patterns exactly learned, so in those points the absolute minimum with error zero is found. All other stable stationary points, both with finite weights and with infinite weights, are either saddle points or (local) maxima.*

5 Instable stationary points with infinite weights

For instable stationary points the gradient of the total error is equal to zero, while the gradient of the error of at least one of the patterns is unequal to zero. In [10] it is proved that all instable stationary points with finite weights are saddle points. In this section we will show that local minima exist when one or more of the weights to the hidden units are equal to $\pm\infty$.

We will divide the problem in classes with respect to the number of terms R_{ij} equal to zero. In section 4.2 we showed that no local minima occur if one of the terms R_{ij} is equal to zero because of the input of the output node being infinite. So we will study here the cases that a number of the patterns is learned, resulting in the corresponding terms R_{ij} being zero.

5.1 Three of the patterns P_{ij} are learned

Since the partial derivative of E with respect to u is equal to $R_{00} + R_{01} + R_{10} + R_{11}$ it is clear that if three of the four terms R_{ij} are zero, the fourth has to be zero too. So in that case we are in a stable stationary point.

5.2 Two of the patterns P_{ij} are learned

There are 6 possibilities to have two of the four patterns P_{ij} learned, but essentially there are two different cases:

- P_{00} and P_{01} are learned and
- P_{00} and P_{11} are learned.

The other possibilities can be obtained from these two cases by transformation of the weights. So we will consider these two cases first.

5.2.1 The patterns P_{00} and P_{01} are learned

In this case $R_{00} = R_{01} = 0$ holds and all first order partial derivatives of E with respect to the weights are equal to zero if in addition:

$$\begin{aligned}
&R_{10} = -R_{11} \neq 0 \\
&f(w_{01} + w_{11}) = f(w_{01} + w_{11} + w_{21}) \\
&f(w_{02} + w_{12}) = f(w_{02} + w_{12} + w_{22}) \\
&v_1 f'(w_{01} + w_{11}) = 0 \\
&v_1 f'(w_{01} + w_{11} + w_{21}) = 0 \\
&v_2 f'(w_{02} + w_{12}) = 0 \\
&v_2 f'(w_{02} + w_{12} + w_{22}) = 0
\end{aligned} \tag{5.1}$$

We are considering the cases where at least one of the weights is infinite. If $w_{01} + w_{11}$ is finite, then necessarily $w_{21} = 0$ and $v_1 = 0$ has to hold. In this case we find (see also the proofs in [10]) that all partial derivatives of the error with respect to combinations of w_{11} and/or w_{21} are equal to zero. But taking also the partial derivative with respect to v_1 results in:

$$\left.\frac{\partial^3 E}{\partial v_1 \partial w_{11} \partial w_{21}}\right|_{\text{stat.pnt}} = R_{11} f''(w_{01} + w_{11}) \neq 0 \text{ if } w_{01} + w_{11} \neq 0$$

$$\left.\frac{\partial^4 E}{\partial v_1 \partial w_{11}^2 \partial w_{21}}\right|_{\text{stat.pnt}} = R_{11} f'''(w_{01} + w_{11}) \neq 0 \text{ if } w_{01} + w_{11} = 0$$

and thus the points with $w_{01} + w_{11}$ finite are saddle points (see theorems A2, A3 and A4 in [7]). Analogously points with $w_{01} + w_{11} + w_{21}$, $w_{02} + w_{12}$ or $w_{02} + w_{12} + w_{22}$ finite are saddle points. So the remaining points that have to be investigated can be divided into the following four cases:

- $w_{01} + w_{11} = w_{01} + w_{11} + w_{12} = w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = \infty$
- $w_{01} + w_{11} = w_{01} + w_{11} + w_{12} = \infty, w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = -\infty$
- $w_{01} + w_{11} = w_{01} + w_{11} + w_{12} = -\infty, w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = \infty$
- $w_{01} + w_{11} = w_{01} + w_{11} + w_{12} = w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = -\infty$

We will consider these cases in the following.

Case 5.2.1.1: $w_{01}+w_{11} = w_{01}+w_{11}+w_{21} = w_{02}+w_{12} = w_{02}+w_{12}+w_{22} = \infty$

In this case equations (2.1) give:

$$\begin{aligned} A_{00} &= u + v_1 f(w_{01}) + v_2 f(w_{02}) = f^{-1}(0.1) \\ A_{01} &= u + v_1 f(w_{01} + w_{21}) + v_2 f(w_{02} + w_{22}) = f^{-1}(0.9) \\ A_{10} &= A_{11} = u + v_1 + v_2 = 0 \end{aligned} \tag{5.2}$$

The corresponding error level is 0.16. Eliminating u results in:

$$\begin{aligned} A_{00} &= -v_1 f(-w_{01}) - v_2 f(-w_{02}) = f^{-1}(0.1) \approx -2.197 \\ A_{01} &= -v_1 f(-w_{01} - w_{21}) - v_2 f(-w_{02} - w_{22}) = f^{-1}(0.9) \approx 2.197 \end{aligned} \tag{5.3}$$

Since $f(x)$ is positive and A_{00} and A_{01} have opposite sign, v_1 and v_2 will have opposite sign and will not be equal to zero in this case. Since $f(x) \in [0, 1]$ it follows that either

- $v_1 \geq f^{-1}(0.9)$ and $v_2 \leq f^{-1}(0.1)$ or
- $v_1 \leq f^{-1}(0.1)$ and $v_2 \geq f^{-1}(0.9)$.

Equations (5.3) have a solution for v_1 and v_2 if the following inequality holds:

$$f(-w_{01})f(-w_{02}-w_{22}) \neq f(-w_{02})f(-w_{01}-w_{21}) \tag{5.4}$$

In order to investigate points with $w_{01}+w_{11} = w_{02}+w_{12} = \infty$, we use the substitution:

$p_1 = e^{-w_{01}-w_{11}}$ and $p_2 = e^{-w_{02}-w_{12}}$

The stationary points considered correspond with $p_1 = p_2 = 0$ and we are interested in the behaviour of the error surface for $p_1 \downarrow 0$ and $p_2 \downarrow 0$. Computation of the partial derivatives of the error E with respect to p_1 and p_2 for p_1 and p_2 equal to zero, choosing w_{01}, w_{21}, w_{02} and w_{22} independent of p_1 and p_2, results in:

$$\left.\frac{\partial E}{\partial p_i}\right|_{p_i = 0} = 0.4f'(0)\, v_i\,(1-e^{-w_{2i}}),\ i \in \{1, 2\} \tag{5.5}$$

Since both p_1 and p_2 are greater than or equal to zero it is clear that if one of the derivatives in equation (5.5) is negative, then the error will decrease if p_1 or p_2 moves away from zero (and $w_{01}+w_{11}$ or $w_{02}+w_{12}$ moves away from infinity, correspondingly). Thus then the stationary point is not a local minimum. The sign of the derivatives in (5.5) is determined by the signs of v_1, v_2, w_{21} and w_{22}. So we can conclude:

Conclusion 5.1 *Stationary points with the patterns P_{00} and P_{01} learned, $u+v_1+v_2 = 0$, $w_{01}+w_{11} = w_{01}+w_{11}+w_{21} = \infty$ and $w_{02}+w_{12} = w_{02}+w_{12}+w_{22} = \infty$, are not local minima if*

- $v_1 \geq f^{-1}(0.9)$ *and* $w_{21} < 0$, *or*
- $v_1 \leq f^{-1}(0.1)$ *and* $w_{21} > 0$, *or*
- $v_2 \geq f^{-1}(0.9)$ *and* $w_{22} < 0$, *or*
- $v_2 \leq f^{-1}(0.1)$ *and* $w_{22} > 0$.

If both derivatives in equation (5.5) are positive, increasing p_1 and/or p_2 will lead to an increase of the error. When p_1 and p_2 are equal to zero the error can only be decreased by altering $u+v_1+v_2$ (see (5.2)), such that the error corresponding to A_{10} and A_{11} decreases, and altering the other weights in order to keep the error corresponding to A_{00} and A_{01} equal to zero. But the error corresponding to A_{10} and A_{11} as a function of $x = u+v_1+v_2$ is equal to:

$$E = \frac{1}{2}(f(x)-0.9)^2 + \frac{1}{2}(f(x)-0.1)^2 \tag{5.6}$$

which attains a minimum for $x = 0$.

So each variation of $u+v_1+v_2$ will increase the error with respect to A_{10} and A_{11}. So here a local minimum is found!

The dimension of the region in which this minimum value is attained follows from (5.2), (5.3) and (5.4): if w_{01}, w_{02}, w_{21} and w_{22} are chosen such that the inequality (5.4) holds, then u, v_1 and v_2 are determined by (5.2) and (5.3). So the dimension of this region of local minima is 4.

Conclusion 5.2 *A 4-dimensional region of local minima with error 0.16 is found if the patterns P_{00} and P_{01} are learned, $u + v_1 + v_2 = 0$, $w_{01} + w_{11} = w_{01} + w_{11} + w_{21} = \infty$ and $w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = \infty$, and*

- $v_1 \geq f^{-1}(0.9)$, $v_2 \leq f^{-1}(0.1)$, $w_{21} > 0$ *and* $w_{22} < 0$, *or*
- $v_1 \leq f^{-1}(0.1)$, $v_2 \geq f^{-1}(0.9)$, $w_{21} < 0$ *and* $w_{22} > 0$.

Finally consider the case that one or both of the derivatives in (5.5) are equal to zero, i.e. $w_{21} = 0$ and/or $w_{22} = 0$. These points are boundary points of the region with saddle points given in conclusion 5.1, so they are saddle points too. In the following we will show that points with $w_{21} = 0$ are saddle points by considering the partial derivatives. If $w_{21} = 0$, then we find from (5.3) that

$$-v_2 \{ f(-w_{02} - w_{22}) - f(-w_{02}) \} = 2f^{-1}(0.9) > 0 \tag{5.7}$$

implying that either

- $v_2 > 0$ and $w_{22} > 0$, or
- $v_2 < 0$ and $w_{22} < 0$.

So conclusion 5.1 can not be applied to conclude that these points are saddle points. Let us consider the second order derivatives of the error with respect to u and p_1. Calculation results in:

$$\left.\frac{\partial^2 E}{\partial u^2}\right|_{p_1 = w_{21} = 0} = 2 \{ f'(f^{-1}(0.1)) \}^2 \tag{5.8}$$

$$\left.\frac{\partial^2 E}{\partial u \partial p_1}\right|_{p_1 = w_{21} = 0} = -2 \{ f'(0) \}^2 v_1 \tag{5.9}$$

$$\left.\frac{\partial^2 E}{\partial p_1^2}\right|_{p_1 = w_{21} = 0} = 2 \{ f'(0) \}^2 v_1^2 \tag{5.10}$$

leading to the following terms in the second order part of the Taylor expansion of the error:

$$\begin{aligned} \Delta E \approx\ & 2 \{ f'(0) \}^2 (v_1 \Delta p_1 - \Delta u)^2 + \\ & -2 [\{ f'(0) \}^2 - \{ f'(f^{-1}(0.1)) \}^2] (\Delta u)^2 \end{aligned}$$

So if Δu is chosen equal to $v_1 \Delta p_1$ it is clear that $\Delta E < 0$ and thus no local minimum exists in the points with $w_{21} = 0$.

So we can sharpen conclusion 5.1 to:

Conclusion 5.3 *Stationary points with the patterns P_{00} and P_{01} learned, $u + v_1 + v_2 = 0$, $w_{01} + w_{11} = w_{01} + w_{11} + w_{21} = \infty$ and $w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = \infty$ are not local minima if*

- $v_1 \geq f^{-1}(0.9)$ *and* $w_{21} \leq 0$, *or*
- $v_1 \leq f^{-1}(0.1)$ *and* $w_{21} \geq 0$, *or*
- $v_2 \geq f^{-1}(0.9)$ *and* $w_{22} \leq 0$, *or*
- $v_2 \leq f^{-1}(0.1)$ *and* $w_{22} \geq 0$.

Now consider the regions where a local minimum is attained. In these regions equations (5.2), (5.3) and (5.4) hold and

- $v_1 \geq f^{-1}(0.9)$, $v_2 \leq f^{-1}(0.1)$, $w_{21} > 0$ *and* $w_{22} < 0$, *or*
- $v_1 \leq f^{-1}(0.1)$, $v_2 \geq f^{-1}(0.9)$, $w_{21} < 0$ *and* $w_{22} > 0$.

It is possible to alter w_{01}, w_{21}, w_{02} and w_{22} smoothly such that inequality (5.4) keeps holding until $w_{21} = 0$. Thus it is clear that the points with (5.2), (5.3) and (5.4) and $w_{21} = 0$ are boundary points of the region where a local minimum is attained.

So the local minima of conclusion 5.2 form a kind of rain gutter where the water can escape in some sink at the end.

So if with on-line learning a movement is caused in the neighbourhood of the region of local minima such that w_{21} or w_{22} is tending to zero, at last a point is reached that is not a local minimum and the learning algorithm escapes at the end. However, this can take a lot of time.

Case 5.2.1.2: $w_{01}+w_{11} = w_{01}+w_{11}+w_{21} = \infty$, $w_{02}+w_{12} = w_{02}+w_{12}+w_{22} = -\infty$

In this case we have:

$$\begin{aligned} A_{00} &= u + v_1 f(w_{01}) + v_2 f(w_{02}) = f^{-1}(0.1) \\ A_{01} &= u + v_1 f(w_{01} + w_{21}) + v_2 f(w_{02} + w_{22}) = f^{-1}(0.9) \\ A_{10} &= A_{11} = u + v_1 = 0 \end{aligned} \tag{5.11}$$

The corresponding error level is again 0.16. Elimination of u results in:

$$\begin{aligned} A_{00} &= -v_1 f(-w_{01}) + v_2 f(w_{02}) = f^{-1}(0.1) \approx -2.197 \\ A_{01} &= -v_1 f(-w_{01} - w_{21}) + v_2 f(w_{02} + w_{22}) = f^{-1}(0.9) \approx 2.197 \end{aligned} \tag{5.12}$$

So clearly v_1 and v_2 will have the same sign, and thus either

- $v_1 \geq f^{-1}(0.9)$ and $v_2 \geq f^{-1}(0.9)$ or

- $v_1 \le f^{-1}(0.1)$ and $v_2 \le f^{-1}(0.1)$.

Consider the behaviour of the error when $w_{01} + w_{11}$ and $w_{02} + w_{12}$ are in the neighbourhood of plus and minus infinity, respectively, by substituting

$$p_1 = e^{-w_{01} - w_{11}} \text{ and } q_2 = e^{w_{02} + w_{12}} \tag{5.13}$$

Calculation of the partial derivatives of the error E with respect to p_1 and q_2 for p_1 and q_2 equal to zero results in:

$$\begin{aligned} \left.\frac{\partial E}{\partial p_1}\right|_{p_1 = 0} &= 0.4f'(0)\, v_1\, (1 - e^{-w_{21}}) \\ \left.\frac{\partial E}{\partial q_2}\right|_{q_2 = 0} &= -0.4f'(0)\, v_2\, (1 - e^{w_{22}}) \end{aligned} \tag{5.14}$$

Similarly to case 1, it is clear that if both derivatives in equation (5.14) are positive, a local minimum will be found.

Conclusion 5.4 *A 4-dimensional region of local minima with error 0.16 is found if the patterns P_{00} and P_{01} are learned, $u + v_1 = 0$, $w_{01} + w_{11} = w_{01} + w_{11} + w_{21} = \infty$ and $w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = -\infty$, and*

- $v_1 \ge f^{-1}(0.9)$, $v_2 \ge f^{-1}(0.9)$, $w_{21} > 0$ *and* $w_{22} > 0$, *or*
- $v_1 \le f^{-1}(0.1)$, $v_2 \le f^{-1}(0.1)$, $w_{21} < 0$ *and* $w_{22} < 0$.

Similarly to case 1 the boundary points of this region with $w_{21} = 0$ or $w_{22} = 0$ are saddle points.

Case 5.2.1.3: $\boldsymbol{w_{01}+w_{11} = w_{01}+w_{11}+w_{21} = -\infty, w_{02}+w_{12} = w_{02}+w_{12}+w_{22} = \infty}$

This case is equivalent to the case considered before by interchanging the two hidden units. This leads with respect to the local minima to the conclusion:

Conclusion 5.5 *A 4-dimensional region of local minima with error 0.16 is found if the patterns P_{00} and P_{01} are learned, $u + v_2 = 0$, $w_{01} + w_{11} = w_{01} + w_{11} + w_{21} = -\infty$ and $w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = \infty$, and*

- $v_1 \ge f^{-1}(0.9)$, $v_2 \ge f^{-1}(0.9)$, $w_{21} > 0$ *and* $w_{22} > 0$, *or*
- $v_1 \le f^{-1}(0.1)$, $v_2 \le f^{-1}(0.1)$, $w_{21} < 0$ *and* $w_{22} < 0$.

Case 5.2.1.4: $\boldsymbol{w_{01}+w_{11} = w_{01}+w_{11}+w_{21} = w_{02}+w_{12} = w_{02}+w_{12}+w_{22} = -\infty}$

All calculations are similar to the calculations made earlier, leading to the following local minima:

Conclusion 5.6 *A 4-dimensional region of local minima with error 0.16 is found if the patterns* P_{00} *and* P_{01} *are learned,* $u = 0$, $w_{01} + w_{11} = w_{01} + w_{11}w_{21} = -\infty$ *and* $w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = -\infty$, *and*

- $v_1 \geq f^{-1}(0.9)$, $v_2 \leq f^{-1}(0.1)$, $w_{21} > 0$ *and* $w_{22} < 0$, *or*
- $v_1 \leq f^{-1}(0.1)$, $v_2 \geq f^{-1}(0.9)$, $w_{21} < 0$ *and* $w_{22} > 0$.

5.2.2 The patterns P_{00} and P_{11} are learned

In this case $R_{00} = R_{01} = 0$ and the first order partial derivatives of E with respect to the weights are equal to zero if:

$$
\begin{aligned}
&R_{01} = -R_{10} \neq 0 \\
&f(w_{01} + w_{21}) = f(w_{01} + w_{11}) \\
&f(w_{02} + w_{22}) = f(w_{02} + w_{12}) \\
&v_1 f'(w_{01} + w_{11}) = 0 \\
&v_1 f'(w_{01} + w_{21}) = 0 \\
&v_2 f'(w_{02} + w_{12}) = 0 \\
&v_2 f'(w_{02} + w_{22}) = 0
\end{aligned}
\tag{5.15}
$$

Similarly to the case with the patterns P_{00} and P_{01} learned, it was proved that no local minima occur if one or more of the terms $w_{0i} + w_{1i}$, $w_{0i} + w_{2i}$, $i \in \{1, 2\}$ is finite. The remaining points that have to be investigated are given in the following four cases:

- $w_{01} + w_{11} = w_{01} + w_{21} = w_{02} + w_{12} = w_{02} + w_{22} = \infty$
- $w_{01} + w_{11} = w_{01} + w_{21} = \infty$, $w_{02} + w_{12} = w_{02} + w_{22} = -\infty$
- $w_{01} + w_{11} = w_{01} + w_{21} = -\infty$, $w_{02} + w_{12} = w_{02} + w_{22} = \infty$
- $w_{01} + w_{11} = w_{01} + w_{21} = w_{02} + w_{12} = w_{02} + w_{22} = -\infty$

In the first case we find:

$$A_{01} = A_{10} = u + v_1 + v_2 = 0 \tag{5.16}$$

which is in contradiction with the first equation of (5.15), so in this case no local minimum is found, since even no stationary points are found. Analogously also the other cases will not result in local minima.

Conclusion 5.7 *No stationary points and thus no local minima occur if the patterns* P_{00} *and* P_{11} *are learned and* $R_{01} = -R_{10} \neq 0$.

Using the transformations 2.1, 2.2 and 2.3 local minima are found if the patterns P_{00} *and* P_{10}, P_{01} *and* P_{11}, or P_{10} *and* P_{11} are learned. The resulting local minima are summarized in the following subsection.

5.3 The local minima for two of the patterns P_{ij} learned

Table 2: Local minima with P_{00} and P_{01} learned

$u + v_1 f(w_{01}) + v_2 f(w_{02}) = f^{-1}(0.1)$ $u + v_1 f(w_{01} + w_{21}) + v_2 f(w_{02} + w_{22}) = f^{-1}(0.9)$			error 0.16
$w_{01} + w_{11} = w_{01} + w_{11} + w_{21} = \infty$ $w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = \infty$	$u + v_1 + v_2 = 0$	&	$v_1 \geq f^{-1}(0.9)$, $w_{21} > 0$, $v_2 \leq f^{-1}(0.1)$, $w_{22} < 0$
or			or
$w_{01} + w_{11} = w_{01} + w_{11} + w_{21} = -\infty$ $w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = -\infty$	$u = 0$		$v_1 \leq f^{-1}(0.1)$, $w_{21} < 0$, $v_2 \geq f^{-1}(0.9)$, $w_{22} > 0$
$w_{01} + w_{11} = w_{01} + w_{11} + w_{21} = \infty$ $w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = -\infty$	$u + v_1 = 0$	&	$v_1 \geq f^{-1}(0.9)$, $w_{21} > 0$, $v_2 \geq f^{-1}(0.9)$, $w_{22} > 0$
or			or
$w_{01} + w_{11} = w_{01} + w_{11} + w_{21} = -\infty$ $w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = \infty$	$u + v_2 = 0$		$v_1 \leq f^{-1}(0.1)$, $w_{21} < 0$, $v_2 \leq f^{-1}(0.1)$, $w_{22} < 0$

Table 3: Local minima with P_{00} and P_{10} learned

$u + v_1 f(w_{01}) + v_2 f(w_{02}) = f^{-1}(0.1)$ $u + v_1 f(w_{01} + w_{11}) + v_2 f(w_{02} + w_{12}) = f^{-1}(0.9)$			error 0.16
$w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty$ $w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = \infty$	$u + v_1 + v_2 = 0$	&	$v_1 \geq f^{-1}(0.9)$, $w_{11} > 0$, $v_2 \leq f^{-1}(0.1)$, $w_{12} < 0$
or			or
$w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = -\infty$ $w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = -\infty$	$u = 0$		$v_1 \leq f^{-1}(0.1)$, $w_{11} < 0$, $v_2 \geq f^{-1}(0.9)$, $w_{12} > 0$
$w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty$ $w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = -\infty$	$u + v_1 = 0$	&	$v_1 \geq f^{-1}(0.9)$, $w_{11} > 0$, $v_2 \geq f^{-1}(0.9)$, $w_{12} > 0$
or			or
$w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = -\infty$ $w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = \infty$	$u + v_2 = 0$		$v_1 \leq f^{-1}(0.1)$, $w_{11} < 0$, $v_2 \leq f^{-1}(0.1)$, $w_{12} < 0$

Table 4: Local minima with P_{01} and P_{11} learned

$u + v_1 f(w_{01} + w_{21}) + v_2 f(w_{02} + w_{22}) = f^{-1}(0.9)$ $u + v_1 f(w_{01} + w_{11} + w_{21}) + v_2 f(w_{02} + w_{12} + w_{22}) = f^{-1}(0.1)$ error 0.16			
$w_{01} = w_{01} + w_{11} = \infty$ $w_{02} = w_{02} + w_{12} = \infty$	$u + v_1 + v_2 = 0$		$v_1 \geq f^{-1}(0.9)\,,\; w_{11} < 0,$ $v_2 \leq f^{-1}(0.1)\,,\; w_{12} > 0$
or		&	or
$w_{01} = w_{01} + w_{11} = -\infty$ $w_{02} = w_{02} + w_{12} = -\infty$	$u = 0$		$v_1 \leq f^{-1}(0.1)\,,\; w_{11} > 0,$ $v_2 \geq f^{-1}(0.9)\,,\; w_{12} < 0$
$w_{01} = w_{01} + w_{11} = \infty$ $w_{02} = w_{02} + w_{12} = -\infty$	$u + v_1 = 0$		$v_1 \geq f^{-1}(0.9)\,,\; w_{11} < 0,$ $v_2 \geq f^{-1}(0.9)\,,\; w_{12} < 0$
or		&	or
$w_{01} = w_{01} + w_{11} = -\infty$ $w_{02} = w_{02} + w_{12} = \infty$	$u + v_2 = 0$		$v_1 \leq f^{-1}(0.1)\,,\; w_{11} > 0,$ $v_2 \leq f^{-1}(0.1)\,,\; w_{12} > 0$

Table 5: Local minima with P_{10} and P_{11} learned

$u + v_1 f(w_{01} + w_{11}) + v_2 f(w_{02} + w_{12}) = f^{-1}(0.9)$ $u + v_1 f(w_{01} + w_{11} + w_{21}) + v_2 f(w_{02} + w_{12} + w_{22}) = f^{-1}(0.1)$ error 0.16			
$w_{01} = w_{01} + w_{21} = \infty$ $w_{02} = w_{02} + w_{22} = \infty$	$u + v_1 + v_2 = 0$		$v_1 \geq f^{-1}(0.9)\,,\; w_{21} < 0,$ $v_2 \leq f^{-1}(0.1)\,,\; w_{22} > 0$
or		&	or
$w_{01} = w_{01} + w_{21} = -\infty$ $w_{02} = w_{02} + w_{22} = -\infty$	$u = 0$		$v_1 \leq f^{-1}(0.1)\,,\; w_{21} > 0,$ $v_2 \geq f^{-1}(0.9)\,,\; w_{22} < 0$
$w_{01} = w_{01} + w_{21} = \infty$ $w_{02} = w_{02} + w_{22} = -\infty$	$u + v_1 = 0$		$v_1 \geq f^{-1}(0.9)\,,\; w_{21} < 0,$ $v_2 \geq f^{-1}(0.9)\,,\; w_{22} < 0$
or		&	or
$w_{01} = w_{01} + w_{21} = -\infty$ $w_{02} = w_{02} + w_{22} = \infty$	$u + v_2 = 0$		$v_1 \leq f^{-1}(0.1)\,,\; w_{21} > 0,$ $v_2 \leq f^{-1}(0.1)\,,\; w_{22} > 0$

It is possible to escape from all these local minima via points with $w_{1i} = 0$ (tables 3 and 4) and $w_{2i} = 0$ (tables 2 and 5) ($i, j \in \{1, 2\}$), respectively.

5.4 Exactly one of the patterns P_{ij} is learned

The case that the pattern P_{00} is learned will be studied first. A number of regions with local minima will be found. Via transformations of the weights corresponding results are derived for the cases that one of the other patterns is learned.

5.4.3 The pattern P_{00} is learned

In this case $R_{00} = 0$ and the equalities for stationary points become:

$$\begin{aligned}
&R_{01} + R_{10} + R_{11} = 0 \\
&R_{01}f(w_{01} + w_{21}) + R_{10}f(w_{01} + w_{11}) + R_{11}f(w_{01} + w_{11} + w_{21}) = 0 \\
&R_{01}f(w_{02} + w_{22}) + R_{10}f(w_{02} + w_{12}) + R_{11}f(w_{02} + w_{12} + w_{22}) = 0 \\
&v_1 f'(w_{01} + w_{21}) = 0 \\
&v_1 f'(w_{01} + w_{11}) = 0 \\
&v_1 f'(w_{01} + w_{11} + w_{21}) = 0 \\
&v_2 f'(w_{02} + w_{22}) = 0 \\
&v_2 f'(w_{02} + w_{12}) = 0 \\
&v_2 f'(w_{02} + w_{12} + w_{22}) = 0
\end{aligned} \tag{5.17}$$

Similarly to the proof in section 5.2.1 it can be shown that no local minima occur if one of the terms $w_{0i} + w_{1i}$, $w_{0i} + w_{2i}$, $w_{0i} + w_{1i} + w_{2i}$, $i \in \{1, 2\}$ is finite. Thus we have to consider the cases where all these terms are equal to plus or minus infinity. If $w_{01} + w_{11} = \infty$ and $w_{01} + w_{21} = -\infty$, then equations (5.17) result in $R_{01} = 0$ or $R_{10} = 0$, which case we don't consider here. Similarly we will not have to consider cases where one of the terms $w_{01} + w_{11}$, $w_{01} + w_{21}$ and $w_{01} + w_{11} + w_{21}$ is going to ∞ ($-\infty$) and the other terms are tending to $-\infty$ (∞). For the weights $w_{02} + w_{12}$, $w_{02} + w_{22}$ and $w_{02} + w_{12} + w_{22}$ the same argument holds.

Thus four cases remain:

- $w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = w_{02} + w_{12} = w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = \infty$
- $w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty$ and $w_{02} + w_{12} = w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = -\infty$
- $w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = -\infty$ and $w_{02} + w_{12} = w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = \infty$

- $w_{01}+w_{11} = w_{01}+w_{21} = w_{01}+w_{11}+w_{21} = w_{02}+w_{12} = w_{02}+w_{22} = w_{02}+w_{12}+w_{22} = -\infty$

We will investigate these cases in the following.

Case 5.4.1.1: $w_{01}+w_{11}= w_{01}+w_{21} = w_{01}+w_{11}+w_{21} = w_{02}+w_{12}= w_{02}+w_{22} = w_{02}+w_{12}+w_{22} = \infty$

In this case equations (2.1) result in

$$\begin{aligned} A_{00} &= u+v_1f(w_{01})+v_2f(w_{02}) = f^{-1}(0.1) \\ A_{01} &= A_{10} = A_{11} = u+v_1+v_2 \end{aligned} \tag{5.18}$$

Using equations (4.1) and (5.18) in (5.17) leads to:

$$\begin{aligned} &R_{01}+R_{10}+R_{11} = \\ &2\,(f(u+v_1+v_2)-0.9)\,f'(u+v_1+v_2) + \\ &(f(u+v_1+v_2)-0.1)\,f'(u+v_1+v_2) = \\ &(3f(u+v_1+v_2)-1.9)\,f'(u+v_1+v_2) = 0 \end{aligned}$$

and thus

$$u+v_1+v_2 = f^{-1}(1.9/3) \tag{5.19}$$

Using (5.19) in order to eliminate u from the equation for A_{00} in (5.18) and using that $1-f(x) = f(-x)$ results in:

$$v_1f(-w_{01})+v_2f(-w_{02}) = f^{-1}(1.9/3)-f^{-1}(0.1) > 0 \tag{5.20}$$

and thus $v_1 > 0$ and/or $v_2 > 0$ has to hold.

For R_{01}, R_{10} and R_{11} we find:

$$\begin{aligned} R_{01} &= R_{10} = (f(f^{-1}(1.9/3))-0.9)f'(f^{-1}(1.9/3)) \approx -0.0619259 \\ R_{11} &= (f(f^{-1}(1.9/3))-0.1)f'(f^{-1}(1.9/3)) \approx 0.123852 = \\ &= -2R_{01} = -2R_{10} \end{aligned} \tag{5.21}$$

Altering $u+v_1+v_2$ while keeping A_{00} constant can only increase the error, since the function $2\,(f(x)-0.9)^2+(f(x)-0.1)^2$ attains a minimum for $f(x) = 1.9/3$.

So let us consider what happens when moving $w_{01}+w_{11}$, $w_{01}+w_{21}$, and/or $w_{01}+w_{11}+w_{21}$ away from infinity.

We split this case into two cases:

- $w_{01} = \infty$
- w_{01} is finite

For $w_{01} = \infty$ considering $p = e^{-w_{01}}$, w_{11} and w_{21} as independent variables results in:

$$\left.\frac{\partial E}{\partial p}\right|_{p=0} = -R_{01} v_1 \left(e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11}-w_{21}}\right) \tag{5.22}$$

So, since $R_{01} < 0$, saddle points are found if

- $v_1 \geq 0$ and $e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11}-w_{21}} \leq 0$ or
- $v_1 \leq 0$ and $e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11}-w_{21}} \geq 0$

(Boundary points of regions of saddle points are saddle points too.)

So a necessary condition to obtain a local minimum with $w_{01} = \infty$, is that

- $v_1 > 0$ and $e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11}-w_{21}} > 0$, or
- $v_1 < 0$ and $e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11}-w_{21}} < 0$.

A similar condition is necessary for a local minimum with $w_{02} = \infty$.

Now suppose w_{01} is finite. Then both w_{11} and w_{21} are infinite, because we are considering stationary points with $w_{01} + w_{11} = w_{01} + w_{21} = \infty$. Considering w_{01}, $p_1 = e^{-w_{11}}$ and $p_2 = e^{-w_{21}}$ as independent variables, yields:

$$\left.\frac{\partial E}{\partial p_1}\right|_{p_1 = p_2 = 0} = \left.\frac{\partial E}{\partial p_2}\right|_{p_1 = p_2 = 0} = -R_{01} v_1 e^{-w_{01}} \tag{5.23}$$

So, since R_{01} is negative, saddle points are found if $v_1 \leq 0$.

Hence, a necessary condition to obtain a local minimum with w_{01} finite, is

- $v_1 > 0$.

Similarly a necessary condition to obtain a local minimum with w_{02} finite is

- $v_2 > 0$.

So local minima are obtained if both the weights connected to the first hidden unit and those connected to the second hidden unit satisfy the given restrictions. Remarking that $v_1 > 0$ and/or $v_2 > 0$ has to hold and that w_{01} or w_{02} has to be finite, because of equations (5.18) and (5.19) leads to:

Conclusion 5.8 *If pattern P_{00} is learned, then regions with local minima with error 0.213333 are found if $u + v_1 + v_2 = f^{-1}(1.9/3)$, $w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = w_{02} + w_{12} = w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = \infty$ and if one of the following conditions is fulfilled:*

- $w_{01} = \infty$, w_{02} *finite and either*
  - $v_1 > 0$, $v_2 > 0$ *and* $e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11}-w_{21}} > 0$ *or*
  - $v_1 < 0$, $v_2 > 0$ *and* $e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11}-w_{21}} < 0$

or

- $w_{02} = \infty$, w_{01} *finite and either*

- $v_1 > 0$, $v_2 > 0$ *and* $e^{-w_{12}} + e^{-w_{22}} - 2e^{-w_{12} - w_{22}} > 0$ *or*
- $v_1 > 0$, $v_2 < 0$ *and* $e^{-w_{12}} + e^{-w_{22}} - 2e^{-w_{12} - w_{22}} < 0$

or

- w_{01} *and* w_{02} *are finite and* $v_1 > 0$ *and* $v_2 > 0$.

Case 5.4.1.2: w_{01}+w_{11} = w_{01}+w_{21} = w_{01}+w_{11}+w_{21} = ∞ and w_{02}+w_{12} = w_{02}+w_{22} = w_{02}+w_{12}+w_{22} = –∞

Analogously to case 1 we find:

$$
\begin{aligned}
&A_{00} \;=\; u + v_1 f(w_{01}) + v_2 f(w_{02}) \;=\; f^{-1}(0.1)\\
&A_{01} \;=\; A_{10} \;=\; A_{11} \;=\; u + v_1\\
&R_{01} + R_{10} + R_{11} \;=\; (3f(u + v_1) - 1.9)\, f'(u + v_1) \;=\; 0\\
&u + v_1 \;=\; f^{-1}(1.9/3)
\end{aligned}
\qquad (5.24)
$$

and R_{01}, R_{10} and R_{11} are given by (5.21).

Removing u from the equation for A_{00}, using the last equation of (5.24), results in:

$$v_1 f(-w_{01}) - v_2 f(w_{02}) \;=\; f^{-1}(1.9/3) - f^{-1}(0.1) > 0 \qquad (5.25)$$

and thus $v_1 > 0$ and/or $v_2 < 0$ has to hold.

Altering $u + v_1$ increases the error. Consideration of the weights connected to the first hidden unit again leads to the necessary conditions for local minima:

- $w_{01} \;=\; \infty$ and either
  - $v_1 > 0$ and $e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11} - w_{21}} > 0$ or
  - $v_1 < 0$ and $e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11} - w_{21}} < 0$

or

- w_{01} finite and $v_1 > 0$.

Let us consider the weights connected to the second hidden unit. We split this case into the two cases:

- $w_{02} \;=\; -\infty$
- w_{02} is finite

If $w_{02} \;=\; -\infty$ considering the independent variables $q \;=\; e^{w_{02}}$, w_{12} and w_{22} gives:

$$\left.\frac{\partial E}{\partial q}\right|_{q\,=\,0} \;=\; R_{01} v_2 \,(e^{w_{12}} + e^{w_{22}} - 2e^{w_{12} w_{22}}) \qquad (5.26)$$

So, since $R_{01} < 0$, a necessary condition to obtain a local minimum with $w_{02} \;=\; -\infty$ is:

- $v_2 > 0$ and $e^{w_{12}} + e^{w_{22}} - 2e^{w_{12} + w_{22}} < 0$ or
- $v_2 < 0$ and $e^{w_{12}} + e^{w_{22}} - 2e^{w_{12} + w_{22}} > 0$.

Now suppose w_{02} is finite and thus both w_{12} and w_{22} are equal to minus infinity. Considering the independent variables w_{02}, $q_1 = e^{w_{12}}$ and $q_2 = e^{w_{22}}$ leads to:

$$\left.\frac{\partial E}{\partial q_1}\right|_{q_1 = q_2 = 0} = \left.\frac{\partial E}{\partial q_2}\right|_{q_1 = q_2 = 0} = R_{01} v_2 e^{w_{02}} \tag{5.27}$$

So a necessary condition to obtain a local minimum if w_{02} is finite and w_{12} and w_{22} are equal to minus infinity is

- $v_2 < 0$.

Since $v_1 > 0$ and/or $v_2 < 0$ has to hold and since $w_{01} = \infty$ and $w_{02} = -\infty$ can not occur because of equation (5.24), the following conclusion results:

Conclusion 5.9 *If pattern P_{00} is learned, then regions with local minima with error 0.213333 will be found if $u + v_1 = f^{-1}(1.9/3)$, $w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty$ and $w_{02} + w_{12} = w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = -\infty$ and if one of the following conditions is fulfilled:*

- $w_{01} = \infty$, w_{02} *finite and either*
  - $v_1 > 0$, $v_2 < 0$ *and* $e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11} - w_{21}} > 0$ *or*
  - $v_1 < 0$, $v_2 < 0$ *and* $e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11} - w_{21}} < 0$

or

- $w_{02} = -\infty$, w_{01} *finite and either*
  - $v_1 > 0$, $v_2 > 0$ *and* $e^{w_{12}} + e^{w_{22}} - 2e^{w_{12} + w_{22}} < 0$ *or*
  - $v_1 > 0$, $v_2 < 0$ *and* $e^{w_{12}} + e^{w_{22}} - 2e^{w_{12} + w_{22}} > 0$

or

- w_{01} *and* w_{02} *are finite,* $v_1 > 0$ *and* $v_2 < 0$.

Also boundary points of the regions with local minima of conclusion 5.9 are saddle points.

Case 5.4.1.3: $w_{01}+w_{11} = w_{01}+w_{21} = w_{01}+w_{11}+w_{21} = -\infty$ and $w_{02}+w_{12} = w_{02}+w_{22} = w_{02}+w_{12}+w_{22} = \infty$

Analogously to case 2 we find the following conclusion:

Conclusion 5.10 *If pattern P_{00} is learned, then regions with local minima with error 0.213333 will be found if $u + v_2 = f^{-1}(1.9/3)$, $w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = -\infty$ and $w_{02} + w_{12} = w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = \infty$ and if one of the following conditions is fulfilled:*

- $w_{02} = \infty$, w_{01} *finite and either*
  - $v_1 < 0$, $v_2 > 0$ *and* $e^{-w_{12}} + e^{-w_{22}} - 2e^{-w_{12} - w_{22}} > 0$ *or*
  - $v_1 < 0$, $v_2 < 0$ *and* $e^{-w_{12}} + e^{-w_{22}} - 2e^{-w_{12} - w_{22}} < 0$

or

- $w_{01} = -\infty$, w_{02} *finite and either*
  - $v_1 > 0$, $v_2 > 0$ *and* $e^{w_{11}} + e^{w_{21}} - 2e^{w_{11} + w_{21}} < 0$ *or*
  - $v_1 < 0$, $v_2 > 0$ *and* $e^{w_{11}} + e^{w_{21}} - 2e^{w_{11} + w_{21}} > 0$

or

- w_{01} *and* w_{02} *are finite,* $v_1 < 0$ *and* $v_2 > 0$.

These regions have boundary points which are saddle points.

Case 5.4.1.4: $w_{01}+w_{11} = w_{01}+w_{21} = w_{01}+w_{11}+w_{21} = w_{02}+w_{12} = w_{02}+w_{22} = w_{02}+w_{12}+w_{22} = -\infty$

Analogously to the previous cases we find:

$$\begin{aligned} A_{00} &= u + v_1 f(w_{01}) + v_2 f(w_{02}) = f^{-1}(0.1) \\ A_{01} &= A_{10} = A_{11} = u \\ R_{01} + R_{10} + R_{11} &= (3f(u) - 1.9) f'(u) = 0 \\ u &= f^{-1}(1.9/3) \end{aligned} \tag{5.28}$$

and R_{01}, R_{10} and R_{11} are given by (5.21). Hence:

$$v_1 f(w_{01}) + v_2 f(w_{02}) = f^{-1}(0.1) - f^{-1}(1.9/3) < 0 \tag{5.29}$$

and thus $v_1 < 0$ and/or $v_2 < 0$ has to hold.

Since $v_1 < 0$ and/or $v_2 < 0$ and $w_{01} = w_{02} = -\infty$ can not hold, we find the conclusion:

Conclusion 5.11 *If pattern* P_{00} *is learned, then regions with local minima with error 0.213333 will be found if* $u = f^{-1}(1.9/3)$, $w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = w_{02} + w_{12} = w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = -\infty$ *and if one of the following conditions is fulfilled:*

- $w_{01} = -\infty$, w_{02} *finite and either*
  - $v_1 > 0$, $v_2 < 0$ *and* $e^{w_{11}} + e^{w_{21}} - 2e^{w_{11} + w_{21}} < 0$ *or*
  - $v_1 < 0$, $v_2 < 0$ *and* $e^{w_{11}} + e^{w_{21}} - 2e^{w_{11} + w_{21}} > 0$

or

- $w_{02} = -\infty$, w_{01} *finite and either*
  - $v_1 < 0$, $v_2 > 0$ *and* $e^{w_{12}} + e^{w_{22}} - 2e^{w_{12} + w_{22}} < 0$ *or*
  - $v_1 < 0$, $v_2 < 0$ *and* $e^{w_{12}} + e^{w_{22}} - 2e^{w_{12} + w_{22}} > 0$

or

- w_{01} *and* w_{02} *are finite,* $v_1 < 0$ *and* $v_2 < 0$.

Also these regions with local minima have boundary points which are saddle points.

So conclusions 5.8, 5.9, 5.10 and 5.11 give the local minima for the case that *pattern* P_{00} is learned and R_{01}, R_{10} and R_{11} are unequal to zero. All these local minima have boundary points which are saddle points.

Using the transformations 1, 2 and 3 from section 5.2 immediately gives the corresponding local minima if one of the other patterns is learned. These local minima are summarized in the next subsection.

5.5 Summary of the local minima with one pattern learned

Table 6: Local minima with pattern P_{00} learned

$u + v_1 f(w_{01}) + v_2 f(w_{02}) = f^{-1}(0.1)$	error 0.213333
$w_{01} + w_{11} = w_{01} + w_{21} =$ $w_{01} + w_{11} + w_{21} = \infty$ $w_{02} + w_{12} = w_{02} + w_{22} =$ $w_{02} + w_{12} + w_{22} = \infty$ $u + v_1 + v_2 = f^{-1}(1.9/3)$	$w_{01} = \infty$, w_{02} finite and $v_1 > 0, v_2 > 0, e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11}-w_{21}} > 0$ or $v_1 < 0, v_2 > 0, e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11}-w_{21}} < 0$
	$w_{02} = \infty$, w_{01} finite and $v_1 > 0, v_2 > 0, e^{-w_{12}} + e^{-w_{22}} - 2e^{-w_{12}-w_{22}} > 0$ or $v_1 > 0, v_2 < 0, e^{-w_{12}} + e^{-w_{22}} - 2e^{-w_{12}-w_{22}} < 0$
	w_{01}, w_{02} finite, $v_1 > 0, v_2 > 0$
$w_{01} + w_{11} = w_{01} + w_{21} =$ $w_{01} + w_{11} + w_{21} = \infty$ $w_{02} + w_{12} = w_{02} + w_{22} =$ $w_{02} + w_{12} + w_{22} = -\infty$ $u + v_1 = f^{-1}(1.9/3)$	$w_{01} = \infty$, w_{02} finite and $v_1 > 0, v_2 < 0, e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11}-w_{21}} > 0$ or $v_1 < 0, v_2 < 0, e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11}-w_{21}} < 0$
	$w_{02} = -\infty$, w_{01} finite and $v_1 > 0, v_2 > 0, e^{w_{12}} + e^{w_{22}} - 2e^{w_{12}+w_{22}} < 0$ or $v_1 > 0, v_2 < 0, e^{w_{12}} + e^{w_{22}} - 2e^{w_{12}+w_{22}} > 0$
	w_{01}, w_{02} finite, $v_1 > 0, v_2 < 0$
$w_{01} + w_{11} = w_{01} + w_{21} =$ $w_{01} + w_{11} + w_{21} = -\infty$ $w_{02} + w_{12} = w_{02} + w_{22} =$ $w_{02} + w_{12} + w_{22} = \infty$ $u + v_2 = f^{-1}(1.9/3)$	$w_{02} = \infty$, w_{01} finite and $v_1 < 0, v_2 > 0, e^{-w_{12}} + e^{-w_{22}} - 2e^{-w_{12}-w_{22}} > 0$ or $v_1 < 0, v_2 < 0, e^{-w_{12}} + e^{-w_{22}} - 2e^{-w_{12}-w_{22}} < 0$
	$w_{01} = -\infty$, w_{02} finite and $v_1 > 0, v_2 > 0, e^{w_{11}} + e^{w_{21}} - 2e^{w_{11}+w_{21}} < 0$ or $v_1 < 0, v_2 > 0, e^{w_{11}} + e^{w_{21}} - 2e^{w_{11}+w_{21}} > 0$
	w_{01}, w_{02} finite, $v_1 < 0, v_2 > 0$
$w_{01} + w_{11} = w_{01} + w_{21} =$ $w_{01} + w_{11} + w_{21} = -\infty$ $w_{02} + w_{12} = w_{02} + w_{22} =$ $w_{02} + w_{12} + w_{22} = -\infty$ $u = f^{-1}(1.9/3)$	$w_{01} = -\infty$, w_{02} finite and $v_1 > 0, v_2 < 0, e^{w_{11}} + e^{w_{21}} - 2e^{w_{11}+w_{21}} < 0$ or $v_1 < 0, v_2 < 0, e^{w_{11}} + e^{w_{21}} - 2e^{w_{11}+w_{21}} > 0$
	$w_{02} = -\infty$, w_{01} finite and $v_1 < 0, v_2 > 0, e^{w_{12}} + e^{w_{22}} - 2e^{w_{12}+w_{22}} < 0$ or $v_1 < 0, v_2 < 0, e^{w_{12}} + e^{w_{22}} - 2e^{w_{12}+w_{22}} > 0$
	w_{01}, w_{02} finite, $v_1 < 0, v_2 < 0$

Table 7: Local minima with pattern P_{01} learned

$u + v_1 f(w_{01} + w_{21}) + v_2 f(w_{02} + w_{22}) = f^{-1}(0.9)$	error 0.213333
$w_{01} = w_{01} + w_{11} =$ $w_{01} + w_{11} + w_{21} = -\infty$ $w_{02} = w_{02} + w_{12} =$ $w_{02} + w_{12} + w_{22} = -\infty$ $u = f^{-1}(1.1/3)$	$w_{01} + w_{21} = -\infty$, $w_{02} + w_{22}$ finite and $v_1 > 0, v_2 > 0, e^{w_{11}} + e^{-w_{21}} - 2e^{w_{11} - w_{21}} > 0$ or $v_1 < 0, v_2 > 0, e^{w_{11}} + e^{-w_{21}} - 2e^{w_{11} - w_{21}} < 0$
	$w_{02} + w_{22} = -\infty$, $w_{01} + w_{21}$ finite and $v_1 > 0, v_2 > 0, e^{w_{12}} + e^{-w_{22}} - 2e^{w_{12} - w_{22}} > 0$ or $v_1 > 0, v_2 < 0, e^{w_{12}} + e^{-w_{22}} - 2e^{w_{12} - w_{22}} < 0$
	$w_{01} + w_{21}, w_{02} + w_{22}$ finite, $v_1 > 0, v_2 > 0$
$w_{01} = w_{01} + w_{11} =$ $w_{01} + w_{11} + w_{21} = -\infty$ $w_{02} = w_{02} + w_{12} =$ $w_{02} + w_{12} + w_{22} = \infty$ $u + v_2 = f^{-1}(1.1/3)$	$w_{01} + w_{21} = -\infty$, $w_{02} + w_{22}$ finite and $v_1 > 0, v_2 < 0, e^{w_{11}} + e^{-w_{21}} - 2e^{w_{11} - w_{21}} > 0$ or $v_1 < 0, v_2 < 0, e^{w_{11}} + e^{-w_{21}} - 2e^{w_{11} - w_{21}} < 0$
	$w_{02} + w_{22} = \infty$, $w_{01} + w_{21}$ finite and $v_1 > 0, v_2 > 0, e^{-w_{12}} + e^{w_{22}} - 2e^{-w_{12} + w_{22}} < 0$ or $v_1 > 0, v_2 < 0, e^{-w_{12}} + e^{w_{22}} - 2e^{-w_{12} + w_{22}} > 0$
	$w_{01} + w_{21}, w_{02} + w_{22}$ finite, $v_1 > 0, v_2 < 0$
$w_{01} = w_{01} + w_{11} =$ $w_{01} + w_{11} + w_{21} = \infty$ $w_{02} = w_{02} + w_{12} =$ $w_{02} + w_{12} + w_{22} = -\infty$ $u + v_1 = f^{-1}(1.1/3)$	$w_{02} + w_{22} = -\infty$, $w_{01} + w_{21}$ finite and $v_1 < 0, v_2 > 0, e^{w_{12}} + e^{-w_{22}} - 2e^{w_{12} - w_{22}} > 0$ or $v_1 < 0, v_2 < 0, e^{w_{12}} + e^{-w_{22}} - 2e^{w_{12} - w_{22}} < 0$
	$w_{01} + w_{21} = \infty$, $w_{02} + w_{22}$ finite and $v_1 > 0, v_2 > 0, e^{-w_{11}} + e^{w_{21}} - 2e^{-w_{11} + w_{21}} < 0$ or $v_1 < 0, v_2 > 0, e^{-w_{11}} + e^{w_{21}} - 2e^{-w_{11} + w_{21}} > 0$
	$w_{01} + w_{21}, w_{02} + w_{22}$ finite, $v_1 < 0, v_2 > 0$
$w_{01} = w_{01} + w_{12} =$ $w_{01} + w_{11} + w_{21} = \infty$ $w_{02} = w_{02} + w_{12} =$ $w_{02} + w_{12} + w_{22} = \infty$ $u + v_1 + v_2 = f^{-1}(1.1/3)$	$w_{01} + w_{21} = \infty$, $w_{02} + w_{22}$ finite and $v_1 > 0, v_2 < 0, e^{-w_{11}} + e^{w_{21}} - 2e^{-w_{11} + w_{21}} < 0$ or $v_1 < 0, v_2 < 0, e^{-w_{11}} + e^{w_{21}} - 2e^{-w_{11} + w_{21}} > 0$
	$w_{02} + w_{22} = \infty$, $w_{01} + w_{21}$ finite and $v_1 < 0, v_2 > 0, e^{-w_{12}} + e^{w_{22}} - 2e^{-w_{12} + w_{22}} < 0$ or $v_1 < 0, v_2 < 0, e^{-w_{12}} + e^{w_{22}} - 2e^{-w_{12} + w_{22}} > 0$
	$w_{01} + w_{21}, w_{02} + w_{22}$ finite , $v_1 < 0, v_2 < 0$

Table 8: Local minima with pattern P_{10} learned

<table>
<tr><td colspan="2">$u + v_1 f(w_{01} + w_{11}) + v_2 f(w_{02} + w_{12}) = f^{-1}(0.9)$ error 0.213333</td></tr>
<tr><td rowspan="3">$w_{01} = w_{01} + w_{21} =$
$w_{01} + w_{11} + w_{21} = -\infty$
$w_{02} = w_{02} + w_{22} =$
$w_{02} + w_{12} + w_{22} = -\infty$
$u = f^{-1}(1.1/3)$</td><td>$w_{01} + w_{11} = -\infty$, $w_{02} + w_{12}$ finite and
$v_1 > 0, v_2 > 0, e^{w_{21}} + e^{-w_{11}} - 2e^{w_{21} - w_{11}} > 0$ or
$v_1 < 0, v_2 > 0, e^{w_{21}} + e^{-w_{11}} - 2e^{w_{21} - w_{11}} < 0$</td></tr>
<tr><td>$w_{02} + w_{12} = -\infty$, $w_{01} + w_{11}$ finite and
$v_1 > 0, v_2 > 0, e^{w_{22}} + e^{-w_{12}} - 2e^{w_{22} - w_{12}} > 0$ or
$v_1 > 0, v_2 < 0, e^{w_{22}} + e^{-w_{12}} - 2e^{w_{22} - w_{12}} < 0$</td></tr>
<tr><td>$w_{01} + w_{11}, w_{02} + w_{12}$ finite, $v_1 > 0, v_2 > 0$</td></tr>
<tr><td rowspan="3">$w_{01} = w_{01} + w_{21} =$
$w_{01} + w_{11} + w_{21} = -\infty$
$w_{02} = w_{02} + w_{22} =$
$w_{02} + w_{12} + w_{22} = \infty$
$u + v_2 = f^{-1}(1.1/3)$</td><td>$w_{01} + w_{11} = -\infty$, $w_{02} + w_{12}$ finite and
$v_1 > 0, v_2 < 0, e^{w_{21}} + e^{-w_{11}} - 2e^{w_{21} - w_{11}} > 0$ or
$v_1 < 0, v_2 < 0, e^{w_{21}} + e^{-w_{11}} - 2e^{w_{21} - w_{11}} < 0$</td></tr>
<tr><td>$w_{02} + w_{12} = \infty$, $w_{01} + w_{11}$ finite and
$v_1 > 0, v_2 > 0, e^{-w_{22}} + e^{w_{12}} - 2e^{-w_{22} + w_{12}} < 0$ or
$v_1 > 0, v_2 < 0, e^{-w_{22}} + e^{w_{12}} - 2e^{-w_{22} + w_{12}} > 0$</td></tr>
<tr><td>$w_{01} + w_{11}, w_{02} + w_{12}$ finite, $v_1 > 0, v_2 < 0$</td></tr>
<tr><td rowspan="3">$w_{01} = w_{01} + w_{21} =$
$w_{01} + w_{11} + w_{21} = \infty$
$w_{02} = w_{02} + w_{22} =$
$w_{02} + w_{12} + w_{22} = -\infty$
$u + v_1 = f^{-1}(1.1/3)$</td><td>$w_{02} + w_{12} = -\infty$, $w_{01} + w_{11}$ finite and
$v_1 < 0, v_2 > 0, e^{w_{22}} + e^{-w_{12}} - 2e^{w_{22} - w_{12}} > 0$ or
$v_1 < 0, v_2 < 0, e^{w_{22}} + e^{-w_{12}} - 2e^{w_{22} - w_{12}} < 0$</td></tr>
<tr><td>$w_{01} + w_{11} = \infty$, $w_{02} + w_{12}$ finite and
$v_1 > 0, v_2 > 0, e^{-w_{21}} + e^{w_{11}} - 2e^{-w_{21} + w_{11}} < 0$ or
$v_1 < 0, v_2 > 0, e^{-w_{21}} + e^{w_{11}} - 2e^{-w_{21} + w_{11}} > 0$</td></tr>
<tr><td>$w_{01} + w_{11}, w_{02} + w_{12}$ finite, $v_1 < 0, v_2 > 0$</td></tr>
<tr><td rowspan="3">$w_{01} = w_{01} + w_{21} =$
$w_{01} + w_{11} + w_{21} = \infty$
$w_{02} = w_{02} + w_{22} =$
$w_{02} + w_{12} + w_{22} = \infty$
$u + v_1 + v_2 = f^{-1}(1.1/3)$</td><td>$w_{01} + w_{11} = \infty$, $w_{02} + w_{12}$ finite and
$v_1 > 0, v_2 < 0, e^{-w_{21}} + e^{w_{11}} - 2e^{-w_{21} + w_{11}} < 0$ or
$v_1 < 0, v_2 < 0, e^{-w_{21}} + e^{w_{11}} - 2e^{-w_{21} + w_{11}} > 0$</td></tr>
<tr><td>$w_{02} + w_{12} = \infty$, $w_{01} + w_{11}$ finite and
$v_1 < 0, v_2 > 0, e^{-w_{22}} + e^{w_{12}} - 2e^{-w_{22} + w_{12}} < 0$ or
$v_1 < 0, v_2 < 0, e^{-w_{22}} + e^{w_{12}} - 2e^{-w_{22} + w_{12}} > 0$</td></tr>
<tr><td>$w_{01} + w_{11}, w_{02} + w_{12}$ finite , $v_1 < 0, v_2 < 0$</td></tr>
</table>

Table 9: Local minima with pattern P_{11} learned

<table>
<tr><td colspan="2">$u + v_1 f(w_{01} + w_{11} + w_{21}) + v_2 f(w_{02} + w_{12} + w_{22}) = f^{-1}(0.1)$ error 0.213333</td></tr>
<tr><td rowspan="3">$w_{01} = w_{01} + w_{11} =$
$w_{01} + w_{21} = \infty$
$w_{02} = w_{02} + w_{12} =$
$w_{02} + w_{22} = \infty$
$u + v_1 + v_2 = f^{-1}(1.9/3)$</td><td>$w_{01} + w_{11} + w_{21} = \infty$, $w_{02} + w_{12} + w_{22}$ finite and
$v_1 > 0, v_2 > 0, e^{w_{11}} + e^{w_{21}} - 2e^{w_{11} + w_{21}} > 0$ or
$v_1 < 0, v_2 > 0, e^{w_{11}} + e^{w_{21}} - 2e^{w_{11} + w_{21}} < 0$</td></tr>
<tr><td>$w_{02} + w_{12} + w_{22} = \infty$, $w_{01} + w_{11} + w_{21}$ finite and
$v_1 > 0, v_2 > 0, e^{w_{12}} + e^{w_{22}} - 2e^{w_{12} + w_{22}} > 0$ or
$v_1 > 0, v_2 < 0, e^{w_{12}} + e^{w_{22}} - 2e^{w_{12} + w_{22}} < 0$</td></tr>
<tr><td>$w_{01} + w_{11} + w_{21}$, $w_{02} + w_{12} + w_{22}$ finite, $v_1 > 0, v_2 > 0$</td></tr>
<tr><td rowspan="3">$w_{01} = w_{01} + w_{11} =$
$w_{01} + w_{21} = \infty$
$w_{02} = w_{02} + w_{12} =$
$w_{02} + w_{22} = -\infty$
$u + v_1 = f^{-1}(1.9/3)$</td><td>$w_{01} + w_{11} + w_{21} = \infty$, $w_{02} + w_{12} + w_{22}$ finite and
$v_1 > 0, v_2 < 0, e^{w_{11}} + e^{w_{21}} - 2e^{w_{11} + w_{21}} > 0$ or
$v_1 < 0, v_2 < 0, e^{w_{11}} + e^{w_{21}} - 2e^{w_{11} + w_{21}} < 0$</td></tr>
<tr><td>$w_{02} + w_{12} + w_{22} = -\infty$, $w_{01} + w_{11} + w_{21}$ finite and
$v_1 > 0, v_2 > 0, e^{-w_{12}} + e^{-w_{22}} - 2e^{-w_{12} - w_{22}} < 0$ or
$v_1 > 0, v_2 < 0, e^{-w_{12}} + e^{-w_{22}} - 2e^{-w_{12} - w_{22}} > 0$</td></tr>
<tr><td>$w_{01} + w_{11} + w_{21}$, $w_{02} + w_{12} + w_{22}$ finite, $v_1 > 0, v_2 < 0$</td></tr>
<tr><td rowspan="3">$w_{01} = w_{01} + w_{11} =$
$w_{01} + w_{21} = -\infty$
$w_{02} = w_{02} + w_{12} =$
$w_{02} + w_{22} = \infty$
$u + v_2 = f^{-1}(1.9/3)$</td><td>$w_{02} + w_{12} + w_{22} = \infty$, $w_{01} + w_{11} + w_{21}$ finite and
$v_1 < 0, v_2 > 0, e^{w_{12}} + e^{w_{22}} - 2e^{w_{12} + w_{22}} > 0$ or
$v_1 < 0, v_2 < 0, e^{w_{12}} + e^{w_{22}} - 2e^{w_{12} + w_{22}} < 0$</td></tr>
<tr><td>$w_{01} + w_{11} + w_{21} = -\infty$, $w_{02} + w_{12} + w_{22}$ finite and
$v_1 > 0, v_2 > 0, e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11} - w_{21}} < 0$ or
$v_1 < 0, v_2 > 0, e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11} - w_{21}} > 0$</td></tr>
<tr><td>$w_{01} + w_{11} + w_{21}$, $w_{02} + w_{12} + w_{22}$ finite, $v_1 < 0, v_2 > 0$</td></tr>
<tr><td rowspan="3">$w_{01} = w_{01} + w_{11} =$
$w_{01} + w_{21} = -\infty$
$w_{02} = w_{02} + w_{12} =$
$w_{02} + w_{22} = -\infty$
$u = f^{-1}(1.9/3)$</td><td>$w_{01} + w_{11} + w_{21} = -\infty$, $w_{02} + w_{12} + w_{22}$ finite and
$v_1 > 0, v_2 < 0, e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11} - w_{21}} < 0$ or
$v_1 < 0, v_2 < 0, e^{-w_{11}} + e^{-w_{21}} - 2e^{-w_{11} - w_{21}} > 0$</td></tr>
<tr><td>$w_{02} + w_{12} + w_{22} = -\infty$, $w_{01} + w_{11} + w_{21}$ finite and
$v_1 < 0, v_2 > 0, e^{-w_{12}} + e^{-w_{22}} - 2e^{-w_{12} - w_{22}} < 0$ or
$v_1 < 0, v_2 < 0, e^{-w_{12}} + e^{-w_{22}} - 2e^{-w_{12} - w_{22}} > 0$</td></tr>
<tr><td>$w_{01} + w_{11} + w_{21}$, $w_{02} + w_{12} + w_{22}$ finite, $v_1 < 0, v_2 < 0$</td></tr>
</table>

5.6 All terms R_{ij} are unequal to zero

The conditions for stationary points become in this case:

$$
\begin{aligned}
& R_{00} + R_{01} + R_{10} + R_{11} = 0 \\
& R_{00} f(w_{01}) + R_{01} f(w_{01} + w_{21}) + \\
& \quad + R_{10} f(w_{01} + w_{11}) + R_{11} f(w_{01} + w_{11} + w_{21}) = 0 \\
& R_{00} f(w_{02}) + R_{01} f(w_{02} + w_{22}) + \\
& \quad + R_{10} f(w_{02} + w_{12}) + R_{11} f(w_{02} + w_{12} + w_{22}) = 0 \\
& v_1 R_{00} f'(w_{01}) = -v_1 R_{01} f'(w_{01} + w_{21}) = \\
& -v_1 R_{10} f'(w_{01} + w_{11}) = v_1 R_{11} f'(w_{01} + w_{11} + w_{21}) \\
& v_2 R_{00} f'(w_{02}) = -v_2 R_{01} f'(w_{02} + w_{22}) = \\
& -v_2 R_{10} f'(w_{02} + w_{12}) = v_2 R_{11} f'(w_{02} + w_{12} + w_{22})
\end{aligned}
\tag{5.30}
$$

If w_{01} is finite, then either $v_1 = 0$ or also $w_{01} + w_{11}$, $w_{01} + w_{21}$ and $w_{01} + w_{11} + w_{21}$ have to be finite. Similarly to previous proofs $v_1 = 0$ and/or $v_2 = 0$ result in saddle points. All terms w_{0i}, $w_{0i} + w_{1i}$, $w_{0i} + w_{2i}$, $w_{0i} + w_{1i} + w_{2i}$, $i \in \{1, 2\}$ finite also result in saddle points (see section 4.2 in [10]). So we have to consider here the stationary points with w_{01}, $w_{01} + w_{11}$, $w_{01} + w_{21}$, and $w_{01} + w_{11} + w_{21}$ infinite and/or w_{02}, $w_{02} + w_{12}$, $w_{02} + w_{22}$, and $w_{02} + w_{12} + w_{22}$ infinite.

Remark that $R_{ij} \neq 0$, $i, j \in \{0, 1\}$, and $w_{01} + (w_{01} + w_{11} + w_{21}) = (w_{01} + w_{11}) + (w_{01} + w_{21})$ and equations (5.30) have to hold. So, we have the following possibilities for w_{01}, $w_{01} + w_{11}$, $w_{01} + w_{21}$, and $w_{01} + w_{11} + w_{21}$ infinite:

- $w_{01} = w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \pm\infty$ or
- $w_{01} = w_{01} + w_{11} = \pm\infty$ and $w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \mp\infty$ or
- $w_{01} = w_{01} + w_{21} = \pm\infty$ and $w_{01} + w_{11} = w_{01} + w_{11} + w_{21} = \mp\infty$.

In the following we will investigate a number of characteristic cases:

- $w_{01} = w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty$ and w_{02}, $w_{02} + w_{12}$, $w_{02} + w_{22}$, and $w_{02} + w_{12} + w_{22}$ finite
- $w_{01} = w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty$ and $w_{02} = w_{02} + w_{12} = w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = \infty$
- $w_{01} = w_{01} + w_{11} = \infty$ and $w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = -\infty$ and w_{02}, $w_{02} + w_{12}$, $w_{02} + w_{22}$, and $w_{02} + w_{12} + w_{22}$ finite
- $w_{01} = w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty$ and $w_{02} = w_{02} + w_{12} = \infty$ and $w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = -\infty$
- $w_{01} = w_{01} + w_{11} = \infty$ and $w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = -\infty$ and $w_{02} = w_{02} + w_{12} = \infty$ and $w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = -\infty$
- $w_{01} = w_{01} + w_{11} = \infty$ and $w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = -\infty$ and w_{02}

$= w_{02} + w_{22} = \infty$ and $w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = -\infty$

We will find a number of regions with local minima. Analogously, the other regions with local minima can be derived.

Case 5.6.1: $w_{01} = w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty$ and w_{02}, $w_{02} + w_{12}$, $w_{02} + w_{22}$, and $w_{02} + w_{12} + w_{22}$ finite

If w_{02}, $w_{02} + w_{12}$, $w_{02} + w_{22}$, and $w_{02} + w_{12} + w_{22}$ are finite and $v_2 \neq 0$ it follows that $w_{12} = 0$ or $w_{22} = 0$ has to hold (see section 4.2 in [10]). If $w_{12} = 0$ it follows from (2.1) that $A_{00} = A_{10}$ and from (5.30) that $R_{00} = -R_{10}$ and thus $A_{00} = A_{10} = 0$. Also we find $A_{01} = A_{11} = 0$ and thus also $w_{22} = 0$ has to hold. Thus $A_{00} = A_{01} = A_{10} = A_{11} = 0$, and $R_{00} = -R_{01} = -R_{10} = R_{11}$ for the stationary points considered here. So taking $p = e^{-w_{01}}$, w_{11} and w_{21} as independent variables results in

$$\left.\frac{\partial E}{\partial p}\right|_{p=0} = -R_{00}v_1(1 - e^{-w_{11}} - e^{-w_{21}} + e^{-w_{11} - w_{21}}) = \\ -R_{00}v_1(1 - e^{-w_{11}})(1 - e^{-w_{21}})$$

and thus, since $R_{00} = (f(0) - 0.1)f'(0) > 0$, if

- $v_1(1 - e^{-w_{11}})(1 - e^{-w_{21}}) < 0$ or equivalently $v_1 w_{11} w_{21} < 0$

decreasing w_{01} away from infinity will result in an increasing error. If $w_{01} = \infty$ then w_{11} and w_{21} have no influence on the error. Considering the second order part of the Taylor series expansion with respect to the weights u, v_1, v_2, w_{02}, w_{12}, and w_{22} leads to:

$$\begin{aligned}\Delta E = {} & 4\{f'(0)\}^2[\Delta u + \Delta v_1 + f(w_{02})\Delta v_2 + v_2 f'(w_{02})\Delta w_{02} + \\ & \tfrac{1}{2}v_2 f'(w_{02})(\Delta w_{12} + \Delta w_{22})]^2 + \\ & \{f'(0)\}^2 v_2^2\{f'(w_{02})\}^2[(\Delta w_{12})^2 + (\Delta w_{22})^2] + \\ & 2v_2(f(0) - 0.1)f'(0)f''(w_{02})\Delta w_{12}\Delta w_{22}\end{aligned} \tag{5.31}$$

The first quadratic term can be made zero by chosing Δu. The remaining terms in ΔE are positive if $\Delta w_{12} \neq 0$ and/or $\Delta w_{22} \neq 0$ if the discriminant of the quadratic equation is negative:

$$[2v_2(f(0) - 0.1)f'(0)f''(w_{02})]^2 - 4[\{f'(0)\}^2 v_2^2\{f'(w_{02})\}^2]^2 < 0$$

resulting in

$$|v_2| > 1.6\frac{|f''(w_{02})|}{\{f'(w_{02})\}^2} \tag{5.32}$$

So under this condition altering w_{12} and/or w_{22} can only increase the error. But if w_{12} and w_{22} are kept constant equal to zero then all terms A_{ij} are equal to $u + v_1 + v_2 f(w_{02})$ and altering this also results in increasing the error. Thus we found the region of local minima:

Conclusion 5.12 *If all terms R_{ij} are unequal to zero, a 5-dimensional region of local minima exists if $v_1 \neq 0$, $v_2 \neq 0$, $w_{01} = w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty$, w_{02} finite, $w_{12} = w_{22} = 0$, $u + v_1 + v_2 f(w_{02}) = 0$ and*

- $v_1 w_{11} w_{21} < 0$ *and*
- $|v_2| > 1.6 |f''(w_{02})| / \{f'(w_{02})\}^2$

Case 5.6.2: $w_{01} = w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty$ and $w_{02} = w_{02} + w_{12} = w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = \infty$

Here we find the following local minima:

Conclusion 5.13 *If all terms R_{ij} are unequal to zero, a 6-dimensional region of local minima exists if $v_1 \neq 0$, $v_2 \neq 0$, $w_{01} = w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty$, $w_{02} = w_{02} + w_{12} = w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = \infty$, $u + v_1 + v_2 = 0$ and*

- $v_1 w_{11} w_{21} < 0$ *and*
- $v_2 w_{12} w_{22} < 0$

Case 5.6.3: $w_{01} = w_{01} + w_{11} = \infty$ and $w_{01} + w_{21} = w_{01} + w_{11} + w_{21} -\infty$ and w_{02}, $w_{02} + w_{12}$, $w_{02} + w_{22}$, and $w_{02} + w_{12} + w_{22}$ finite

From (5.30) it follows that $R_{00} = -R_{10}$ and $R_{11} = -R_{01}$. Also it can be derived that $w_{12} = 0$ or $w_{22} = 0$ (see section 4.2 in [10]). If $w_{12} = 0$, it follows from (5.30) that $R_{00} = -R_{01} = -R_{10} = R_{11}$. If $w_{22} = 0$ then either $w_{12} = 0$ or $w_{12} = -w_{02}$, because of (5.30). If $w_{22} = 0$ and $w_{12} = -w_{02}$, then (5.30) leads to $(R_{00} - R_{11})(2f(w_{02}) - 1) = 0$. Thus either $R_{00} = R_{11}$ or $w_{02} = 0$. In both cases we find again that $R_{00} = -R_{01} = -R_{10} = R_{11}$.

Considering $p = e^{-w_{01}}$, $q = e^{w_{01} + w_{21}}$ and w_{11} as independent variables yields

$$\frac{\partial E}{\partial p_1} = -R_{00} v_1 (1 - e^{-w_{11}})$$

$$\frac{\partial E}{\partial q_1} = -R_{00} v_1 (1 - e^{w_{11}})$$

Thus saddle points are found if $v_1 w_{11} \geq 0$ and if $v_1 w_{11} \leq 0$. So no local minima are found in this case.

Case 5.6.4: $w_{01} = w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty$ and $w_{02} = w_{02} + w_{12} = \infty$ and $w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = -\infty$

Here we find from equations (2.1) that $A_{00} = A_{10} = u + v_1 + v_2$ and that $A_{10} = A_{11} = u + v_1$. From (5.30) it follows that $R_{00} = -R_{10}$ and $R_{01} = -R_{11}$, and thus all terms A_{ij} are equal to zero and we find that $v_2 = 0$. Thus these points are saddle points.

Case 5.6.5: $w_{01} = w_{01} + w_{11} = \infty$ and $w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = -\infty$ and $w_{02} = w_{02} + w_{12} = \infty$ and $w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = -\infty$

In this case we find $A_{00} = A_{10} = u + v_1 + v_2 = 0$, $A_{11} = A_{01} = u = 0$. Thus we have $R_{00} = -R_{01} = -R_{10} = R_{11}$ and the same argument as in case 5.6.3 can be used to prove that no local minima are found in this case.

Case 5.6.6: $w_{01} = w_{01} + w_{11} = \infty$ and $w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = -\infty$ and $w_{02} = w_{02} + w_{22} = \infty$ and $w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = -\infty$.

From (5.30) it follows that $R_{00} = -R_{01} = -R_{10} = R_{11}$. So the same argument as in case 5.6.3 shows that no local minima occur in this case.

So only case 5.6.1 and case 5.6.3 and cases similar to these cases will result in local minima. In tables 10 and 11 these local minima are summarized.

All these regions with local minima have boundary points that are saddle points.

5.7 Some concrete examples

Lisboa and Perantonis [4] give 5 examples of points which are local minima. The fifth example is one with finite weights and is not a real local minimum but a saddle point, as proved in [10]. The other points are examples of local minima with some of the weights from the input units to the hidden units equal to plus or minus infinity. The points they have found have finite weights, but the error does not further decrease because of numerical saturation. The numerical saturation occurs as soon as the input of the hidden nodes results in a value of the transfer function $f(x)$ very close to 0 or 1. The four examples they give are shown in table12.

From the table it is clear that saturation of $f(x)$ occurs for $x < -11$ and $x > 11$. Examples 1 and 2 are of the category that exactly two patterns are learned and the other two are not. The first example is of the class (see table 3 in section 5.3):

$$A_{00} = u + v_1 f(w_{01}) + v_2 f(w_{02}) = f^{-1}(0.1)$$

Table 10: Local minima with one of the hidden nodes saturated for all patterns

error 0.32
$w_{01} = w_{01}+w_{11} = w_{01}+w_{11} = w_{01}+w_{11}+w_{21} = \infty$ w_{02} finite, $w_{12} = w_{22} = 0$ $u+v_1+v_2 f(w_{02}) = 0$ $v_1 w_{11} w_{21} < 0$ $\|v_2\| > 1.6\|f''(w_{02})\| / \{f'(w_{02})\}^2$
$w_{01} = w_{01}+w_{11} = w_{01}+w_{11} = w_{01}+w_{11}+w_{21} = -\infty$ w_{02} finite, $w_{12} = w_{22} = 0$ $u+v_2 f(w_{02}) = 0$ $v_1 w_{11} w_{21} > 0$ $\|v_2\| > 1.6\|f''(w_{02})\| / \{f'(w_{02})\}^2$
w_{01} finite, $w_{11} = w_{21} = 0$ $w_{02} = w_{02}+w_{12} = w_{02}+w_{22} = w_{02}+w_{12}+w_{22} = \infty$ $u+v_1 f(w_{01})+v_2 = 0$ $\|v_1\| > 1.6\|f''(w_{01})\| / \{f'(w_{01})\}^2$ $v_2 w_{12} w_{22} < 0$
w_{01} finite, $w_{11} = w_{21} = 0$ $w_{02} = w_{02}+w_{12} = w_{02}+w_{22} = w_{02}+w_{12}+w_{22} = -\infty$ $u+v_1 f(w_{01}) = 0$ $\|v_1\| > 1.6\|f''(w_{01})\| / \{f'(w_{01})\}^2$ $v_2 w_{12} w_{22} > 0$

$$A_{10} = u+v_1 f(w_{01}+w_{11}) + v_2 f(w_{02}+w_{12}) = f^{-1}(0.9)$$

$$A_{01} = A_{11} = u+v_2 = 0$$

$$w_{01}+w_{21} = w_{01}+w_{11}+w_{21} = -\infty,\ w_{02}+w_{22} = w_{02}+w_{12}+w_{22} = \infty$$

$$v_1 \le f^{-1}(0.1),\ w_{11} < 0,\ v_2 \le f^{-1}(0.1),\ w_{12} < 0$$

The second example belongs to the class (see table 5 in section 5.3):

$$A_{10} = u+v_1 f(w_{01}+w_{11}) + v_2 f(w_{02}+w_{12}) = f^{-1}(0.9)$$

$$A_{11} = u+v_1 f(w_{01}+w_{11}+w_{21}) + v_2 f(w_{02}+w_{12}+w_{22}) = f^{-1}(0.1)$$

$$A_{00} = A_{01} = u+v_1+v_2 = 0$$

$$w_{01} = w_{01}+w_{21} = \infty,\ w_{02} = w_{02}+w_{22} = \infty$$

Table 11: Local minima with both hidden units saturated for all patterns

error 0.32
$w_{01} = w_{01}+w_{11} = w_{01}+w_{21} = w_{01}+w_{11}+w_{21} = \infty$ $w_{02} = w_{02}+w_{12} = w_{02}+w_{22} = w_{02}+w_{12}+w_{22} = \infty$ $u+v_1+v_2 = 0$ $v_1w_{11}w_{21} < 0$ $v_2w_{12}w_{22} < 0$
$w_{01} = w_{01}+w_{11} = w_{01}+w_{21} = w_{01}+w_{11}+w_{21} = \infty$ $w_{02} = w_{02}+w_{12} = w_{02}+w_{22} = w_{02}+w_{12}+w_{22} = -\infty$ $u+v_1 = 0$ $v_1w_{11}w_{21} < 0$ $v_2w_{12}w_{22} > 0$
$w_{01} = w_{01}+w_{11} = w_{01}+w_{21} = w_{01}+w_{11}+w_{21} = -\infty$ $w_{02} = w_{02}+w_{12} = w_{02}+w_{22} = w_{02}+w_{12}+w_{22} = \infty$ $u+v_2 = 0$ $v_1w_{11}w_{21} > 0$ $v_2w_{12}w_{22} < 0$
$w_{01} = w_{01}+w_{11} = w_{01}+w_{21} = w_{01}+w_{11}+w_{21} = -\infty$ $w_{02} = w_{02}+w_{12} = w_{02}+w_{22} = w_{02}+w_{12}+w_{22} = -\infty$ $u = 0$ $v_1w_{11}w_{21} > 0$ $v_2w_{12}w_{22} < 0$

$$v_1 \geq f^{-1}(0.9),\ w_{21} < 0,\ v_2 \leq f^{-1}(0.1),\ w_{22} > 0$$

Examples 3 and 4 have exactly one pattern learned. The third example belongs to the class (see conclusion 5.10 and/or table 6 in section 5.5):

$$A_{00} = u + v_1 f(w_{01}) + v_2 f(w_{02}) = f^{-1}(0.1)$$

$$A_{01} = A_{10} = A_{11} = u + v_2 = f^{-1}(1.9/3)$$

$$w_{01} + w_{11} = w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = -\infty,$$

$$w_{02} + w_{12} = w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = \infty$$

w_{01} and w_{02} finite, $v_1 < 0$, $v_2 > 0$

The fourth example belongs to the class (see table 7 in section 5.5):

Table 12: The local minima found by Lisboa and Perantonis

	Example 1	**Example 2**	**Example 3**	**Example 4**
$f(A_{00})$	0.1	0.5	0.1	0.366
$f(A_{01})$	0.5	0.5	0.633	0.9
$f(A_{10})$	0.9	0.9	0.633	0.366
$f(A_{11})$	0.5	0.1	0.633	0.366
u	5.05670	1.54884	0.10897	–0.54656
v_1	–2.78335	4.59262	–3.42122	0.66331
v_2	–5.05670	–6.14146	0.43758	4.23784
w_{01}	1.41913	12.59865	1.39427	–10.28005
w_{11}	–5.52058	–11.52144	–13.70896	–10.42464
w_{21}	–13.69016	–1.10568	–13.70896	8.55786
w_{02}	4.73579	11.70733	5.121702	–10.97265
w_{12}	–4.50867	–11.89991	6.01491	–12.71414
w_{22}	12.27468	4.01044	6.01491	11.47785

$$A_{01} = u + v_1 f(w_{01} + w_{21}) + v_2 f(w_{02} + w_{22}) = f^{-1}(0.9)$$

$$A_{00} = A_{10} = A_{11} = u = f^{-1}(1.1/3)$$

$$w_{01} = w_{01} + w_{11} = w_{01} + w_{11} + w_{21} = -\infty$$

$$w_{02} = w_{02} + w_{12} = w_{02} + w_{12} + w_{22} = -\infty$$

$w_{01} + w_{21}$ and $w_{02} + w_{22}$ finite and $v_1 > 0$, $v_2 > 0$

Rumelhart and McClelland [6] give the following example of a local minimum:

$u = -0.8$, $v_1 = -4.5$, $v_2 = 5.3$, $w_{01} = 2.0$, $w_{11} = -2.0$, $w_{21} = 9.2$, $w_{02} = -0.1$, $w_{12} = 4.3$, $w_{22} = 9.2$,

resulting in the outputs: $f(A_{00}) = 0.096$, $f(A_{01}) = 0.4999$, $f(A_{10}) = 0.898$ and $f(A_{11}) = 0.500$.

This local minimum belongs to the class described in table 3 in section 5.3, satisfying the conditions:

$$A_{00} = u + v_1 f(w_{01}) + v_2 f(w_{02}) = f^{-1}(0.1)$$
$$A_{10} = u + v_1 f(w_{01} + w_{11}) + v_2 f(w_{02} + w_{12}) = f^{-1}(0.9)$$
$$A_{01} = A_{11} = u + v_1 + v_2 = 0$$
$$w_{01} + w_{21} = w_{01} + w_{11} + w_{21} = \infty, w_{02} + w_{22} = w_{02} + w_{12} + w_{22} = \infty$$
$$v_1 \le f^{-1}(0.1), w_{11} < 0, v_2 \ge f^{-1}(0.9), w_{12} > 0$$

6 Conclusions

In this paper it is proved that the error surface of the two-layer XOR network with two hidden units has a number of regions with local minima. These regions of local minima occur for combinations of the weights from the inputs to the hidden nodes such that one or both hidden nodes are saturated (give output 0 or 1) for at least two of the patterns. However, boundary points of these regions of local minima are saddle points. From these results it can be concluded that from each finite point in weight space a strictly decreasing path exists to a point with error zero. In the neighbourhood of a region of local minima, this path will combine a decreasing step towards the local minimum and a step “parallel” to the region of local minima towards the boundary points that are saddle points. Section 5.3, 5.5, and 5.6 contain a summary of all regions of local minima for the considered XOR network.

The results of this paper can explain a number of experimentally found results. For example the fact that with a higher numerical precision less “local minima” are met. This fact can be explained by the observation that with a higher numerical precision the learning algorithm will less soon get stuck and cannot proceed further because of the saturation of the hidden units. As long as the learning algorithm makes some movement in the neighbourhood of a region of local minima, the possibility exists that a point near the boundary of such a region will be met, and a path to the global minimum will be found. Testing this idea we started in the four “local minima” given by Lisboa and Perantonis (see section 5.7) using on-line backpropagation and “double” precision. As a result, in three of the four cases the algorithm found at last a point with error zero, so the algorithm escaped from the neighbourhood of the region with local minima. In the fourth case the path chosen by the algorithm leaded away from the boundary points which are saddle points.

7 References

[1] E.K. Blum; “Approximation of Boolean Functions by Sigmoidal Networks: Part I: XOR and other Two-variable functions”. In: *Neural Computation* 1, 532-540, 1989.

[2] L.G.C. Hamey; “The analysis of local minima in feed-forward neural networks”. *Neural Networks*, 1994. (submitted)

[3] L.G.C. Hamey; *Analysis of the Error Surface of the XOR Network with Two Hidden Nodes*, Computing Report 95/167C, Department of Computing, Macquarie University, NSW 2109 Australia, 1995. Accessible by FTP from ftp.mpce.mq.edu.au in pub/comp/techreports/950167.hamey.ps.

[4] P.J.G. Lisboa and S.J. Perantonis; “Complete solution of the local minima in the XOR problem”, *Network 2*, pp. 119–124, 1991.

[5] L. Prechelt; *A study of Experimental Evaluations of Neural Network Learning Algorithms: Current Research Practice*, Technical Report 19/94, Fakultät für Informatik, Universität Karlsruhe, 1994.

[6] D.E. Rumelhart, J.L. McClelland and the PDP Research Group; *Parallel Distributed Processing, Volume 1*. The MIT Press, Cambridge, Massachusetts, 1986.

[7] I.G. Sprinkhuizen-Kuyper and E.J.W. Boers; “The Error Surface of the simplest XOR Network has only global Minima”. *Neural Computation* Vol. 8, nr. 6, 1996. An earlier version is Technical Report 94-21, Dept. of Computer Science, Leiden University, The Netherlands.

[8] I.G. Sprinkhuizen-Kuyper and E.J.W. Boers; “Classification of all stationary points on a neural network error surface”. In: J.C. Bioch and S.H. Nienhuys-Cheng, eds., *Proceedings of Benelearn-94: 4th Belgian-Dutch Conference on Machine Learning*, pp. 192-201, Report EUR-CS-94-05, Dept. of Computer Science, Fac. of Economics, Erasmus University, Rotterdam, The Netherlands, June 1994. Also as Technical Report 94-19, Dept. of Computer Science, Leiden University, The Netherlands.

[9] I.G. Sprinkhuizen-Kuyper and E.J.W. Boers; *A Comment on a Paper of Blum: Blum’s “local minima” are saddle points*. Technical Report 94-34, Dept. of Computer Science, Leiden University, The Netherlands.

[10] I.G. Sprinkhuizen-Kuyper and E.J.W. Boers; *The Error Surface of the 2-2-1 XOR Network: The finite stationary Points.* Technical Report 95-39, Dept. of Computer Science, Leiden University, The Netherlands.