

① 由于模型可辨认性 (model identifiability) 问题，神经网络和任意具有多个等效参数化潜变量的模型都会具有多个局部极小值。② 如果一个足够大的训练集可以唯一确定一组模型参数，那么该模型被称为可辨认的。③ 带有潜变量的模型通常是不可辨认的，因为通过相互交换潜变量我们能得到等价的模型。例如，考虑神经网络的第一层，我们可以交换单元 i 和单元 j 的传入权重向量、传出权重向量而得到等价的模型。如果神经网络有 m 层，每层有 n 个单元，那么会有 $n!^m$ 种排列隐藏单元的方式。这种不可辨认性被称为权重空间对称性 (weight space symmetry)。

除了权重空间对称性，很多神经网络还有其他导致不可辨认的原因。例如，在任意整流线性网络或者maxout网络中，我们可以将传入权重和偏置乘以 α 倍，然后将传出权重乘以 $\frac{1}{\alpha}$ 倍，而保持模型等价。这意味着，如果代价函数不包括如权重衰减这种直接依赖于权重而非模型输出的项，那么整流线性网络或者maxout网络的每一个局部极小点都在等价的局部极小值的 $(m \times n)$ 维双曲线上。



