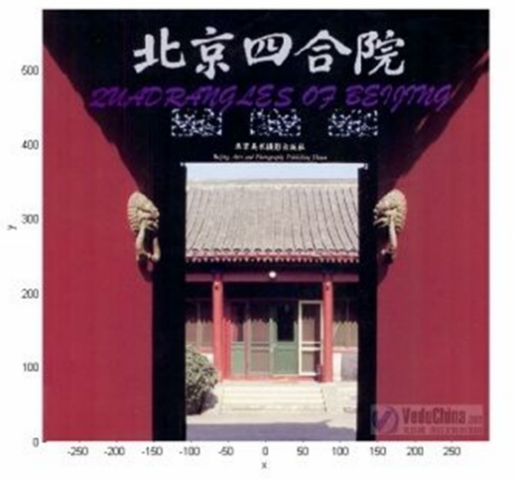
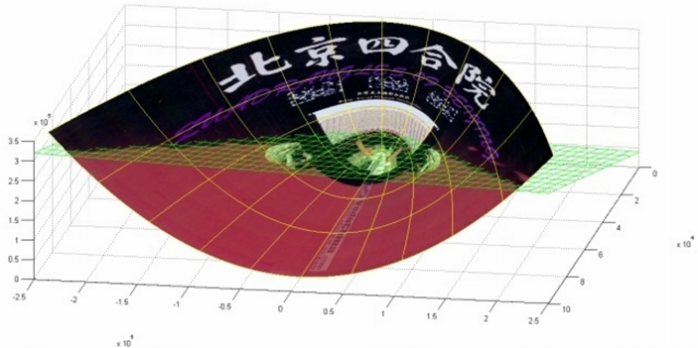
基于2D-KPCA 的拉普拉斯特征映射的量子机器学习

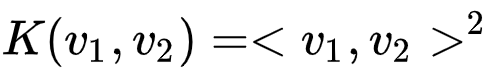
当数据集在原始特征中不是线性可分的时候，引入核函数的策略：通过映射函数将原始特征空间映射为更高维的空间，在原始空间中不可分的数据在高维空间中可能变成线性可分。

核函数的作用：隐含着一个从低维空间到高维空间的映射，而这个映射可以把低维空间中线性不可分的两类点变成线性可分的。

用一张图片直观地解释这一思想。

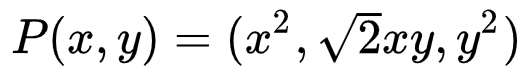
 

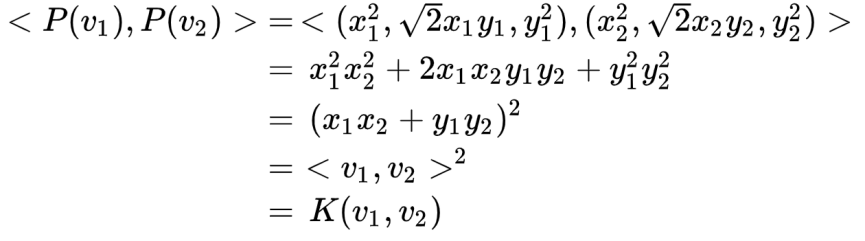
考虑一种核函数

（即内积平方）

点坐标

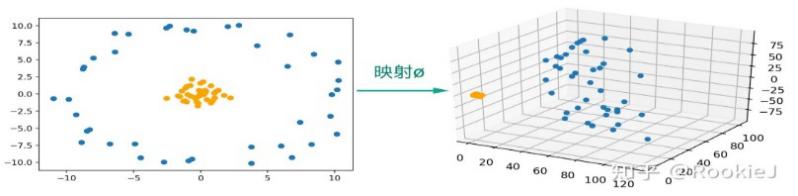
这个核函数对应着一个二维空间到三维空间的映射





1. 思想来源

非线性问题转化为更高维度空间上的线性可分方法



2、与PCA相比的创新部分：

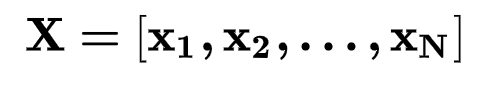
①为了更好地处理非线性数据，引入非线性映射函数Φ，将原空间中的数据映射到高维空间，这个Φ可以是隐性的，我们不知道，也不需要知道它的具体形式是啥。

②引入了一个定理：空间中的任一向量（哪怕是基向量），都可以由该空间中的所有样本线性表示（局部可以有整体表示），这点对KPCA很重要！

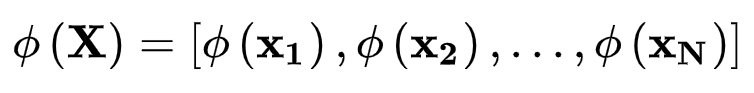
1. 数学模型

假设中心化后的样本集合X（d\*N，N个样本，维数d维，样本按列排），现

将X映射到高维空间，得到Φ(𝑥), 即由：



到

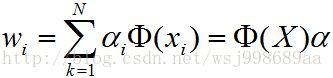


假设D（D >> d）维向量，𝑤\_𝑖 (𝑖=1,…,𝑑),为高维空间中的特征向量, 𝜆\_𝑖 (𝑖=1,…,𝑑)

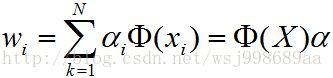
为对应的特征值，高维空间中的PCA如下：



利用②的定理(空间中的任一向量（哪怕是基向量），都可以由该空间中的所有样本线性表示)，将特征向量𝑤\_𝑖 (𝑖=1,…,𝑑), 利用样本集合Φ(𝑥)线性表示









左右对称，构造



定义矩阵,则K为𝑁∗𝑁的对称半正定矩阵，其𝑖行𝑗列的元素为





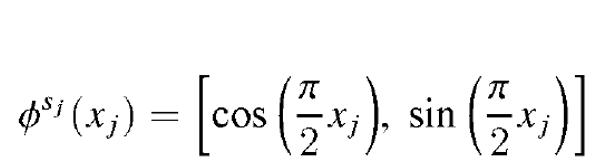
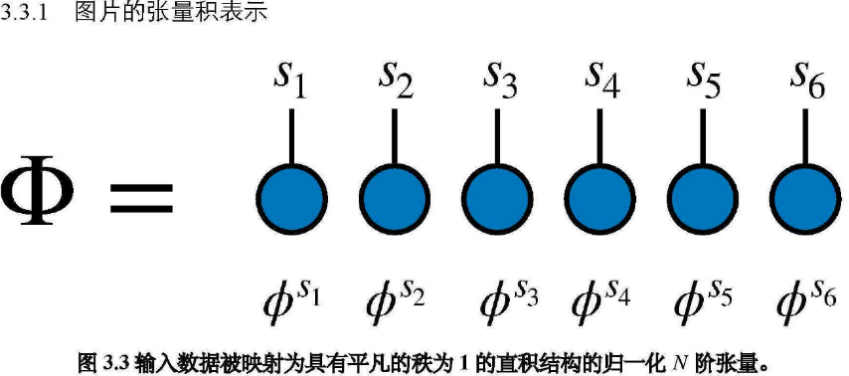
因为矩阵K中的元素可由得到，不需要显式地定义映射，而只需定义特征空间中向量的点积.定义函数 则称为核函数

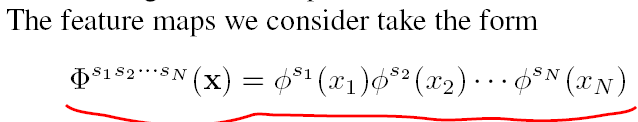
1. 缺点

核函数 K 维数易受图像矩阵维数影响，当图像矩阵规模大时，K2DPCA 会得到一个大规模的核函数，存储及计算麻烦，且导致识别速度降低。

1. 改进

利用约化密度矩阵对图片进行特征提取得到一个可以处理的核函数。（主要改进就是先初始化数据，然后再按照3的模型处理）

1. 像素点表示为
2. 



1. 将图片转换为张量积表示除了帮助简单地我们将图片输入张量网络

模型。我们还可以在这个表示之下利用张量网络中约化密度矩阵的技术对图片进行局域的特征提取。从而对图片信息进行有效地压缩和粗粒化。具体的实现方法和张量网络中最经典的ＤＭＲＧ（密度矩阵重组）算法利用约化密度矩阵做最优截断非常像。

**定义1(系统的密度算子)** 设量子系统以概率处于状态, 称 为一个系综(ensemble), 其中, 且, 系统的密度算子为



它是一个迹为1的半正定厄米算子。

**约化密度算子**

**出发点**： 对于一个大的量子体系而言, 我们感兴趣的物理量只与体系的一部分有关。约化密度算子是分析复合量子系统不可缺少的工具。

**定义2(约化密度算子)** 假设有物理系统A和B, 其状态由密度算子描

述, 针对系统A和B的约化密度算子定义为：

其中是一个算子映射, 称为在系统B上的偏迹(partial trace)。

**定义3(偏迹)** 偏迹定义为：



 一般来说, 复合系统处于纯态, 其子系统也可能为混合态。例如前面提到的Bell

态是一个两量子比特系统纯态



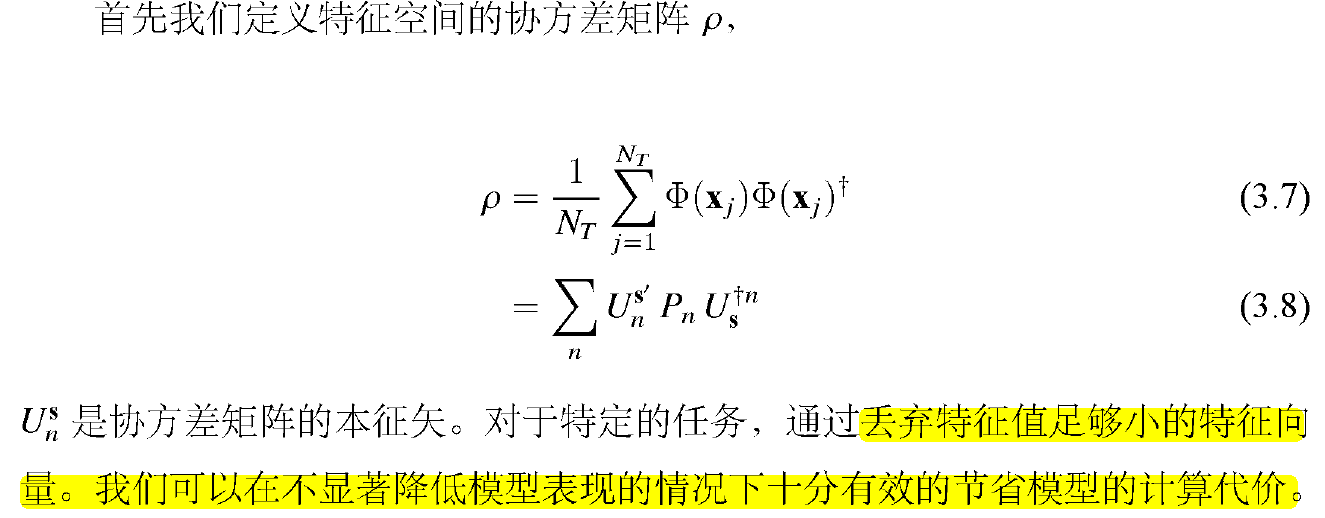
其密度算子为

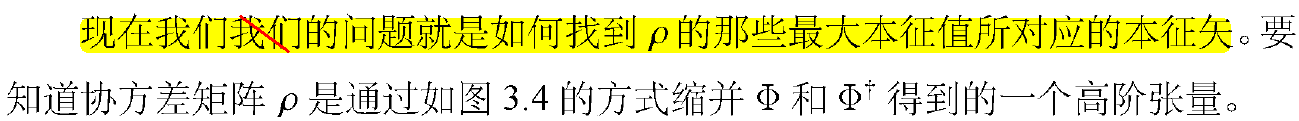


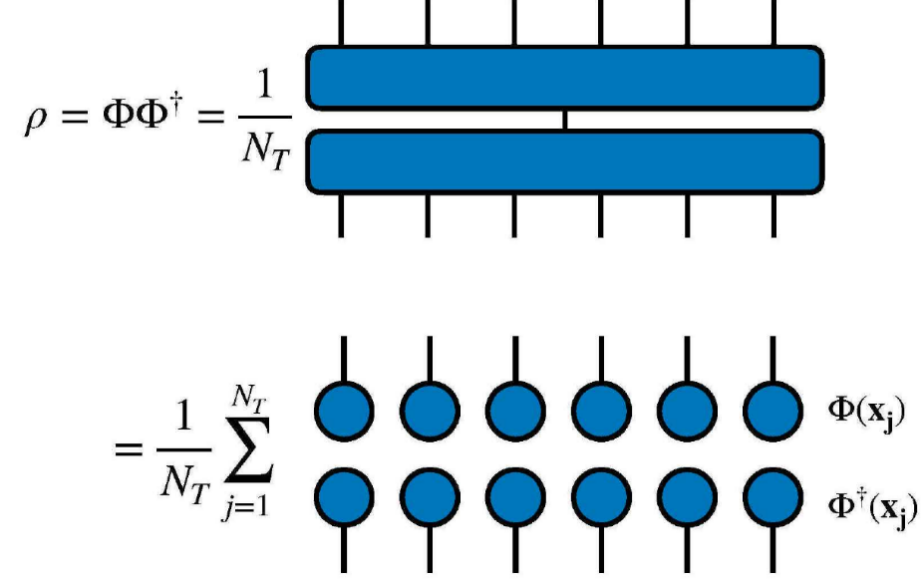


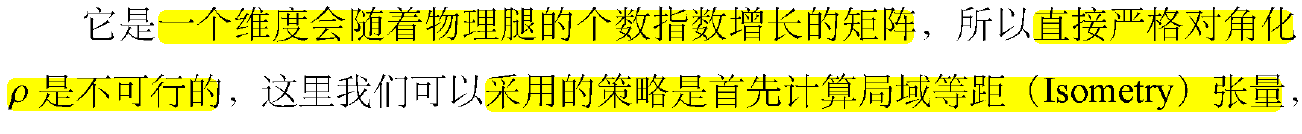
描述量子比特1的密度算子为

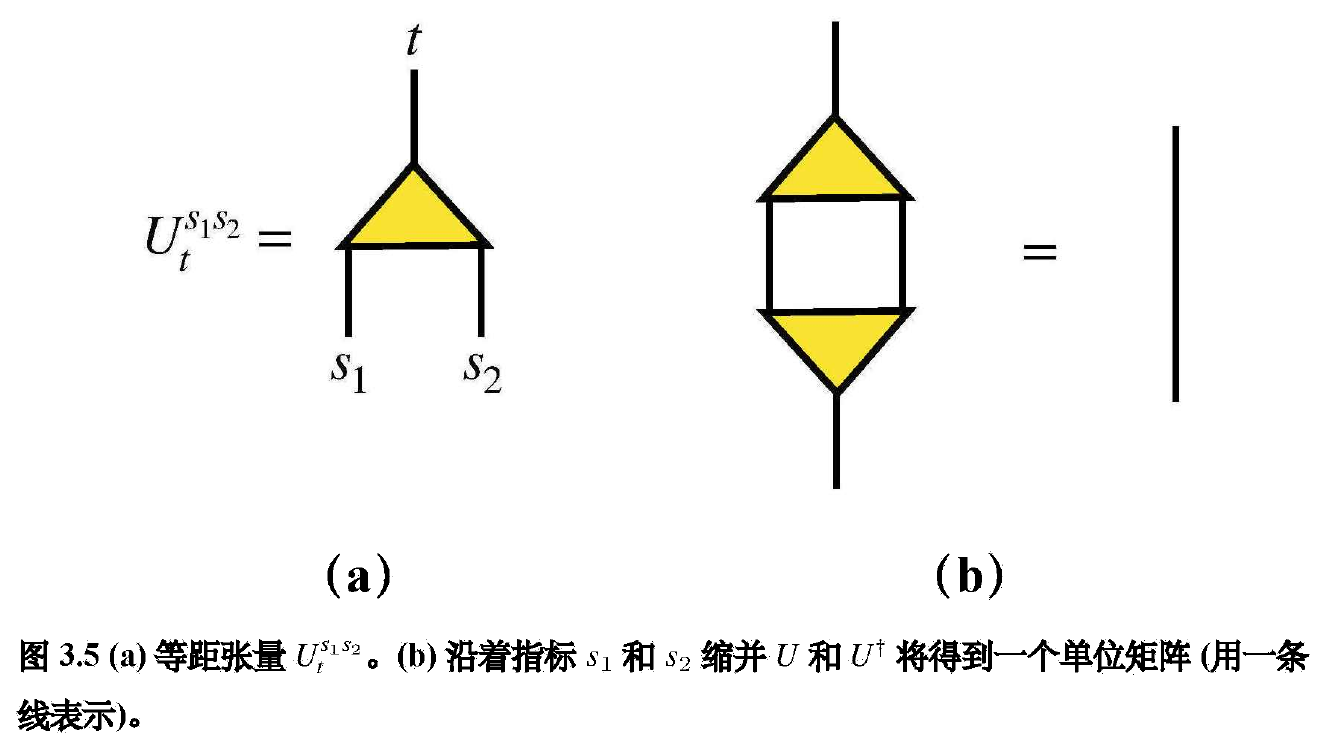


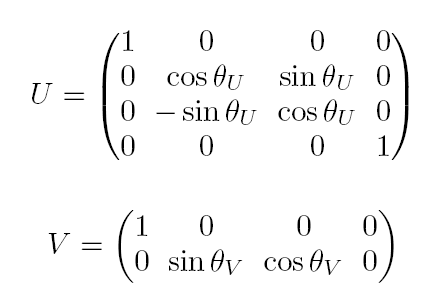


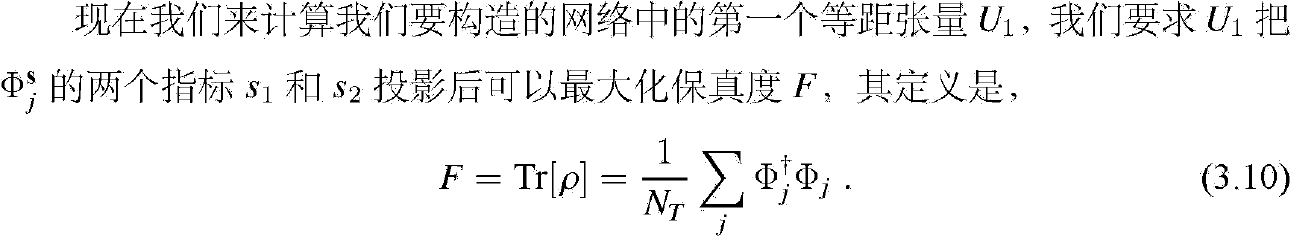


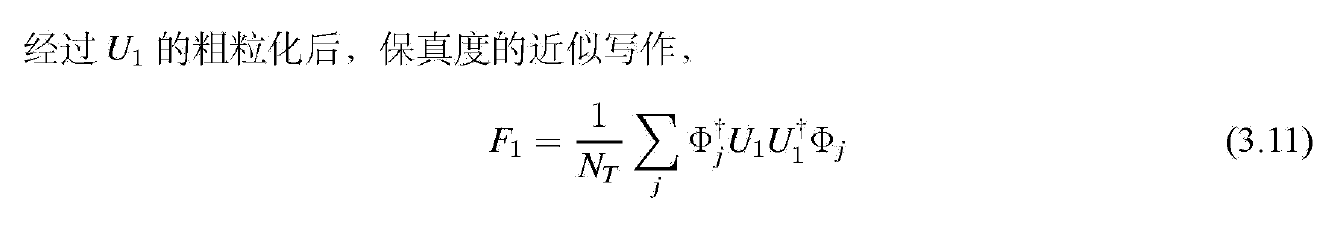


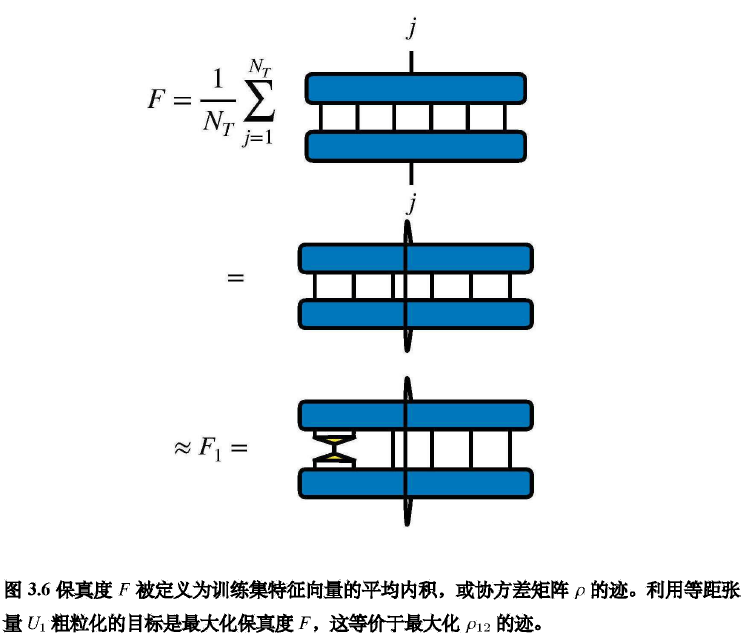


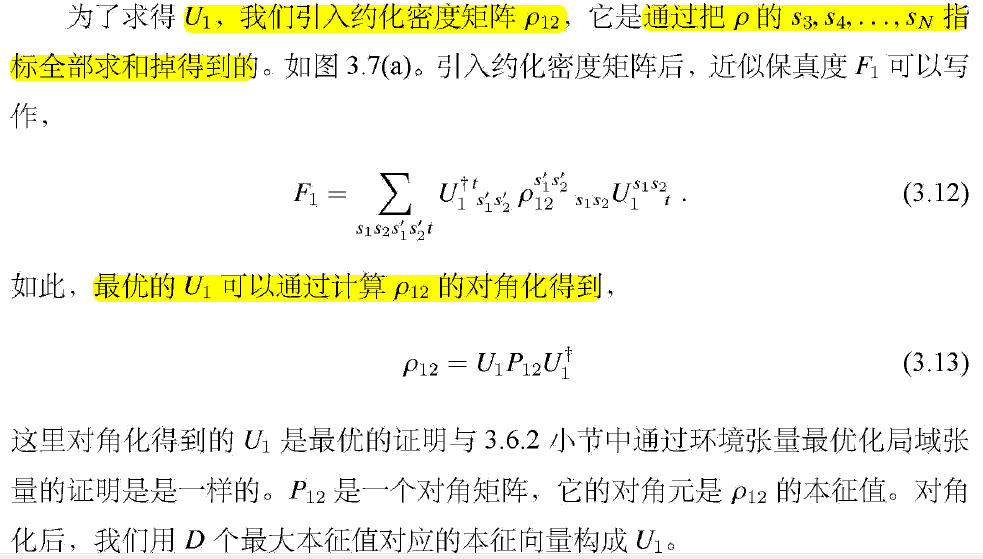


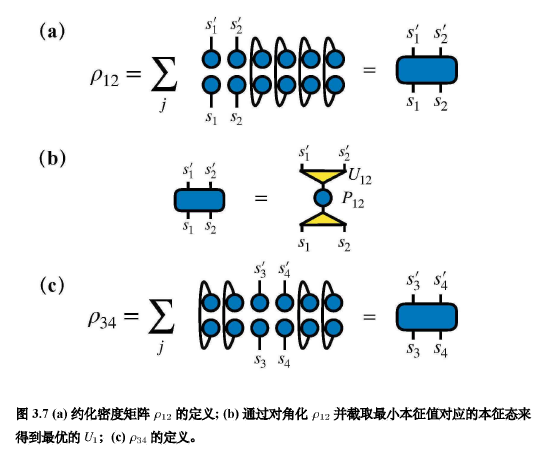


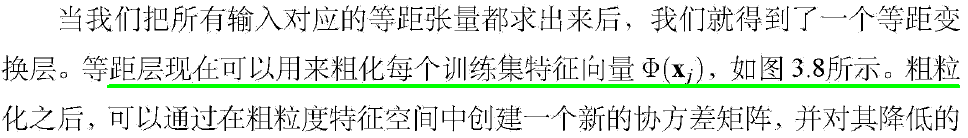


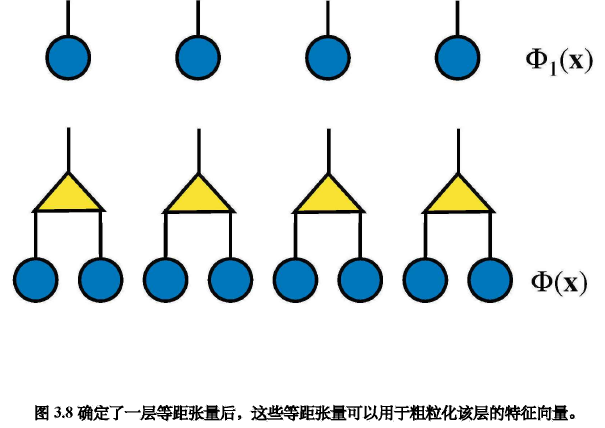


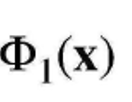
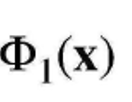










**最后将作为训练数据带入模型3，当然也可以用同样的方法压缩，直到得到一个可处理的数据。**