降维

特征降维(Dimensionality Reduction)在机器学习中有广泛的应用，比如文件压缩(Compression)、数据预处理(Preprocessing)等。最常见的降维算法有主成分分析法(Principal components analysis，简称 PCA)，通过对协方差矩阵进行特征分解而得到数据的主要成分，但是 PCA 本质上是一种线性变换，提取特征的 能力极为有限。

# 主成分分析（PCA）

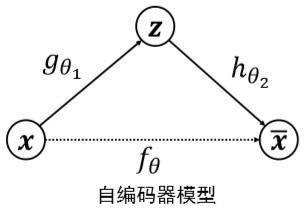
# 自编码器（Auto Encoder）

考虑有监督学习中神经网络的功能：

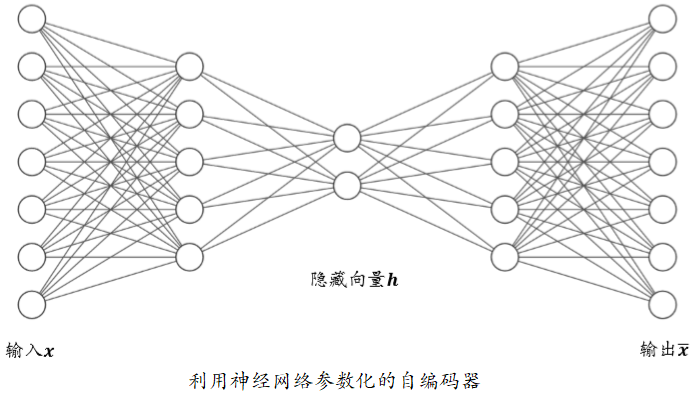
其中，是输入的特征向量维度，是网络输出的向量长度。对于分类问题，网络模型通过把长度为的输入特征向量𝒙变换到长度为的输出向量𝒐，这个过程可以看成是特征降维的过程，把原始的高维输入向量𝒙变换到低维的变量𝒐。

利用神经网络的强大非线性表达能力去学习到低维的数据表示，问题的关键在于，训练神经网络一般需要一个显式的标签数据(或监督信号)，但是无监督的数据没有额外的标注信息，只有数据𝒙本身。

因此，需要利用数据𝒙本身作为监督信号来指导网络的训练，即希望神经网络能够学习到映射，需要把切分成两部分，前面的子网络尝试学习映射关系:，后面的子网络尝试学习映射关系，如下图所示。把看成一个数据编码(Encode)的过程，把高维度的输入𝒙编码成低维度的隐变量𝒛(Latent Variable，或隐藏变量)，称为Encoder 网络(编码器)；看成数据解码(Decode)的过程，把编码过后的输入𝒛解码为高维度的𝒙，称为Decoder网络(解码器)。

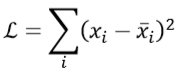


编码器和解码器共同完成了输入数据𝒙的编码和解码过程，整个网络模型叫做自动编码器(Auto-Encoder)，简称自编码器。如果使用深层神经网络来参数化和函数，则称为深度自编码器(Deep Auto-encoder)，如下图所示。



自编码器能够将输入变换到隐藏向量𝒛，并通过解码器重建(Reconstruct，或恢复)出𝒙，即𝒙 ≈ 𝒙，那么，自编码器的优化目标可以写成：

其中表示和的距离度量，称为重建误差函数。最常见的度量方法有欧氏距离 (Euclidean distance)的平方，计算方法如下：



它和均方误差原理上是等价的。自编码器网络和普通的神经网络并没有本质的区别，只是训练的监督信号由标签𝒚变成了自身𝒙。借助于深层神经网络的非线性特征提取能力，自 编码器可以获得良好的数据表示，相对于PCA等线性方法，自编码器性能更加优秀，甚至可以更加完美的恢复出输入𝒙。

# 自编码器变种

一般而言，自编码器网络的训练较为稳定，但是由于损失函数是直接度量重建样本与真实样本的底层特征之间的距离，而不是评价重建样本的逼真度和多样性等抽象指标，因此在某些任务上效果一般，如图片重建，容易出现重建图片边缘模糊，逼真度相对真实图片仍有不小差距。为了尝试让自编码器学习到数据的真实分布，产生了一系列的自编码器变种网络。下面将介绍几种典型的自编码器变种模型。

## Denoising Auto-Encoder

为了防止神经网络记忆住输入数据的底层特征，Denoising Auto-Encoders给输入数据添加随机的噪声扰动，如给输入𝒙添加采样自高斯分布的噪声𝜀：

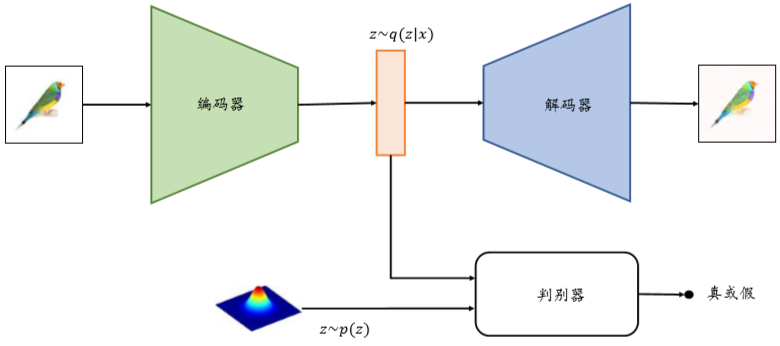
添加噪声后，网络需要从学习到数据的真实隐藏变量z，并还原出原始的输入𝒙。模型的优化目标为：

## Dropout Auto-Encoder

自编码器网络同样面临过拟合的风险，Dropout Auto-Encoder 通过随机断开网络的连接 来减少网络的表达能力，防止过拟合。Dropout Auto-Encoder 实现非常简单，通过在网络层中插入 Dropout 层即可实现网络连接的随机断开。

## Adversarial Auto-Encoder

为了能够方便地从某个已知的先验分布中𝑝(𝒛)采样隐藏变量𝒛，方便利用𝑝(𝒛)来重建输 入，对抗自编码器(Adversarial Auto-Encoder)利用额外的判别器网络(Discriminator，简称D网络)来判定降维的隐藏变量𝒛是否采样自先验分布𝑝(𝒛)，如下图所示。判别器网络的输出为一个属于[0,1]区间的变量，表征隐藏向量是否采样自先验分布𝑝(𝒛)：所有采样自先 验分布𝑝(𝒛)的𝒛标注为真，采样自编码器的条件概率𝑞(𝒛|𝒙)的𝒛标注为假。通过这种方式训练，除了可以重建样本，还可以约束条件概率分布𝑞(𝒛|𝒙)逼近先验分布𝑝(𝒛)。

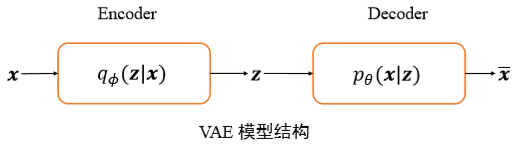


对抗自编码器是从生成对抗网络算法衍生而来。

## 变分自编码器

基本的自编码器本质上是学习输入𝒙和隐藏变量𝒛之间映射关系，它是一个判别模型 (Discriminative model)，并不是生成模型(Generative model)。那么能不能将自编码器调整为生成模型，方便地生成样本呢？

给定隐藏变量的分布P(𝒛)，如果可以学习到条件概率分布P(𝒙|𝒛)，则通过对联合概率分布P(𝒙,𝒛) = P(𝒙|𝒛)P(𝒛)进行采样，生成不同的样本。变分自编码器(Variational AutoEncoders，简称 VAE)就可以实现此目的，如下图所示。



从神经网络的角度VAE，则VAE和前面的自编码器一样，非常好理解；但是VAE的理论推导稍复杂，接下来先从神经网络的角度去阐述 VAE，再从概率角度去推导VAE。

这部分推导过于复杂，建议看原书（深度学习TensorFlow，12.4.1节 VAE原理）

参考：

龙书。TensorFlow深度学习