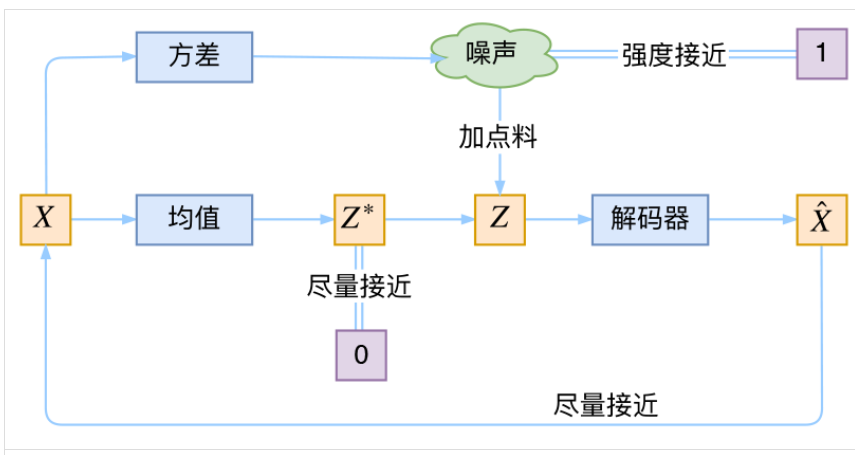
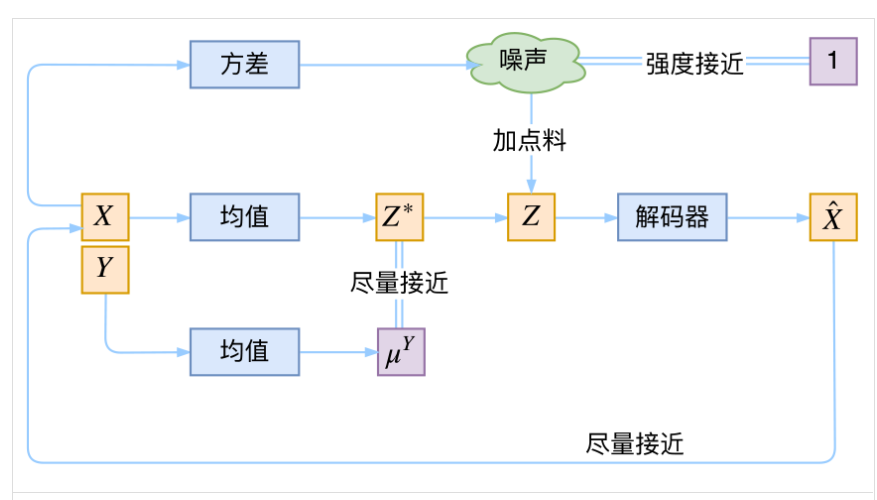
简单来说，VAE就是为每个输入拟合一个专属的正态分布(μ，σ)，loss中的一部分的目的就是将所有输入的正态分布都训练成(0,1)。其中P(z)不是正态分布，p(Z|X)是正态分布！也就是说拟合出来的均值和方差是拟合P(z|x)，间接回答了为什么是专属的正态分布。

单独看每个输入就是这样的：

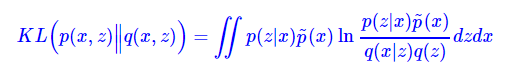


如果加入分类，就变成这样了：

也就是每个类别最后的均值不同，这也是“为什么要训练成(0,1)”的一个补充，我的理解在VAE中方差是噪声，类似对抗训练，而均值是归一到某个度量上，不知道说的对不对。

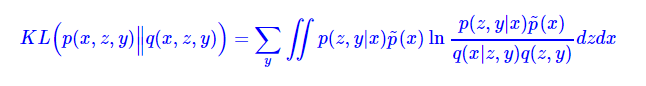


加入聚类的VAE



普通VAE的loss：

Z为隐藏层状态，即编码向量。



加入聚类功能后的loss：

也就是用（z,y）换掉了z，换言之，在隐藏层不仅仅编码z同时编码y，y就是类别的离散变量。最终y即相当于聚类后得到的向量。

为什么觉得这样的聚类比k-means这样的好，因为在VAE编码过程中不断的最小化先验分布和最大化互信息，后者互信息很重要。