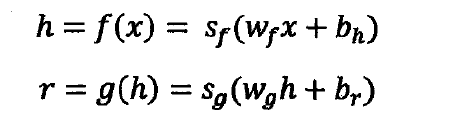
目前进度汇报

因为我对于深度学习还是刚接触，完全是零基础的。所以现在还是处于学习的进度。目前是在阅读冯方向的硕士论文——基于深度学习的多模态检索研究，就说一下现在我的了解和认识。

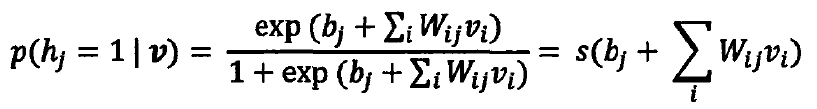
所谓的多模态是指从不同的模态，即信息的表现方式（文字、图片、视频、音频等），能够找到我们想要的其他模态。目前的功能预想是以文字搜索图片，以图片搜索文字。那么就要建立不同模态的关联，有两种策略。一种是建立共享层→模态关联，另一种是建立公共表示空间。

一、无监督学习：单层组件的栈式叠加。

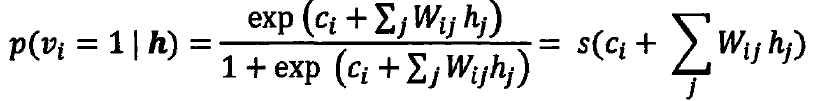
自编码器AE：我认为是输入（n目）→表示→重构层（n目）。输入x→编码器f（x）→表示h→解码器g（h）→重构r。其中

利用s函数是logistic激活函数。

限制玻尔兹曼机RBM：概率图模型。有输入层v和表示层h，其中v和h之间有权重w，c表示输入层的偏置值，b表示表示层的偏置值。论文是以二值随机变量来说明的（v，h属于{0,1}）。那么可以用公式：

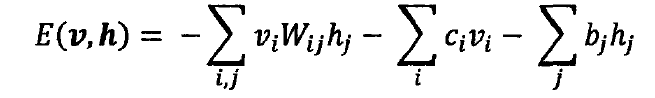


进行对输入v进行编码变成表示层h。同样，也可以用公式：

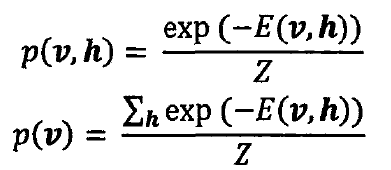


对表示层h解码得到原来的输入v。

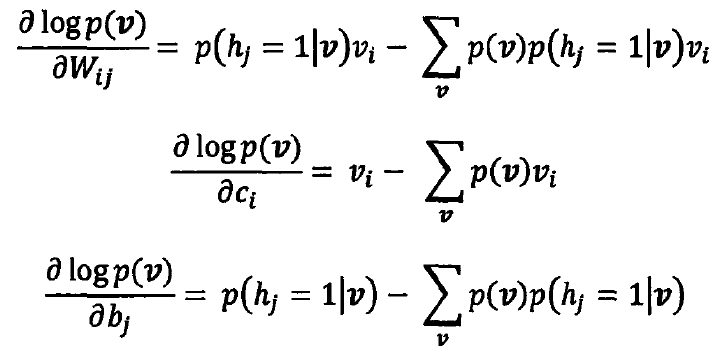
能量模型：用于拟合输入数据，确定变量之间的联系程度。能量模型也要捕获变量之间的相关性，变量之间的相关程度决定了能量的高低。把变量的相关关系用图表示出来，并引入概率测度方式就构成了概率图模型的能量模型。论文给出的能量函数：



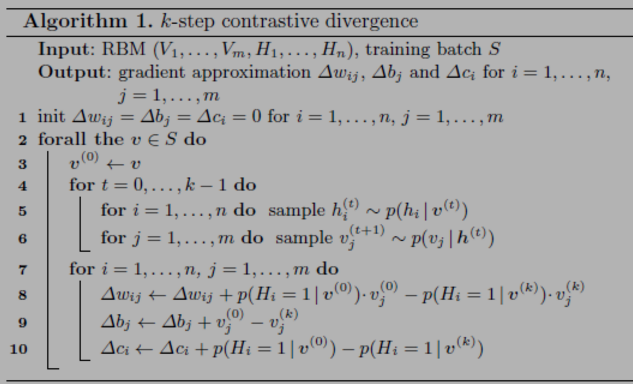
它用于二值随机变量，那么可以从v和h得到它们连接的能量。有了能量函数，就能描述两个节点的联合概率：



分母z好像是用来归一化的。那么有了联合概率分布，就要求一个最大的似然函数，论文给出的公式：（这个没怎么看明白）

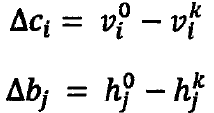


论文中是用了一个CD-k算法来进行抽样得到期望，下面是我找到的CD-k算法的伪代码：



这个算法大概就是开始时初始化v0为输入样本，然后用条件概率分布来抽样得到接下来的h0.。我在网上看到的解释：在CD算法一开始，可见单元的状态就被设置为一个训练样本，并用上面的几个条件概率来对隐藏节点的每个单元都从{0,1}中抽取到相应的值，然后再利用 来对可视节点的每个单元都从{0,1}中抽取相应的值，这样就得到了v1了，一般v1就够了，就可以拿来估算梯度了。那么进行了抽样后就可以计算参数的值了，论文中如此计算：

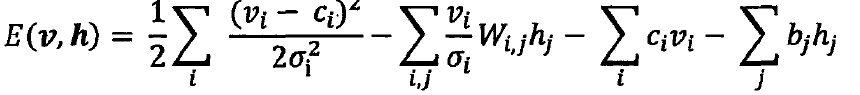




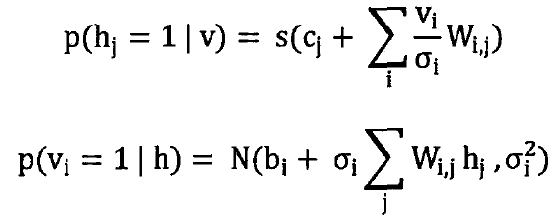
总觉得和CD-k算法有点不同。

扩展限制玻尔兹曼机：改变能量函数就可以建模各种数据分布。原因是显示数据不可能像二值随机分布那么简单，肯定是很复杂的，所以我们要对玻尔兹曼机进行扩展。

高斯限制玻尔兹曼机：能量函数：

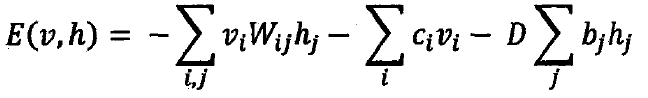


条件概率分布如下;



可以用这个GRBM来建立输入特征。

离散值时：



1. 深层模型

基于自编码器的深层模型：

栈式自编码器SAE：由底向上得到表示层，逐个依赖前一个的表示层。

深层自编码器DAE：即SAE训练完毕后再镜像展开所有的自编码器。

基于RBM的深层模型：

深度信念网络DBN：先随机初始化最顶部hn，然后用Gibbs采样，得到hn-1，一次用这个方法最终得到输入x。

深度玻尔兹曼机DBM：有自顶向下和自底向上两种，表达能力强。