# 四种模型对比总结

## 符号说明

 ：词语实例的数目。

 ：句子实例的数目。

： 表示文本中的第个词语

 ： 表示个句子和对应个人工标注的词性序列。

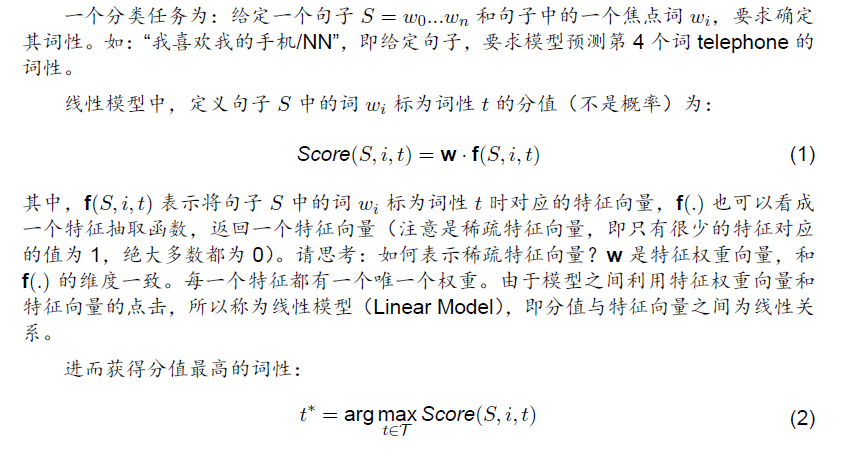
 ：表示第个句子，由个词语构成。

 ：表示第个句子对应的词性序列。

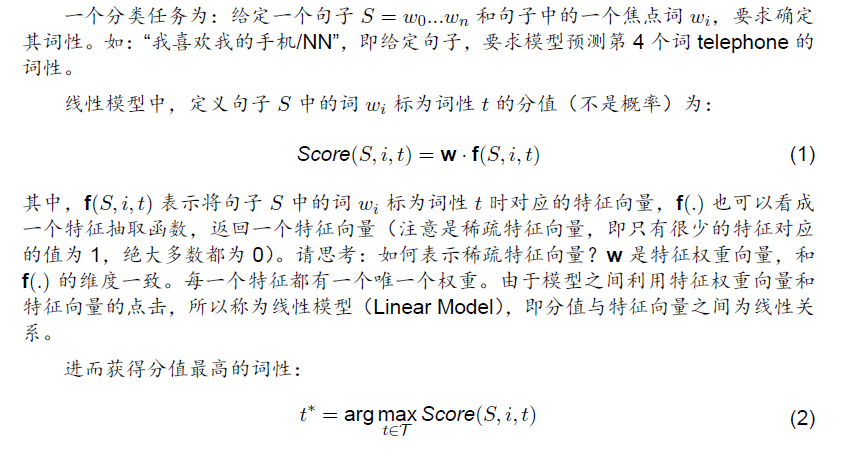
## 线性模型（Linear Model）

线性模型将句子中的每个词孤立地来看待：

1. 得分



1. 预测

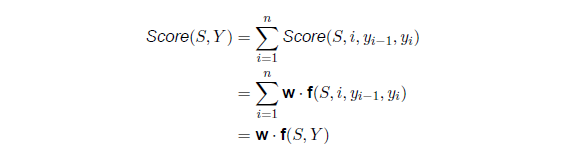


## 全局线性模型（Global Linear Model）

进一步地，考虑到句子的整体性，我们将句子当作一个单位，而词只是其中的部分结构，得到了全局线性模型。

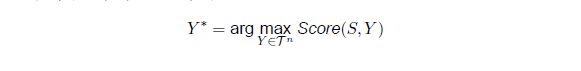
1. 得分

全局限性模型的得分计算的是整个句子的得分：



1. 预测

预测阶段也是对整个句子的词性序列进行预测：



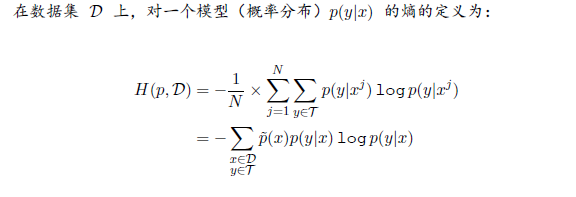
具体的计算方法可以用Viterbi算法。

## 对数线性模型（Log Linear Model）

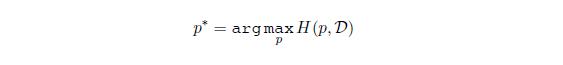
对数线性模型的理论基础来自于最大（条件）熵理论，以期在已知信息的条件下，得到最均衡的条件概率，而这也正是最符合事实的，由于自然世界的事务在无干预的条件下总是朝着熵增的方向变化。具体的模型如下：

（1） 最大熵

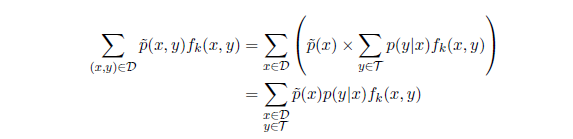
定义条件熵函数：



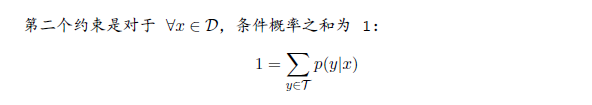
目标函数则为



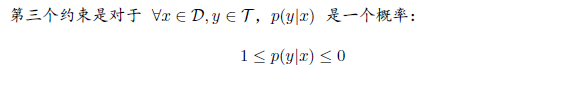
约束1：



约束2：



约束3：

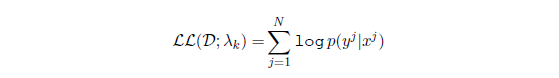


（2）最大似然估计

根据对偶问题理论：

问题可以转化为求最大似然估计的问题：

似然函数定义为：



目标函数（最大似然函数）：

C:\Users\Administrator\AppData\Roaming\Tencent\Users\1536946657\TIM\WinTemp\RichOle\[{V~6176T(NHW6]C`_(_7_2.png

注：该对偶问题的理论的具体证明请参见李航老师的《统计机器学习》对于最大熵和最大似然估计及其中对偶问题的证明过程。

## 条件随机场（CRF）

同样的，对数线性模型的一点不足就是只是将每个词语孤立地考虑，当我们将一个句子当作一个基本的单位时，就有了CRF（条件随机场）模型。

条件熵为：



目标函数为：



与对数线性模型相似，该问题也存在对偶问题：

似然函数：



目标函数：

