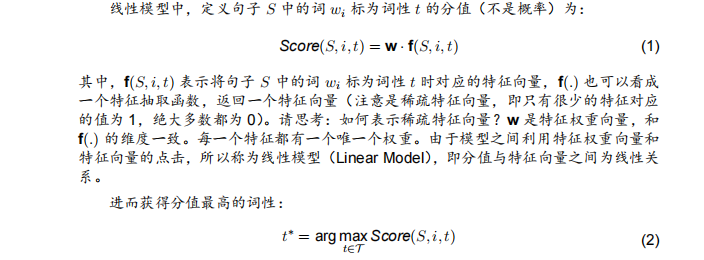
线性模型的词性标注是独立预测词性，和HMM不同



稀疏特征向量用三元组形式存储。

（i,j,value）

特征向量的数学意义我了解的不够清晰，补习了一下

矩阵乘法对应了一个变换,是把任意一个向量变成另一个方向或长度都大多不同的新向量。在这个变换的过程中,原向量主要发生旋转、伸缩的变化。如果矩阵对某一个向量或某些向量只发生伸缩变换,不对这些向量产生旋转的效果,那么这些向量就称为这个矩阵的特征向量,伸缩的比例就是特征值。

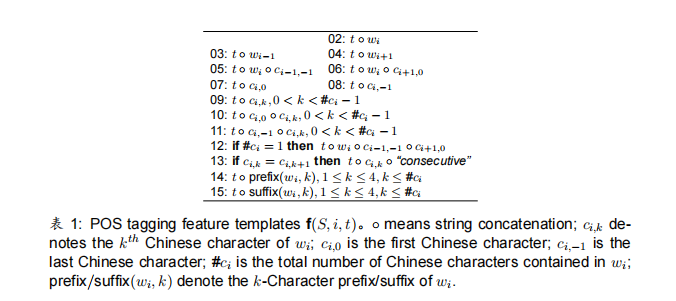
特征抽取就是减少特征数据集中的属性(或者称为特征)的数目。特征向量里特征值为0时就将这条属性减去了。

权重向量乘特征向量最终可以获得一个得分，其意义为将句子 S 中的词 wi 标为词性 t 时计算出的一个评价分数

f(.)在线性模型中的意义---？一开始没理解，后面看了特征模板后才懂，f函数得到的结果实质上是通过特征模板构造出对应各个词性的特征向量，通过将其和特征空间内的正确特征向量对比，计算正确的个数，这个个数可以代表分数，个数越多说明是这个词性的概率越大。

特征模板

通过模板处理出所有具体特征，以加入特征空间



参数理解

T >> 词性

Wi >> 词汇

Ci-1,-1 >> 前一个词的尾字

Ci+1,0 >> 后一个词的首字

t ◦ prefix(wi, k), 1 ≤ k ≤ 4, k ≤ #ci >> 前四个字

t ◦ suffix(wi, k), 1 ≤ k ≤ 4, k ≤ #ci >> 后四个字

特征空间有四维，很难实际想象出来。具体加入特征空间的为以下特征向量：

02 词性 词

03 词性 前一个词

04 词性 后一个词

05 词性 词 前一个字

06 词性 词 后一个字

07 词性 词的第一个字

08 词性 词的后一个字

For 词里除了首尾字外所有字：

09 词性 字

1. 词性 词的首字 字
2. 词性 词的尾字 字

若词长度为1（即词为单字）

1. 词性 词 前一个字 后一个字

同样还是For 词里除了首尾字外所有字：

如果和前一个字相同

13 词性 字 ‘consecutive’

对词里的第2，3，4，5个字，如果存在

14 词性 词的前2，3，4，5个字（word[：i]一一对应)

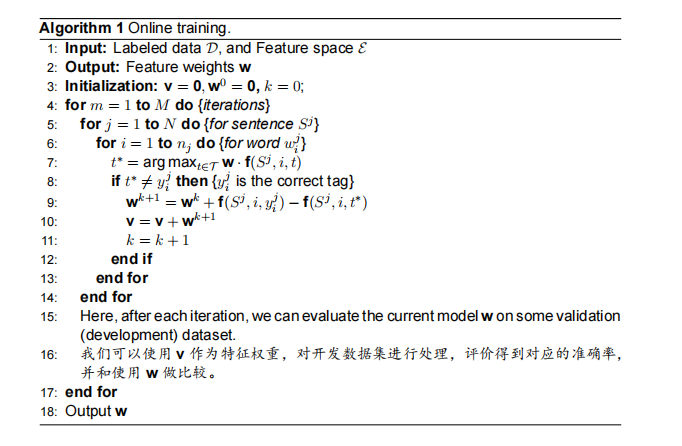
对词的倒数2，3，4，5个字，如果存在

1. 词性 词的后2，3，4，5个字（word[-i：]一一对应)

如果词本身长度小于等于4：

14 词性 词

15 词性 词



权重更新函数

正确词性的f得分减去预测词性的f得分，得到差值，差值越大，权重变更幅度就越大，差值较小时权重变化也同时变小，一次次更新，最后会回归到一个合适的权重

v为w多次变化的平均值，效果更好更稳定。