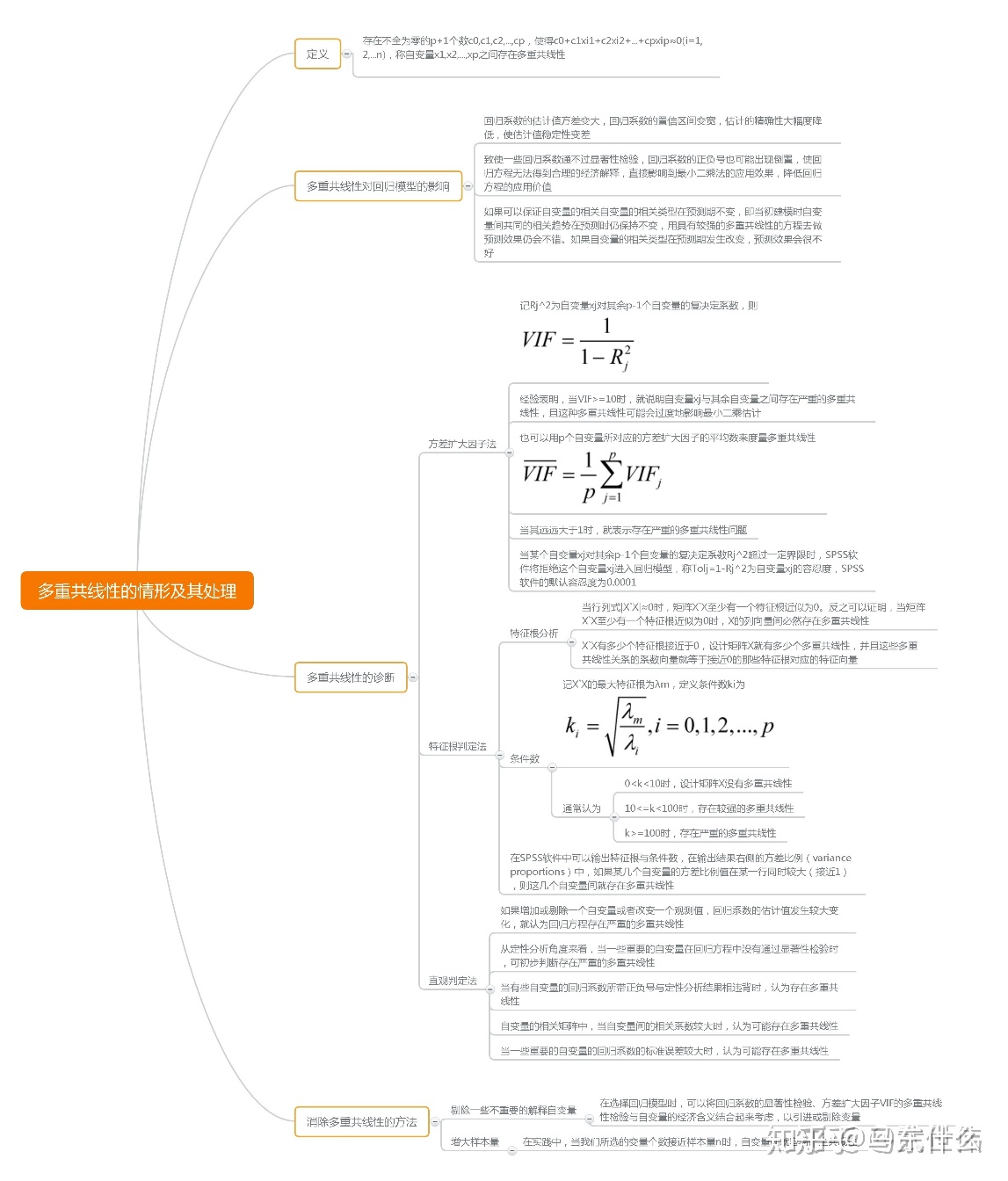
1. 为什么不直接用线性回归做分类

* 线性回归对于数据分布更为敏感，一些复杂分布的数据和异常值会对分类结果造成很大的影响，而且回归结果无法保证分布在(0,1)之间，不具有分类结果概率的性质。
* 线性回归的结果稳定性较差，sigmoid函数在x = 0周围梯度很大，可以使数据快速的从x = 0附近排开，让数据点映射到远离x = 0处去，做到将线性变换的连续结果转化为分类结果。

1. 多重共线性会给逻辑回归带来什么问题，为什么以及如何解决

* 多重共线性是指回归过程中至少有两个特征强相关或完全一致的问题，这个问题会导致逻辑回归的w数值解不稳定，参数估计的方差变大，影响特征的可解释性
* 具体原因不太清楚，直观理解是多重共线性相当于给模型添加了无意义的“假特征”，更容易造成数据过拟合。（网上都没有这么说，不过我看解决的方法很多都是解决过拟合的手段）
* PCA降维，逐步回归法，L1L2正则化，增大样本数量。
* 

1. L1L2正则化会给模型带来什么影响，为什么。

* L1是，L2是，是结构风险最小化的一种手段，目的是减少模型自由度，控制其复杂程度从而减小过拟合风险。
* （理解的比较肤浅没有公式推导）模型中系数w越多，其VC维越高，模型越复杂，在同样数量数据集上越容易造成过拟合，为了降低其模型复杂度，降低VC维，便希望可以减少w。L1L2正则化可以在原优化目标函数中增加约束条件，分别以作为惩罚项，当惩罚项与原损失函数交点的位置就是最优条件（二者相等取最小值），其中L1函数尖点（感觉不是专业说法，导数突变那个点）较多，可能会获得稀疏解，即使得某一些维度的w变成0，达到了目的；对于L2不存在尖点，但是增加了约束条件也会限制模型的自由度，可以防止很多个w在保证损失函数j不变的情况下任意取值，提高模型的稳定程度。

1. 逻辑回归本身只具有线性表达能力，如何让模型学会非线性关系

* 可以将输出的结果再次进行逻辑回归，形成多层神经元网络，对损失函数反向传播，梯度下降优化出w，得到非线性分类结果