统计学习：第二章

1 Minsky 与 Papert 指出：感知机因为是线性模型，所以不能表示复杂的函数, 如异或 (XOR)。验证感知机为什么不能表示异或。

答:异或 XOR 满足:XOR(0,1)=1,XOR(1,0)=1,XOR(1,1)=0, XOR(0,0)=0。

假设 $y = \text{sign}(ax_1 + bx_2 + c)$ 满足条件，则有 $a + c > 0, b + c > 0, c < 0, a + b + c < 0$, 由 $b + c > 0, c < 0$ 可得 $b > 0$，由 $a + c > 0, a + b + c < 0$ 可得 $b < 0$，矛盾，故知不存在可以表示异或的感知机。

2 模仿例题 2.1, 构建从训练数据集求解感知机模型的例子

3 证明以下定理：样本集线性可分的充分必要条件是正实例点集所构成的凸壳与负实例点集所构成的凸壳互不相交

答：若样本线性可分，则存在超平面 $S : wx + b = 0$，有对于正实例集合 $wx_i + b > 0$, 负实例集合 $wx'_i + b < 0$。采用反证法，若二者凸壳相交，则存在 $x = \sum_{i=1}^{j} \lambda_i x_i = \sum_{i=1}^{k} \lambda'_i x'_i$，而 $wx + b = \sum_{i=1}^{j} \lambda_i (wx_i + b) = \sum_{i=1}^{k} \lambda'_i (x'_i + b)$。第一个等号右边为正数的和，第二个等号右边为负数的和，即正又负，矛盾。因此，二者凸壳不相交。

反之，定义 $d = min(\{||x - x'|||x \in$ 正样本闭包$, x` \in$ 负样本闭包$\})$。闭包为闭集，故 d 良定且为正数，任取 x, x' 距离为 d，则令 $w = x - x'$,$b = (||x||^2 - ||x'||^2)/2$ 即可。