

Artificial Intelligence Homework 11

李云志 2015K8009929014

2017 年 11 月 18 日

1 optimal K

Answer 1 *Using dynamic programming*

首先我们定义 $OPT(i, j)$ 为数据集中前 i 个元素通过 j 次划分能够达到的最小目标函数值。

注意到，目标函数的定义为 $\sum_{j=1}^T [(A_j + \frac{\lambda}{2})\omega_j^2 - 2B_j\omega_j] + \gamma T + const$

所以子划分的目标函数之间可以通过简单加和的方式得到。

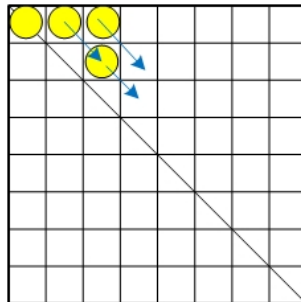
所以我们定义动态规划的递推表达式如下：

$$OPT(i, j) = \min(OPT(i - tmp, j - 1) + obj(tmp, i) + \gamma)$$

其中 $obj(tmp, i)$ 为训练数据中的子集 (tmp, i) 的目标函数值， $tmp \in [1, i - 1]$

我们可以通过事先计算得到 $OPT(i, i), OPT(i, 1), i \in [1, n]$ ，所以通过动态规划可以完成整个表的填写。

填写过程如下图：



最后我们寻找 $OPT(n, j)$ 这一列中的最小值即可，对应的纵坐标即为我们要找的最佳 k 划分的值。

可以看到，我们通过填写 $O(n^2)$ 的表，并且每一次填写经历 $O(n)$ 的循环即可，所以时间复杂度为 $O(n^3)$ ，同时我们的递归表达式保证了我们解的最优性。