应用运筹学基础：组合优化 (4) - 近似算法选讲 (2)

这一节课证明了 bin packing 问题 first fit 算法的渐进比为 1.7。

均摊体积

如果有 n 个物品，每个物品的体积都是 0.51，我们可以分析出最优目标函数值的下界至少约为 $n/2$。可是这个下界太松了（事实上最优目标函数值就是 n），只能用来证明近似比为 2。怎样才能证明渐进比为 1.7 呢?

聪明的数学家们不知怎么就想到了“均摊体积”的方法。设第 i 件物品体积为 a_i，定义权重 $w(a_i)$ 如下：$w(a_i) = \frac{6}{5}a_i + v(a_i)$ 称 v 为 bonus，定义为：$v(a_i) = \begin{cases} 0 & a_i \le \frac{1}{6} \\ \frac{3}{5}(a_i - \frac{1}{6}) & \frac{1}{6} < a_i \le \frac{1}{3} \\ \frac{1}{10} & \frac{1}{3} < a_i \le \frac{1}{2} \\ \frac{2}{5} & a_i > \frac{1}{2} \end{cases}$

证明思路

记 $w(I)$ 为 bin packing 的一个实例 I 的权重总和，$\mathrm{FF}(I)$ 表示对实例 I 运用 first fit 算法得到的目标函数值，$\mathrm{OPT}(I)$ 表示实例 I 的最优目标函数值。

再记 B 为 first fit 算法得到的方案，B^* 为最优方案，$c(B_j)$ 表示第 j 个 bin 中物品的体积总和，$w(B_j)$ 表示第 j 个 bin 中物品的权重总和。

我们容易得到以下等式 $w(I) = \sum_{i=1}^{n} w(a_i) = \sum_{j=1}^{\mathrm{FF}(I)} w(B_j) = \sum_{j=1}^{\mathrm{OPT}(I)} w(B_j^*)$

如果我们能证明 $\forall c(B_j^*) \le 1, w(B_j^*) \le 1.7$，根据 $w(I) = \sum_{j=1}^{\mathrm{OPT}(I)} w(B_j^*)$，我们就能得到 $w(I) \le 1.7\mathrm{OPT}(I)$；如果我们还能证明所有 $w(B_j)$ 的均值都至少为 1，那么根据 $w(I) = \sum_{j=1}^{\mathrm{FF}(I)} w(B_j)$，我们就能证明 $\mathrm{FF}(I) \le w(I) \le 1.7\mathrm{OPT}(I)$。不过我们这里要证明一个弱一点的结论：除了两个 bin 以外，其它 bin $w(B_j)$ 的均值都至少为 1。后面我们会看到，这个结论将会推导出 $\mathrm{FF}(I) \le w(I) + 0.8 \le 1.7\mathrm{OPT}(I) + 0.8$，就能证明 first fit 算法 1.7 的近似比。

第一步：证明均摊体积不超过 1.7

第一步的证明比较容易，根据权重的定义可以直接推导出来。对于一个 bin，分以下情况讨论。

1. 如果所有物品体积 c 均有 $c \le \frac{1}{6}$

这个情况下，bin 的权重就是 bin 中物品体积总和的 1.2 倍，不会超过 1.7。

2. 如果存在物品体积 c 有 $\frac{1}{6} < c \le \frac{1}{2}$

很显然，这种物品在一个 bin 内至多有 5 个，那么 bonus 不会超过 $\frac{1}{10} \times 5 = \frac{1}{2}$，权重也不会超过 1.7。

3. 如果存在两个物品体积 c_1 和 c_2 有 $c_1 > \frac{1}{2}$ 且 $\frac{1}{3} < c_2 \le \frac{1}{2}$

很显然，其它物品的体积都不会超过 $\frac{1}{6}$，没有 bonus；c_1 和 c_2 带来的 bonus 恰为 0.5，权重不会超过 1.7。

4. 如果存在三个物品体积 c_1，c_2 和 c_3 有 $c_1 > \frac{1}{2}$，$\frac{1}{6} < c_2, c_3 \le \frac{1}{3}$ 且 $c_2 + c_3 < \frac{1}{2}$

很显然，其它物品的体积都不会超过 $\frac{1}{6}$，没有 bonus；c_2 和 c_3 带来的 bonus 为 $\frac{3}{5}(c_2 - \frac{1}{6}) + \frac{3}{5}(c_3 - \frac{1}{6}) < 0.1$，再加上 c_1 带来的 bonus 0.4，权重不会超过 1.7。

第二步：证明除两个 bin 以外，其它 bin 权值均值至少为 1

我们首先去掉权值至少为 1 的 bin，考虑那些权值不足 1 的 bin。容易证明，权值不足 1 的 bin 有以下性质：

\1. 不含体积至少为 0.5 的物品；

\2. 一个 bin 内不会包含两个体积至少为 1/3 的物品；

\3. bin 的体积之和小于 5/6。

据此容易推出：

\1. 除了最后一个 bin，其它 bin 中至少有两个物品；

\2. 除了最后两个 bin，其它 bin 的体积之和都大于 2/3（如果有一个 bin 的体积之和不超过 2/3，由于是 first fit 算法，后面 bin 里的物品体积肯定至少为 1/3 但不足 1/2；而后面至少还有两个 bin，这就违反了"一个 bin 内不会包含两个体积至少为 1/3 的物品"的性质）。

下面证明一个引理：如果两个 bin B_1 和 B_2 满足 B_1 在 B_2 前面、$w(B_1), w(B_2) < 1$、$c(B_1) \geq \frac{2}{3}$ 以及 B_2 有至少两个物品，则 $\frac{6}{5}c(B_1) + v(B_2) \geq 1$。

引理的证明，只需要分类讨论 B_2 里体积最小的物品的体积 c' 即可：

首先，$c' \geq \frac{1}{6}$，不然 B_1 将会与"bin 的体积之和小于 5/6"的性质矛盾；

其次，$c' < \frac{1}{3}$，不然 B_2 将会与"一个 bin 内不会包含两个体积至少为 1/3 的物品"的性质矛盾；

另外，$c' > 1 - c(B_1)$，不然 c' 就会放进 B_1 里。

那么只可能有 $\frac{1}{6} \leq c' < \frac{1}{3}$，则
$$\begin{array}{cc} & \frac{6}{5}c(B_1) + v(B_2) \\ \geq & \frac{6}{5}c(B_1) + 2 \times v(c') \\ > & \frac{6}{5}c(B_1) + \frac{6}{5}(1 - c(B_1) - \frac{1}{6}) \\ = & 1 \end{array}$$

假设 first fit 得到的方案中，权重之和小于 1 的 bin 按先后顺序为 $B_1, B_2, \ldots, B_k$，那么
$$\begin{array}{cc} & w(B_1) + w(B_2) + \cdots + w(B_{k-2}) + w(B_{k-1}) + w(B_k) \\ = & v(B_1) + (\frac{6}{5}c(B_1) + v(B_2)) + \cdots + (\frac{6}{5}c(B_{k-2}) + v(B_{k-1})) + (\frac{6}{5}c(B_{k-1}) + \frac{6}{5}c(B_k)) + v(B_k) \\ \geq & (k-2) + \frac{6}{5} \end{array}$$
也就是说，除了最后两个 bin，其它的 bin 权值均值都至少为 1。再补个 0.8，再加上权值本来就至少为 1 的 bin，那么所有的 bin 权值均值就都至少为 1 了。这就完成了渐进比为 1.7 的证明。