Home Page
Title Page
Page 1 of 75
Go Back
Full Screen
Close
Quit

运筹通论II

刘克
中科院数学与系统科学研究院 北京100190

邮箱地址: kliu@amss.ac.cn

第三部分　马氏决策—空集装箱调配问题

Home Page
Title Page
Page 3 of 75
Go Back
Full Screen
Close
Quit

1 单港口的问题与建模

2 无限阶段折扣准则

3 无限阶段平均准则

4 数值例子

5 多港口空集装箱的调配问题

1 港口的问题与建模

集装箱配送业务是一个竞争的服务行业.为了增加公司的竞争能力,一方面是降低自身的成本,一方面要尽可能的满足顾客的需求.如果公司保有一个船队专门负责调配空集装箱,费用是十分昂贵的而且是不现实的. 做好公司自有的和外租的集装箱存储计划则是有效可行的方案.由于贸易的不平衡性,在配送线路上有的港口会积累大量的空集装箱,而此时在另一些港口上面临着空集装箱的严重短缺.特别的是, 在面对空集装箱的不确定性的需求下,如何合理的调配空集装箱是香港许多集装箱配送公司面临的共同问题.

香港南华早报1997年6月27日报道:2005年香港将需要3千1百万个TEU(twenty-foot equivalent units)单位, 到2011年需要3千6百万个TEU单位,而到2016年会增加到4千万个TEU单位.因此,为公司提供较好的方案就会为其带来巨大的经济利益.

为了解决这个问题,我们利用MDP的方法,为公司提供最好的方案.首先我们针对一个港口建立MDP的模型, 寻求最优的(费用最低的)平稳策略.然后再将其推广到多个港口的问题中去.

首先我们考虑一个公司在单港口的运作规律.例如香港一家公司在远东的很多港口都有分公司,他们每天要向香港总部提供更新的订货信息,如需要的空集装箱数目和需要运走的装满的集装箱数目等等.他们还要提供随时更新的集装箱的状态信息,例如:空的,装满的,在途的,托运的或者代理的等等.公司将这些信息存入计算机进行处理.当一艘船离开迪拜(离开中东开往远东的最后一站)进入远东范围时,开始收集托运的集装箱并且同时收集每个港口的集装箱状态信息以为将来提供补充的空集装箱做准备,这包括确认一些潜在的缺货问题.

由于供应空集装箱需要10到15天的延迟时间,所以托运船队需要预测10到15天的需求.但是通常来说,这种预测是十分困难的,经常会受到很多因素的影响,使得变化非常的大.但是这些分公司可以相互调配空集装箱.如果急需的话, 也可以向外部公司临时租赁空集装箱,但是这会造成较高的租赁费用.

考虑到在一个港口卸下来的满货集装箱一个周期后可以将货物腾空,实际上这种集装箱可以认为有一个周期的延迟使用时间.但是我们认为每个周期到货的满集装箱数也是不确定的,这是由于很多原因造成船期的不确定性的影响.我们要针对这个问题建立MDP模型并求解出具有特殊结构的最优策略.

一般的运输问题比较通常的物流分配问题会更复杂一些,文献[37]总结了1960年以来的许多静态优化问题. 而文献[133]则利用求解动态资源分配的方法求解运输问题,实际上可以看成为一类特殊的供需库存问题. 类似的文献还有[132],[69],[134]和[33]等等. 还有一些文献是关于库存理论的,尽管不是针对集装箱调配的,但是思想可以借鉴过来,当然需要就我们所面对的问题做出必要的变化.例如考虑租赁设备的文献有[165];考虑货币平衡问题的有文献[70],[54]和[125]等等.

我们这里的模型主要有延迟时间,复杂的费用结构以及运来空集装箱和运走空集装箱都是可行的决策规则等特点, 造成前面提到的模型都不适用.因此,我们需要最简单的模型开始分析到实际上比较适用的模型上去. 下面我们介绍一些需要用到的记号.

n	离散时间的阶段数(周期数);
x	每周期开始时港口拥有的可使用的空集装箱数;
d	决策变量(正值表示进口空集装箱,负值表示出口);
y	在阶段$n+$时刻的可用空箱数;
Z_1	每周期进口的满箱数,为独立同分布的随机变量;
P_1	Z_1取值的概率,其中$P_1(Z_1=z_1)=p_1(z_1), 0\le z_1\le K$;
F_1	Z_1的分布, 其中 $F_1(Z_1\le z)=\sum_{z_1=0}^{z}p_1(z_1)$;
Z_2	每周期出口的满箱数,为独立同分布的随机变量;
P_2	Z_2取值的概率,其中$P_2(Z_2=z_2)=p_2(z_2), 0\le z_2\le K$;
F_2	Z_2的分布, 其中 $F_2(Z_2\le z)=\sum_{z_2=0}^{z}p_2(z_2)$;
c_h	一周期一个空箱的期望存储费用;
c_u	进口一个空箱的期望费用;
c_t	出口一个空箱的期望费用;
c_p	租用一个空箱一周期的期望费用;
K	一周期最大吞吐量限制;
β	折扣因子$0<\beta<1$

其中:d取正值表示进口空集装箱数,取负值表示出口空集装箱数;y包括原有的空箱,上周期进口的满箱以及本周期进口的空箱. 我们这里不考虑固定费用.

对给定的时间周期$0 \le n \le N$,状态就是当前的空集装箱库存数,记为x,可用的行动为d.那么下一个阶段开始时的状态就是

$$x' = x + d - Z_2 + Z_1 = y - Z_2 + Z_1,$$

它说明在周期$0 \le n \le N$时到港的满箱Z_1只有在周期$n+1$时才可以使用.

那么,对于状态x,第n周期的最小折扣费用就是

$$
\begin{aligned}
V_\beta^n(x) \;=\; & \min_y\{c_u(y-x)^+ + c_t(x-y)^+ + L(y) \\
& + \beta\sum_{z_1=0}^{K}\sum_{z_2=0}^{K} V_\beta^{n+1}(y+z_1-z_2)p_1(z_1)p_2(z_2)\},
\end{aligned}
$$

其中$V_\beta^N(x)=0$,对于一切x,而$x^+=\max\{x,0\}$.参数$x'=y+z_1-z_2$表示被租用的箱子在船队到达时就归还了.

$L(y)$表示期望存储费用与租赁费用之和,具体定义为

$$L(y) = c_h E(y - Z_2)^+ + c_p E(Z_2 - y)^+. \tag{1}$$

很明显$(y-c)^+$和$(c-y)^+$对于常数c来说,关于y是凸函数.取数学期望,压缩和求和都保持函数的凸性, 所以$L(y)$关于y是凸函数.做$L(y)$的差分得到:

$$\begin{aligned} \Delta L(y) &= L(y+1) - L(y) \\ &= \begin{cases} -c_p, & y < 0, \\ (c_p + c_h) F_2(y) - c_p, & 0 \le y \le K, \\ c_h, & y > K. \end{cases} \end{aligned} \tag{2}$$

对固定的$0 \le n \le N$和任意的状态y,定义函数$G_n(y)$为

$$G_n(y) = L(y) + \beta \sum_{z_1=0}^{K} \sum_{z_2=0}^{K} V_\beta^{n+1}(y + z_1 - z_2) p_1(z_1) p_2(z_2),$$

及$G_n(y)$的差分

$$\begin{aligned} \Delta G_n(y) &= G_n(y+1) - G_n(y) \\ &= \Delta L(y) + \beta \sum_{z_1=0}^{K} \sum_{z_2=0}^{K} \Delta V_\beta^{n+1}(y + z_1 - z_2) p_1(z_1) p_2(z_2). \qquad (3) \end{aligned}$$

在有限阶段问题中,如果在第n周期决策者已经决定将空集装箱数调整到y(n时刻的状态)并从这以后决策者总使用最优的决策,这样可以利用定理2.3知道$G_n(y)$表示了最小期望折扣费用(注意到这里与定理2.3有点区别, 那就是这里含有折扣因子而定理2.3的叙述中折扣因子是1.但是这个区别不是本质的,因为一般的折扣有限阶段问题可以完全的被平行推出).读者也可以参见文献[185]或者文献[173]. 所以差分$\Delta G_n(y)$表述了在第n周期开始时如果决策者将状态从y提升到$y+1$时的期望折扣费用的差异.

性质7.1：对所有的$0 \le n \le N, V_\beta^n(x)$关于$x$是凸函数.

证明思路:用归纳法可以直接证明,或者参见文献[107].

根据性质7.1,在周期n,我们可以按照下面的方式调整行动:如果$\Delta G_n(y) > c_t$, 决策者保持空集装箱水平y比$y+1$要好;如果$\Delta G_n(y) < -c_u$,则相反.这样就可以提供一个有效的方法确定最优的行动了.我们令

$$D_n(\beta) = \min\{x : \Delta G_n(x) > c_t\}, \tag{4}$$

$$U_n(\beta) = \min\{x : \Delta G_n(x) > -c_u\}. \tag{5}$$

那么在周期n的最优决策规则具有特定的形式,称为$(U_n(\beta), D_n(\beta))$决策规则.

特别地，称$U_n(\beta)$为第n阶段的上限控制量,而$D_n(\beta)$被称为下限控制量.也就是说,如果在周期n开始时系统的状态为x(在n阶段选取行动之前的状态),最优的行动$d(x)$应该满足

$$d(x) = \begin{cases} U_n(\beta) - x & x < U_n(\beta) & \text{增加空集装箱到 } U_n(\beta), \\ 0 & U_n(\beta) \le x \le D_n(\beta) & \text{不做调整}, \\ x - D_n(\beta) & D_n(\beta) < x & \text{减少空集装箱到 } D_n(\beta). \end{cases} \tag{6}$$

仅仅根据$\Delta G_n(y)$的单调性不能保证公式(4)以及(5)定义的两个量是有限的. 如果$U_n(\beta) = -\infty$或者$D_n(\beta) = \infty$最优行动有可能总是保持不调整. 为了方便,我们不排除公式(6)中表示了这种退化的情形.

性质7.2：对所有的$0 \leq n \leq N$,阶段n的最优决策规则具有形式$(U_n(\beta), D_n(\beta))$.

证明思路:用归纳法可以直接证明,或者参见文献[107].

从性质7.1和性质7.2知道

$$V_\beta^n(x)=\begin{cases} c_u(U_n(\beta)-x)+G_n(U_n(\beta)), & x<U_n(\beta),\\ G_n(x), & U_n(\beta)\le x\le D_n(\beta),\\ c_t(x-D_n(\beta))+G_n(D_n(\beta)), & x>D_n(\beta).\end{cases}$$

$$\Delta V_\beta^n(x)=\begin{cases} -c_u, & x<U_n(\beta),\\ \Delta G_n(x), & U_n(\beta)\le x<D_n(\beta),\\ c_t, & x\ge D_n(\beta).\end{cases} \tag{7}$$

令$y^*=\min\{y:\Delta L(y)>0\}$,容易知道必有:$0<y^*<K$.

性质7.3：对所有的$0 \leq n \leq N$,有$U_n(\beta) \leq y^*$和$D_n(\beta) \geq y^*$.

证明思路:将y^*代入公式(4),可以直接证明,或者参见文献[107].

下面我们分析一下c_u, c_p, c_t,和c_h之间的关系.如果$c_u > c_p$,说明了租用一个空箱子要比运来一个费用低,很明显这时候没有必要补充空集装箱.同样,如果$c_t > c_h$,说明了存储一个空箱子要比运走一个费用低, 所以没有必要运走多余的空集装箱.

性质7.4: 如果$c_u < \frac{c_p}{1-\beta}$和$c_t < \frac{c_h}{1-\beta}$成立.则存在一个有限的$N^*$满足:当$N \geq N^*$时, 对于一切的$0 \leq n \leq N - N^*$有$D_n(\beta) < \infty$和$U_n(\beta) > -\infty$.

证明思路:证明参见文献[107].

性质7.4表明了使得$D_n(\beta)$和$U_n(\beta)$有限的条件.由于参数c_u, c_p, c_t,和c_h都是正的实数,所以当β充分的接近于1的时候,总有$c_u < \frac{c_p}{1-\beta}$和$c_t < \frac{c_h}{1-\beta}$成立.所以我们假设下面的条件成立并不是一个实质的限制.我们总认为

$$c_u < \frac{c_p}{1-\beta} \quad \text{和} \quad c_t < \frac{c_h}{1-\beta} \tag{8}$$

在本章成立.

性质7.5：如果n阶段的最优限制决策规则是有限数对$(U_n(\beta), D_n(\beta))$, 那么$n+1$阶段的最优决策规则$(U_{n+1}(\beta), D_{n+1}(\beta))$也是有限的, 而且存在常数$K$,使得

$$D_{n+1}(\beta) \leq D_n(\beta) + K, \qquad U_{n+1}(\beta) \geq U_n(\beta) - K.$$

.

证明思路:证明参见文献[107].

有了性质7.3,性质7.4和性质7.5,我们知道如果条件(8)式成立,则当n足够大的时候,控制限$(U_n(\beta), D_n(\beta))$都是有限的.

下面我们将有限阶段模型的结果推广到无限阶段的折扣情形.然后利用折扣因子β趋于1的技术,得到长期平均意义下的最优平稳策略和最优的平均准则值函数,以解决就单港口的优化问题.

2 无限阶段折扣准则

我们主要考虑当n趋于无穷时,准则值函数$V_{\beta}^{n}(x)$,下限控制量$U_n(\beta)$和上限控制量$D_n(\beta)$的极限情况. 根据前一节的性质和这个问题的设定条件,已经知道对无限阶段折扣问题总存在有限的非退化解.但是我们还需要建立以下的两个结论: (i) 对于每个有限的状态,存在一个策略使得在这个策略下无限阶段折扣问题的期望折扣费用是有限的; (ii) 无限阶段折扣问题的两个控制限$U(\beta)$和$D(\beta)$都是有限数.

首先,我们说明无限阶段的最优折扣费用是有限的.定义

$$V_\beta(x)=\lim_{n\to\infty}V_\beta^n(x).$$

则,我们有下面的性质.

性质7.6：对一切的有限状态x,当n趋于无穷时,$V_\beta^n(x)$的极限是存在的.

证明思路:性质的证明可以直接从下面的引理7.1和引理7.2得到.

引理7.1：对一切的有限状态x,集合$\{V_\beta^n(x)\}$有上界.

证明思路:对每个固定的x,我们采用每个阶段都使空集装箱的可用水平保持在x这个策略,它的值函数应该比最优的值函数不会好,也就是说它是与n无关的.因此,这个策略的值函数就是$V_\beta^n(x)$的一个上界.

引理7.2: . 对任意的$n \geq 0$和固定的状态x,值函数$V_\beta^n(x)$是单调非降的,也就是说满足: $V_\beta^{n+1}(x) \geq V_\beta^n(x)$.

证明思路:利用归纳法直接证明.

下面考虑控制限$U_n(\beta)$和$D_n(\beta)$的收敛问题.记

$$
\begin{aligned}
G(y) &= L(y) + \beta \sum_{z_1=0}^{K} \sum_{z_2=0}^{K} V_\beta(y + z_1 - z_2) p_1(z_1) p_2(z_2), \\
U(\beta) &= \min\{z : G(z) > -c_u\}, \\
D(\beta) &= \min\{z : G(z) > c_t\}.
\end{aligned}
$$

再定义函数g具有形式:

$$g(x) = \min_y \{c_u(y-x)^+ + c_t(x-y)^+ + G(y)\} \tag{9}$$

因为$V_\beta(\cdot) = \lim_{n\to\infty} V_\beta^n(\cdot)$是凸函数, 因此,公式(9)确定的最优解也具有形式$(U(\beta),\ D(\beta))$.我们只需要说明$U(\beta)$和$D(\beta)$是有限的, 以及$V_\beta(\cdot) = g(\cdot)$就可以了.

引理7.3：. 如果$0 < y < D_n(\beta)$,那么

$$\Delta G_n(y) \geq \sum_{i=0}^{A-1} \beta^i c_h - \beta^A c_p - \beta^{A+1} c_u,$$

其中A是小于或者等于y/K的最大整数.

证明思路:将A分为小于1,等于1和大于1分别考虑.详细证明参见文献[107].

对固定的β,定义

$$\hat{N}(\beta) = \min\{N : \sum_{i=0}^{N-1} \beta^i c_h - \beta^N c_p - \beta^{N+1} c_u > c_t\}. \tag{10}$$

由于$\frac{c_h}{1-\beta} > c_t$,所以我们知道$\hat{N}(\beta) < \infty$.取固定的$\beta_2 \in (\beta, 1)$,存在$N_1 < \infty$满足

$$\sum_{i=0}^{N_1-1} \beta_2^i c_h - c_p - c_u \geq c_t. \tag{11}$$

我们总有下面的引理.

引理7.4：. 对于所有的$\beta < \beta_2 \leq \beta_1 < 1$,我们有

$$\hat{N}(\beta_1) \leq N_1 < \infty, \tag{12}$$

其中$\hat{N}$通过(10)定义,而N_1是由(11)所定义.

证明思路:证明参见文献[107].

性质7.7：. 对一切的$n > \hat{N}$,我们有$D_n(\beta) \leq (\hat{N}+2)K$.

证明思路:利用反证法,或者参见文献[107].

利用完全类似(证明$D_n(\beta)$一致有界)的方法,我们可以得到对于$U_n(\beta)$一致有界的结论.这里我们就不再重复了.

最后总结为如下的性质.

性质7.8: . 当$n \to \infty$时,$U_n(\beta)$的极限$U(\beta)$和$D_n(\beta)$的极限$D(\beta)$都是有限的,而且存在一个有限的数N,使得$-NK < U(\beta) \leq D(\beta) < NK$.

下面还需要讨论$V_\beta(\cdot)$和$g(\cdot)$之间的关系.

性质7.9：. $V_\beta(\cdot)$满足方程

$$V_\beta(x) = \min_y\{c_u(y-x)^+ + c_t(x-y)^+ + G(y)\} = g(x).$$

证明思路:利用标准的分析方法,证明两者之差小于任意给定的$\epsilon > 0$.详细证明参见文献[107].

结论:对任意接近于1的折扣因子β我们已经证明了两个控制限的这种策略对$(U(\beta), D(\beta))$在最为一般的策略类中也是最优策略.

3 无限阶段平均准则

由于最终我们需要考虑长期运行下的平均费用最低,回忆第4章对于平均准则的定义, 对任意的策略$\pi \in \Pi$,

$$\bar{V}(x,\pi) = \limsup_{N\to\infty} \frac{V_N(x,\pi)}{N+1} \tag{13}$$

这里$V_N(x,\pi)$是用策略π且折扣因子β为1的N阶段的总期望费用,x则表示初始的出发状态. 这里还需要解释的就是,我们现在考虑的是寻求费用最小的最优策略,所以定义公式(13)与定义式(4.1) 有所不同,也就是在公式(13)中我们考虑的是费用的上极限准则.

对于费用的上极限准则,第4章中的所有结果都是成立的.需要改动的只是寻求极大的地方都改为寻求极小.根据定理4.1, 文献[63]中定理2.8.3或者文献[32]中更一般的结论,当状态空间和行动集合都是有限集合时,

$$\bar{V}(x) = \lim_{\beta \to 1} (1-\beta) V_\beta(x). \tag{14}$$

不仅如此,对于充分接近于1的β来说,β折扣最优策略也是平均准则下的最优策略.

在利用从折扣模型得到平均模型的这个过程中,由于c_u, c_p, c_t和c_h都是大于0的数,所以当β接近于1的时候, 条件(8)自然满足,最优策略具有形式$(U(\beta), D(\beta))$.当然了,β折扣最优策略的结构依然保持. 也就是说,平均准则最优策略也具有β折扣最优策略的结构性质,即具有(U, D)形式,而且这里的控制限U和D对于充分接近于1的折扣因子β一致的成立.

不幸的是,我们这里建立的模型不能够直接利用前面的结论,其原因就是状态空间和可用的行动集合都不是有限的集合.但是,在性质7.8的保证下以及向外租赁的空集装箱没有量的限制的条件下,我们可以将状态空间分成两个部分.

具体来说,对于固定的折扣因子β,令

$$S(\beta) = [U(\beta) - K, D(\beta) + K],$$

那么一部分就是$S(\beta)$; 另一部分就是$S/S(\beta)$, 其中$(U(\beta), D(\beta))$是对应于β折扣模型的最优策略.利用条件(8), 引理7.4,性质7.7,以及关于$D(\beta)$的相应性质,结合性质7.8,我们知道$U(\beta)$和$D(\beta)$关于β接近于1时是一致有界的.

如果不考虑瞬时状态而仅限于考虑状态空间$S(\beta_0)$ 并限制可用的行动集合保证$S(\beta_0)$是一个闭集合(可用的行动集合使状态空间$S(\beta_0)$是闭的表示:在每个状态$x \in S(\beta_0)$上的任意可行的行动实施以后,系统的状态变化不会转移到状态空间$S(\beta_0)$以外的状态上去), 其中

$$S(\beta_0) = [U(\beta_0) - K, D(\beta_0) + K],$$

而$U(\beta_0)$和$D(\beta_0)$分别表示任取一串趋于1的折扣因子β_n, 对应的最优控制上下限的极限点.这些结论的正确性由前面的分析保证了.

限制在状态空间$S(\beta_0)$上的问题就可以转化为一个有限状态空间和有限行动集合的MDP问题.这样我们通过有限阶段折扣模型,经过无限阶段折扣模型最后到无限阶段平均模型的分析就完成了,得到的结论就是存在最优的$(U(\beta_0), D(\beta_0))$策略(也记为(U, D)策略). 一般的来说,平均准则的两个控制限有可能退化为一个点,即$U = D$.

为了具体的求解这个问题,我们可以利用状态空间截尾(类似于有限状态逼近)的方法,并且限定每个状态上的可用行动是使得截尾后MDP问题中的状态空间是闭的.利用上面的方法和公式(13) 可以得到平均准则下的两点控制限最优策略(U, D).

实际上,在考虑长期平均准则的意义下,忽略一些瞬时状态是没有任何影响的,所以状态空间截尾的方法是有效的.但是还有一个问题,就是确定状态空间截尾的长度.基于长期折扣模型的最优策略也是两点控制限的策略,以及当折扣因子充分接近于1时折扣最优策略就是平均准则下的最优策略等因素,我们可以利用折扣模型的两点控制策略来估计平均准则最优策略的两点控制限的界.最后,可以求得平均准则最优策略的两点控制限.

4 数值例子

我们在这一节里给出一个数值例子,说明如何确定这两个最优控制限U和D.这是香港一家公司的实际数据.具体的参数为:

$$M = 2000 \text{ 单位},\ K = 100 \text{ 单位},\ c_u = 1200,\ c_t = 800,\ c_p = 260,\ c_h = 150,$$

其中M是截尾状态空间的界,而费用c_u, c_t, c_p和c_h都是按照每个集装箱计算的.到港和出港的集装箱是服从同一个正态分布的随机变量,这个分布函数的参数分别为:均值$\mu = 70$和标准差$\sigma = 10$.对于$\beta = 0.9999$和$\beta = 0.999$分别计算了它们无限阶段的折扣最优准则值(费用)函数$V^\beta(\cdot)$.

针对这个具体的问题,我们利用下面的算法完成最优费用函数的计算.

集装箱问题截尾状态空间的算法

步骤1: 令$i = 1$且对一切$x \in [-M, M]$,$V_\beta^0(x) = 0$以及$\epsilon = 0.001$

$$L(y) = \begin{cases} c_pEZ_2 - c_py, & y < 0, \\ (c_p + c_h)\sum_{z_2=0}^{y-1} F_2(z_2) - c_py + c_pEZ_2, & 0 \le y \le K, \\ c_hy - c_hEZ_2, & y > K. \end{cases}$$

和

$$G_i(y) = L(y) + \beta \sum_{z_1=0}^{K} \sum_{z_2=0}^{K} V_\beta^{i-1}(y + z_1 - z_2)p_1(z_1)p_2(z_2).$$

步骤2：计算

$$\begin{aligned} U_i(\beta) &= \min\{z : G_i(z) > -c_u\}, \\ D_i(\beta) &= \min\{z : G_i(z) > c_t\}. \end{aligned}$$

根据下面的规则求$V_\beta^i(y)$的值:
1)如果$U_i(\beta) = -\infty$，$D_i(\beta) = \infty$,则令$V_\beta^i(y) = G_i(y)$;
2)如果$U_i(\beta) = -\infty$，$D_i(\beta) < \infty$,则令

$$V_\beta^i(y) = \begin{cases} G_i(y), & y \le D_i(\beta), \\ c_t(y - D_i(\beta)) + G_i(D_i(\beta)), & y > D_i(\beta); \end{cases}$$

(步骤2续)

3)如果$U_i(\beta) > -\infty,\ D_i(\beta) < \infty$,则令

$$V_\beta^i(y) = \begin{cases} c_t(U_i(\beta) - y) + G_i(U_i(\beta)), & y < U_i(\beta). \\ G_i(y), & U_i(\beta) \le y \le D_i(\beta), \\ c_t(y - D_i(\beta)) + G_i(D_i(\beta)), & y > D_i(\beta); \end{cases}$$

4)如果$U_i(\beta) > -\infty,\ D_i(\beta) = \infty$,则令

$$V_\beta^i(y) = \begin{cases} c_t(U_i(\beta) - y) + G_i(U_i(\beta)), & y < U_i(\beta). \\ G_i(y), & y \ge U_i(\beta); \end{cases}$$

其中$y \in [-M, M]$.

步骤3： 当

$$\max_{x\in[-M,M]} |V_i(x) - V_{i-1}(x)| > \epsilon$$

时令$i+1 \Rightarrow i$后返回到步骤2,否则就停止.

由此,我们得到$U(0.999) = U(0.9999) = 54$, $y^* = 73$, $D(0.999) = D(0.9999) = 104$.很明显,对于两个折扣因子都有$U(\beta) < y^* < D(\beta)$,和相应的值$V_{0.999}(\cdot)$和$V_{0.9999}(\cdot)$. 特别的是对于折扣因子0.9999,最小的函数值$V_{0.9999}(79) = 62877464.41$.

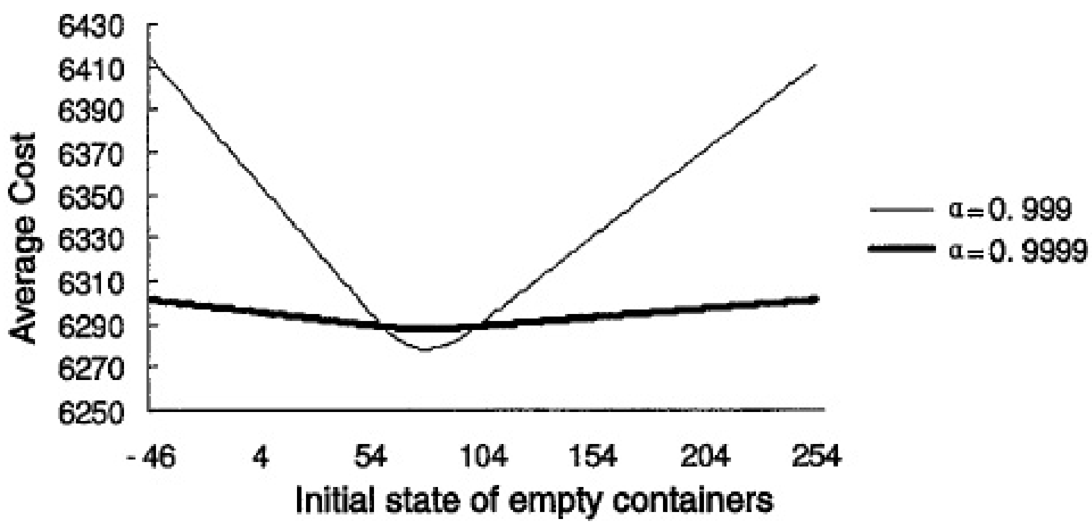

Figure 1. The graph of the function $(1-\alpha)V^{\alpha}(\cdot)$.

我们现在考虑平均准则模型.首先,我们从折扣模型的最优控制限$(U, D) = (54, 104)$扩大到范围$[U' - K, D' + K]$, 其中(U', D')为$(0, 150)$.此时的截尾状态空间为$\bar{S} = \{s : U' - K \le s \le D' + K\}$. 对每个限制在$\bar{S}$上的状态$s$,可用的行动集合为$A(s) = \{a : U' \le a \le D'\}$,它的含义是:如果$s < U'$(或者$D' < s$),可用的行动是增加(减少)$s$不超出范围$[U', D']$. 因此我们已经定义好一个有限状态的平均准则MDP问题,利用标准的平均准则值迭代算法,我们最后得到两点最优的控制限策略为$(\bar{U}, \bar{D}) = (54, 104)$,最优值函数$\bar{V}(\cdot)$是6288.7955.

值得一提的是,0.9999折扣值函数的最小值在点$x = 79$出现,且是62877464.41.如果利用$(1-\beta)V_\beta(79)$来估计的话,$0.0001 \times 62877464.41 \approx 6288.795508$, 是非常接近于平均准则的最优值的.

5 多港口空集装箱的调配问题

类似于单港口的情形,我们先介绍一下所使用的符号.

n 离散时间的阶段数(周期数);

N 所考虑的港口总数;

M 公司拥有的集装箱总数;

K^i 港口i一周期的最大吞吐量;

Z_1^i 港口i每周期进口的满箱数,为独立同分布的随机变量;

P_1^i Z_1^i取值的概率,其中$P_1^i(Z_1^i = z) = p_1^i(z), 0 \le z \le K^i$;

F_1^i Z_1^i的分布, 其中 $F_1^i(Z_1^i \le z) = \sum_{z_1=0}^{z} p_1^i(z_1)$;

Z_2^i 港口i每周期出口的满箱数,为独立同分布的随机变量;

P_2^i Z_2^i取值的概率,其中$P_2^i(Z_2^i = z) = p_2^i(z), 0 \le z \le K^i$;

F_2^i Z_2^i的分布, 其中 $F_2^i(Z_2^i \le z) = \sum_{z_2=0}^{z} p_2^i(z_2)$;

c_h^i 港口i一周期一个空箱的期望存储费用;

c_u^i 港口i进口一个空箱的期望费用;

c_t^i 港口i出口一个空箱的期望费用;

c_p^i 港口i租用一个空箱一周期的期望费用.

根据上面的记号,我们可以定义一个MDP.如果用x_i表示在港口i每个周期开始的空集装箱数,$i = 1, 2, \cdots, N$. 定义系统的状态为

$$x = (x_1, x_2, \cdots, x_N),$$

它满足条件$\sum_{i=1}^{N} x_i = M$.令

$$S = \{x | \sum_{i=1}^{N} x_i = M\}$$

为所有可能的状态全体.

对每个状态$x \in S$,我们令$a = (a_1, a_2, \cdots, a_N)$为一个可用的行动,其中$a_i$表示在港口$i$此时的决策变量(行动),当其为正时表示运来数量为$a_i$的空集装箱,为负值时则表示运走数量为$a_i$的空集装箱.我们这里的模型没有考虑运输所需要的时间,只考虑费用对空集装箱调配的影响.因此,我们有

$$\sum_{i=1}^{N} a_i = 0.$$

为了方便,我们用$A(x)$表示在状态$x \in S$时的可用行动全体的集合.

如果当前的状态是$x \in S$,所采用的行动是$a \in A(x)$,那么当前实际上可以使用的空集装箱数目就是向量$y = (y_1, y_2, \cdots, y_N)$,其中$y_i = x_i + a_i$表示在港口$i$的可用数目.也可以这样说,$y$表示在周期+时刻每个港口的可用空箱数目.因此,在下一个决策周期开始时状态为$x' = (x'_1, x'_2, \cdots, x'_N)$的概率就是

$$p(x'|x, a) = p(x'|y) = \prod_{i=1}^{N} p(x'_i|y), \tag{15}$$

其中$x'_i = y_i - z_2^i + z_1^i$,以及

$$p(x'_i|y) = p_2^i(z_2^i)p_1^i(z_1^i) = p_1^i(z_1^i)p_2^i(y_i + z_1^i - x'_i), \tag{16}$$

$i = 1, 2, \cdots, N$.状态$x'_i = y_i - z_2^i + z_1^i$表示在港口$i$被租用的箱子在船队到达港口$i$时就归还了.

类似的,一步期望费用$c(x,a)$可以写成为

$$c(x,a)=\sum_{i=1}^{N}[c_u^i(a_i)^+ + c_t^i(-a_i)^+ + c_h^i E(y_i-Z_2^i)^+ + c_p^i E(Z_2^i-y_i)^+], \quad (17)$$

其中$(a)^+=\max\{a,0\}$.我们依然用$L(y_i)$记在港口i的一步期望存储费用加上外租费用,$i=1,2,\cdots,N$,即:

$$L(y_i)=c_h^i E(y_i-Z_2^i)^+ + c_p^i E(Z_2^i-y_i)^+, \quad (18)$$

类似于单港口的情形,对任意的折扣因子$\beta\in(0,1)$,0到n阶段的最优值函数$V_\beta^n(x)$满足

$$
\begin{aligned}
V_\beta^n(x) &= \min_{A(x)}\Bigg\{\sum_{i=1}^{N}[c_u^i(a_i)^+ + c_t^i(-a_i)^+ + L(y_i)]+ \\
&\quad +\beta\sum_{z_1^1,z_2^1=0}^{K^1}\cdots\sum_{z_1^N,z_2^N=0}^{K^N} V_\beta^{n-1}(x')\prod_{j=1}^{N} p_1^j(z_1^j)p_2^j(z_2^j)\Bigg\}, \qquad (19)
\end{aligned}
$$

其中$x'=(y_1+z_1^1-z_2^1, y_2+z_1^2-z_2^2,\cdots,y_N+z_1^N-z_2^N)$为下一个阶段的某个状态.

类似于单港口的情形,对任意的折扣因子$\beta \in (0,1)$,0到n阶段的无限阶段最优折扣值函数$V_\beta^*(x)$分别满足

$$\begin{aligned} V_\beta^*(x) &= \min_{A(x)} \Bigg\{ \sum_{i=1}^{N} [c_u^i(a_i)^+ + c_t^i(-a_i)^+ + L(y_i)] + \\ & + \beta \sum_{z_1^1, z_2^1=0}^{K^1} \cdots \sum_{z_1^N, z_2^N=0}^{K^N} V_\beta^*(x') \prod_{j=1}^{N} p_1^j(z_1^j) p_2^j(z_2^j) \Bigg\}, \end{aligned} \tag{20}$$

其中$x' = (y_1 + z_1^1 - z_2^1, y_2 + z_1^2 - z_2^2, \cdots, y_N + z_1^N - z_2^N)$为下一个阶段的某个状态.

最后,我们的目的就是要获得长期平均费用最小的策略π^*,它满足:

$$\bar{V}(x,\pi^*) = \min_{\pi\in\Pi} \limsup_{n\to\infty} \frac{V_\beta^n(x,\pi)}{n+1}, \tag{21}$$

其中$\beta = 1$.这样我们就一般的建立了多港口的MDP问题,求解后就可以得到最优的控制策略π^*.但是,这样做的问题在于N和M很大时,MDP的状态空间会非常的巨大,使得求解的问题变得几乎不可能.不仅如此,最优策略的结构也不能保证具有易于操作的形式.

如果我们孤立的考虑一个港口i时,已经知道对于平均的最优准则,存在形如(U_i^*, D_i^*)的最优平稳策略.在考虑两个港口的情形时就变得复杂起来,原因在于两个港口的各项费用并不一致,会造成一个港口已经超过了它的控制限(需要调整空集装箱),但是此时另一个港口不一定也需要调整空集装箱的库存水平.

对于港口i来说,如果控制限(U_i^*, D_i^*) 变成为(U_i, D_i^*)或者(U_i^*, D_i)的时候,长期平均的费用函数都会增加.因此,我们可以找到两个增加率Δ_1^i和Δ_2^i,而用$V(U_i, D_i^*) = \bar{V}(U_i^*, D_i^*) + \Delta_1^i|U_i^* - U_i|$和$V(U_i^*, D_i) = \bar{V}(U_i^*, D_i^*) + \Delta_2^i|D_i^* - D_i|$来近似$\bar{V}(U_i, D_i^*)$和$\bar{V}(U_i^*, D_i)$ 的值,可以使我们得到一个近似的调整算法.

多港口空集装箱任一状态的调配方案近似算法

步骤1: 对于每个$i=1,2,\cdots,N$,分别记Δ_1^i和Δ_2^i为港口i的左侧和右侧费用近似增加率; (U_i^*,D_i^*)为单独考虑港口i时的最优控制限策略;对任一状态$x=(x_1,x_2,\cdots,x_N)\in S$,定义三个状态的集合: $I=\{i|x_i<U_i^*\}$, $J=\{i|U_i^*\le x_i\le D_i^*\}$和$K=\{i|x_i>D_i^*\}$.

步骤2: 如果$K\neq\emptyset$而且$I\neq\emptyset$,进入步骤3;如果$K=\emptyset$,而且$I\neq\emptyset$, 进入步骤4;如果$I=\emptyset$,而且$K\neq\emptyset$,进入步骤5;否则,进入步骤6.

步骤3: 取$\Delta_2^{i^*}=\max_{i\in K}\Delta_2^i$以及$\Delta_1^{j^*}=\max_{i\in I}\Delta_1^i$.
将港口i^*的空集装箱下调到$x'_{i^*}=x_{i^*}-\min\{x_{i^*}-D^*_{i^*},U^*_{j^*}-x_{i^*}\}$
将港口j^*的空集装箱上调到$x'_{j^*}=x_{j^*}+\min\{x_{i^*}-D^*_{i^*},U^*_{j^*}-x_{i^*}\}$
如果$x'_{i^*}=D^*_{i^*}$,则将状态i^*从K中移到J
如果$x'_{j^*}=U^*_{j^*}$,则将状态j^*从I中移到J
回到步骤2.

步骤4: 若$I = \emptyset$,进入步骤5,否则取$\Delta_2^{i^*} = \min_{i\in J} \Delta_2^i$以及$\Delta_1^{j^*} = \max_{i\in I} \Delta_1^i$.
如果$\Delta_2^{i^*} \geq \Delta_1^{j^*}$,进入步骤5.否则,
将港口i^*的空集装箱下调到$x'_{i^*} = x_{i^*} - \min\{x_{i^*} - U^*_{i^*}, U^*_{j^*} - x_{i^*}\}$
将港口j^*的空集装箱上调到$x'_{j^*} = x_{j^*} + \min\{x_{i^*} - U^*_{i^*}, U^*_{j^*} - x_{i^*}\}$
如果$x_{j^*} = D^*_{j^*}$,则将状态j^*从I中移到J
重复步骤4.

步骤5: 若$K = \emptyset$,进入步骤6,否则取$\Delta_1^{i^*} = \min_{i\in J}\Delta_1^i$以及$\Delta_2^{j^*} = \max_{i\in K}\Delta_2^i$.
如果$\Delta_1^{i^*} \geq \Delta_2^{j^*}$,进入步骤6.否则,
将港口i^*的空集装箱上调到$x'_{i^*} = x_{i^*} + \min\{x_{j^*} - D^*_{j^*}, D^*_{i^*} - x_{i^*}\}$
将港口j^*的空集装箱下调到$x'_{j^*} = x_{j^*} - \min\{x_{j^*} - D^*_{j^*}, D^*_{i^*} - x_{i^*}\}$
如果$x_{j^*} = D^*_{j^*}$,则将状态j^*从K中移到J
重复步骤5.

步骤6: 令$y = x'$,最优行动$a = y - x$,停止.

经过上述算法,我们可以从状态x得到一个可用的调整行动a和集装箱调整后的状态y.

文献[106]针对具体一些实际的例子模拟了我们给出的策略的效果,比较分别单独考虑每个港口的两点最优控制限策略的结果,费用比非常接近,是决策者能够接受的结果.

谢谢大家！