Санкт-Петербургский политехнический университет Петра Великого
Институт компьютерных наук и технологий
Кафедра компьютерных систем и программных технологий

РАСЧЕТНОЕ ЗАДАНИЕ
Тема: «Марковские модели принятия решений»
Дисциплина: «Методы оптимизации и принятия решений»
Вариант № 24

Выполнила студентка гр. 13541/3: Сагадеева С.А.
(подпись__________)
Руководитель, к.т.н., доц. Сиднев А. Г.
(подпись)__________
“16 ” апреля 2019 г.

Санкт-Петербург
2019

Оглавление

Задание

Задача 40, в) из книги Г. Вагнера «Основы исследования операций»:
Капитан Р., служащий в одной судоходной компании, командует судном, совершающим регулярные рейсы между двумя портами A и B. Предположим, что продолжительность рейса составляет 1 сутки. Каждое утро капитан должен решить, стоит ли ему загружать судно имеющимся в наличии грузом и отправляться в порт назначения или обождать сутки в надежде, что на следующий день может подвернуться более выгодный груз. Пусть затраты на один рейс составляют c_1, а затраты, связанные с суточным простоем судна в порту, составляют c_2, где $c_1 > c_2$. Предположим, что в порту A имеется два вида грузов, стоимостью a_1 и a_2, где $a_1 > a_2$ Обозначим вероятность того, что груз вида a_1 имеется в наличии, символом p_a. (откуда $(1 - p_a)$ есть вероятность того, что имеется только груз вида a_2). Предположим также, что наличие груза в рассматриваемый день не зависит от его наличия в предыдущие дни (таким образом, если капитан не уходит в рейс, то все равно сохраняется вероятность p_a получения груза a_1 на следующий день). Аналогично пусть стоимость грузов в порту B составляет b_1 и b_2, где $b_1 > b_2$, и пусть p_b — вероятность наличия груза b_1.

1. построить марковскую модель принятия решений для данной задачи;
2. для случая N= ∞ найти и перечислить все стационарные стратегии;
3. построенную моделью снабдить разумными численными данными и найти оптимальную стратегию для N=3.

Выполнение расчетного задания

Марковская модель принятия решения

Пусть имеются следующие состояния:

S_{a1} - судно находится в порту A и в наличии имеется груз a_1
S_{a2} - судно находится в порту A и в наличии имеется груз a_2
S_{b1} - судно находится в порту B и в наличии имеется груз b_1
S_{b2} - судно находится в порту B и в наличии имеется груз b_2

Множество решений:

1. X_1 - капитан принимает решение отправляться в порт назначения с имеющимся грузом
2. X_2 - капитан принимает решение ждать сутки

Множество допустимых решений для каждого из 4-х состояний:
S_{a1}: X_1

$S_{a2}: X_1, X_2$
$S_{b1}: X_1$

$S_{b2}: X_1, X_2$

Количество стационарных стратегий = 1* 2* 2 * 1 = 4:

$$\begin{matrix} X1 & X1 & X1 & X1 \\ X1 & X2 & X1 & X2 \\ X1 & X1 & X1 & X1 \\ X1 & X1 & X2 & X2 \end{matrix}$$

Теперь можно построить матрицы переходных вероятностей и расходов для каждого из допустимых решений. Далее индексы при P и R определяют номер допустимого решения, которому соответствуют. Т.е. P1 для Sa1, R2 – для Sa2 и т.п.

Матрицы переходов строились с учетом того, что строки и столбцы соответствуют состояниям в моменты времени i и i+1 соответственно.

1 допустимое решение: матрица переходных вероятностей

$$P_1 = \begin{pmatrix} 0 & 0 & p_b & 1-p_b \\ 0 & 0 & p_b & 1-p_b \\ p_a & 1-p_a & 0 & 0 \\ p_a & 1-p_a & 0 & 0 \end{pmatrix}$$

Нулями обозначены те переходы, вероятность которых равна нуля, так как в из S_{a1} в S_{a1}, из S_{a2} в S_{a1}, из S_{a1} в S_{a2}, из S_{a2} в S_{a2} капитан отправляется в порт B в любом случае, соответственно остаться в порту A – невозможно.

Аналогично для случаев из S_{b1} в S_{b1}, из S_{b2} в S_{b1}, из S_{b1} в S_{b2}, из S_{b2} в S_{b2} вероятности переходов равны нулю, поскольку опять же стремясь в порт B, капитан не может оказаться в порту A.

В порту B капитан окажется с вероятностью наличия груза b1, поэтому вероятность перехода из состояния S_{a1}/S_{a2} в S_{b2} равна $1 - p_b$.

В порту A капитан окажется с вероятностью того, что в порту будет груз a1, поэтому вероятность перехода из состояния S_{b1}/S_{b2} в S_{a2}, равна $1 - p_a$.

Матрица доходов для данного решения имеет следующий вид:

$$R_1 = \begin{pmatrix} 0 & 0 & a_1 - c_1 & a_1 - c_1 \\ 0 & 0 & a_2 - c_1 & a_2 - c_1 \\ b_1 - c_1 & b_1 - c_1 & 0 & 0 \\ b_2 - c_1 & b_2 - c_1 & 0 & 0 \end{pmatrix}$$

При переходе из S_{a1} в S_{b1}/S_{b2} будут получены деньги за груз a1, однако с учетом затрат в размере c1 мы получим итоговую прибыль a1-c1.

В случае перевозки груза a2 в пункт b деньги будут получены за груз a2, но с учетом затрат результирующий доход составит a2-c2.

При перевозке груза из S_{b1} в S_{a1}/S_{a2} капитан получит деньги за груз b_1, однако с учетом затрат на рейс, в результате получит b1-c1.

При перевозке груза из состояния S_{b2} в состояние S_{a1}/S_{a2}, капитан получит деньги за груз b_2, но при этом затраты на один рейс составят c_1, поэтому итоговый доход будет равен $b_2 - c_1$.

2 допустимое решение: матрица переходных вероятностей

В данном сценарии капитан принимает решение подождать сутки более ценный груз:

$$P_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ p_a & 1-p_a & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & p_b & 1-p_b \end{pmatrix}$$

Так как груз а1 более ценен чем а2, то ожидать сутки его смысла нет, поэтому переходы из S_{a1} в S_{a1} и из S_{a1} в S_{a2} равны 0.

Вероятность переходов из S_{b1} в S_{b1} и из S_{b1} в S_{b2} тоже равны 0 в силу аналогичной выше причины неравной ценности грузов.

Вероятность перехода из состояния S_{a2} в S_{a1} равна p_a (ценный груз вида a_1 будет в наличии). Вероятность перехода из состояния S_{a2} в S_{a2} равна $1-p_a$, поскольку именно с такой вероятностью будет в наличии груз вида a_2.

Остальные вероятности переходов равны нулю, так как капитан принял решение ожидать груз в пункте А или В соответственно.

Матрица доходов

$$R_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -c_2 & -c_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -c_2 & -c_2 \end{pmatrix}$$

Доходов в данном сценарии нет, так как капитан решил остаться в порту, затратив ресурсы на простой в размере с2.

Применение метода итераций по стратегиям

Рассмотрим все стационарные стратегии поведения. При 4 состояниях и 2 допустимых решениях количество стационарных решений было получено ранее и оно равно 4. В нашей задаче не имеет смысла рассматривать стратегии, при которых находясь в порту А при наличии груза a_1) и находясь в порту В при наличии груза b_1), судно принимает решение о простое, так как при наличии более ценного груза не имеет смысла ждать еще один день.

Для нахождения оптимальной стратегии при N=3 необходимо снабдить модель численными данными. Пусть:

$$a_1 = 100$$
$$a_2 = 80$$
$$b_1 = 90$$
$$b_2 = 55$$
$$c_1 = 50$$
$$c_2 = 25$$

$$p_a = 0.5$$
$$p_b = 0.7$$

В рассматриваемом случае горизонт планирования N=3, а элементы матриц переходных вероятностей и доходов не зависят от номера этапа, поэтому рассмотрим каждую стратегию в соответствии с определенными числовыми показателями. В рамках данной задачи ожидаемые доходы обусловлены только переходом изучаемой системы S из одного состояния в другое в рамках 4х стратегий:

Стратегия 1 $\{X_1, X_1, X_1, X_1\}$: $S(a_1) \rightarrow B$ (рейс из порта А в порт В с грузом a_1):

1) Находясь в состоянии S_{a1}, отправляться в порт В с грузом a_1
2) Находясь в состоянии S_{a2}, отправляться в порт В с грузом a_2
3) Находясь в состоянии S_{b1}, отправляться в порт А с грузом b_1
4) Находясь в состоянии S_{b2}, отправляться в порт А с грузом b_2

$$P_1 = \begin{pmatrix} 0 & 0 & 0.7 & 0.3 \\ 0 & 0 & 0.7 & 0.3 \\ 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \end{pmatrix} R_1 = \begin{pmatrix} 0 & 0 & 50 & 50 \\ 0 & 0 & 30 & 30 \\ 40 & 40 & 0 & 0 \\ 5 & 5 & 0 & 0 \end{pmatrix}$$

Стратегия 2 $\{X_1, X_2, X_1, X_1\}$: $S(a_2) \rightarrow B$ (рейс из порта А в порт В с грузом a_2)

1) Находясь в состоянии S_{a1}, отправляться в порт В с грузом a_1
2) Находясь в состоянии S_{a2}, ожидать сутки груз вида a_1
3) Находясь в состоянии S_{b1}, отправляться в порт А с грузом b_1
4) Находясь в состоянии S_{b2}, отправляться в порт А с грузом b_2

$$P_2 = \begin{pmatrix} 0 & 0 & 0.7 & 0.3 \\ 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \end{pmatrix} R_2 = \begin{pmatrix} 0 & 0 & 50 & 50 \\ -25 & -25 & 0 & 0 \\ 40 & 40 & 0 & 0 \\ 5 & 5 & 0 & 0 \end{pmatrix}$$

Стратегия 3 $\{X_1, X_1, X_1, X_2\}$: $S(b_1) \rightarrow A$ (рейс из порта В в порт А с грузом b_1)

1) Находясь в состоянии S_{a1}, отправляться в порт В с грузом a_1
2) Находясь в состоянии S_{a2}, отправляться в порт В с грузом a_2
3) Находясь в состоянии S_{b1}, отправляться в порт А с грузом b_1
4) Находясь в состоянии S_{b2}, ожидать сутки груз вида b_1

$$P_3 = \begin{pmatrix} 0 & 0 & 0.7 & 0.3 \\ 0 & 0 & 0.7 & 0.3 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.7 & 0.3 \end{pmatrix} R_3 = \begin{pmatrix} 0 & 0 & 50 & 50 \\ 0 & 0 & 30 & 30 \\ 40 & 40 & 0 & 0 \\ 0 & 0 & -25 & -25 \end{pmatrix}$$

Стратегия 4 $\{X_1, X_2, X_1, X_2\}$:$S(b_2) \rightarrow A$ (рейс из порта В в порт А с грузом b_2)

1) Находясь в состоянии S_{a1}, отправляться в порт B с грузом a_1
2) Находясь в состоянии S_{a2}, ожидать сутки груз вида a_1
3) Находясь в состоянии S_{b1}, отправляться в порт A с грузом b_1
4) Находясь в состоянии S_{b2}, ожидать сутки груз вида b_1
5) $P_4 = \begin{pmatrix} 0 & 0 & 0.7 & 0.3 \\ 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.7 & 0.3 \end{pmatrix} R_4 = \begin{pmatrix} 0 & 0 & 50 & 50 \\ -25 & -25 & 0 & 0 \\ 40 & 40 & 0 & 0 \\ 0 & 0 & -25 & -25 \end{pmatrix}$

Алгоритм метода итераций по стратегиям включает в себя следующие шаги [1]:

1. Шаг оценки параметров. Для произвольной стратегии s с матрицами P_s и R_s решаем систему из t уравнений

$$f^s(i) - \alpha \sum_{j=1}^{m} p_{ij}^s f^s(j) = v_i^s, i = 1,2,..., m$$

относительно m неизвестных $f^s(1), f^s(2),..., f^s(m)$.

2. Шаг улучшения стратегии. Для каждого состояния i определяем альтернативу k, обеспечивающую

$$\max_k \left\{ v_i^k + \alpha \sum_{j=1}^{m} p_{ij}^k f^s(j) \right\}, i = 1,2,..., m,$$

где $f^s(j)$ имеют значения, определенные на шаге оценки параметров. Если улучшенная стратегия t совпадает со стратегией s, то алгоритм закончен. В этом случае стратегия t оптимальна. В противном случае полагаем $s = t$ и повторяем шаг оценки параметров [2].

Таким образом, для нашей задачи:
Выберем в качестве произвольной стратегии τ стратегию №1, тогда при $\alpha = 0.9$:

$$P_1(\tau) = \begin{pmatrix} 0 & 0 & 0.7 & 0.3 \\ 0 & 0 & 0.7 & 0.3 \\ 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \end{pmatrix} R_1(\tau) = \begin{pmatrix} 0 & 0 & 50 & 50 \\ 0 & 0 & 30 & 30 \\ 40 & 40 & 0 & 0 \\ 5 & 5 & 0 & 0 \end{pmatrix}$$

$$f_\tau(1) - 0.9\big(0.7 * f_\tau(3) + 0.3 * f_\tau(4)\big) = 50$$
$$f_\tau(2) - 0.9\big(0.7 * f_\tau(3) + 0.3 * f_\tau(4)\big) = 30$$
$$f_\tau(3) - 0.9 * \big(0.5 * f_\tau(1) + 0.5 * f_\tau(2)\big) = 40$$
$$f_\tau(4) - 0.9 * \big(0.5 * f_\tau(1) + 0.5 * f_\tau(2)\big) = 5$$

Скрипт для решения системе в Matlab:

```
A = [1 0 -0.63 -0.27; 0 1 -0.63 -0.27; -0.45 -0.45 1 0; -0.45 -0.45 0 1];
```

```
b = [50; 30; 40; 5];
x = inv(A)*b
```

Полученные решения:

$$f_{\tau}(1) = 360.2632$$
$$f_{\tau}(2) = 340.2632$$
$$f_{\tau}(3) = 355.2368$$
$$f_{\tau}(4) = 320.2368$$

Улучшим стратегию при помощи полученных ранее результатов:

	$\varphi_j(X_i) = \nu_j(X_i) + \alpha \sum_{k=1}^{4} p_{jk}(X_i) F_\tau(k)$		$\max_G \varphi_j$	X_{*j}
	$i = 1$	$i = 2$		
S_1	$50 + 0.9 *$ (0.7*355.2368+0.3*320.2368) = 360.2632	$50 + 0.9 *$ (0.7*355.2368+0.3*320.2368) = 360.2632	360.2632	X_1
S_2	$30 + 0.9 *$ (0.7*355.2368+0.3*320.2368) = 340.2632	$-25 + 0.9 *$ (0.5*360.2632+0.5*340.2632) = 290.2368	340.2632	X_1
S_3	$40 + 0.9 * (0.5 * 360.2632 + 0.5 * 340.2632)$ = 355.2368	$40 + 0.9 *$ (0.5*360.2632+0.5*340.2632) = 355.2368	355.2368	X_1
S_4	$5 + 0.9 * (0.5 * 360.2632 + 0.5 * 340.2632)$ = 320.2368	$5 + 0.9 *$ (0.5*360.2632+0.5*340.2632) = 320.2368	320.2368	X_1

Так как в данном подходе стратегия была выбрана нами, а результат подтверждает ее достаточную оптимальность, то при N=∞ мы приходим к выводу, что капитану следует отправляться в противоположный порт без простоев.

Решение методом итераций по стратегиям с учетом дисконтирования при N=3

Так как горизонт планирования конечен, то для оптимальных ожидаемых доходов должны быть выполнены условия:

$$f_{N+1}(j) \equiv 0, j = 1..m$$

Оптимальный ожидаемый доход fi(j) на каждом из этапов (у нас их будет 3) определяется составляющей, определяемой по формуле [3]:

$$f_i(j) = \max_{X_{l_i} \epsilon G} \left(\nu_j(X_{l_i}) + \alpha \sum_{k=1}^{m} p_{jk}(i+1|X_{l_i})\, f_{i+1}(k) \right)$$

Где

$$\nu_j(X_{l_i}) = \sum_{k=1}^{m} p_{jk}(i+1 \mid X_{l_i}) r_{jk}(i+1 \mid X_{l_i})$$

И их совокупностью с учетом переходных вероятностей. Ранее было получено количество стационарных стратегий = 1* 2* 2 * 1 = 4:

$$\begin{matrix} X1 & X1 & X1 & X1 \\ X1 & X2 & X1 & X2 \\ X1 & X1 & X1 & X1 \\ X1 & X1 & X2 & X2 \end{matrix}$$

В используемом нами подходе матрицы P1, R1, P2, R2 не зависят от номера этапа. С учетом этого, ожидаемые доходы, обусловленные переходом системы из одного возможного состояния в другое при $\alpha = 0{,}9$ приведены ниже:

$\nu_1(X_1)$ = 0,7*50+0,3*50 = 50

$\nu_2(X_1)$ = 0,7*30+0,3*30 = 30

$\nu_3(X_1)$ = 0,5*40+0,5*40 = 40

$\nu_4(X_1)$ =0,5*5+0,5*5 = 5

$\nu_1(X_2)$ = 0,7*50 +0,3*50 = 50

$\nu_2(X_2)$ = -0,5*25 – 0,5*25 = -25

$\nu_3(X_2)$ = 0,5*40+0,5*40 = 40

$\nu_4(X_2)$ =0,5*5+0,5*5 = 5

Этап 3:

	$\nu_j(X_i)$		$\max_G \varphi_j$	X_{*j}
	$i = 1$	$i = 2$		
S_1	50	50	50	$X1$
S_2	30	-25	30	X_1
S_3	40	40	40	X_1, X2
S_4	5	5	5	X_1, X2

$$f_{\tau}(1) = 50$$
$$f_{\tau}(2) = 30$$
$$f_{\tau}(3) = 40$$
$$f_{\tau}(4) = 5$$

Этап 2:

	$\varphi_j(X_i) = \nu_j(X_i) + \alpha \sum_{k=1}^{4} p_{jk}(X_i) F_3(k)$		$\max_G \varphi_j$	X_{*j}
	$i = 1$	$i = 2$		
S_1	$50 + 0.9 *$ (0.7*40+0.3*5) = 76,55	$50 + 0.9 *$ (0.7*40+0.3*5) = 76,55	76,55	X_1, X2
S_2	$30 + 0.9 *$ (0.7*40+0.3*5) = 56,55	$-25 + 0.9 *$ (0.5*50+0.5*30) = 11	56,55	X_1
S_3	$40 + 0.9 * (0.5 * 50 + 0.5 * 30) = 76$	$40 + 0.9 *$ (0.5*50+0.5*30) = 76	76	X_1, X2
S_4	$5 + 0.9 * (0.5 * 50 + 0.5 * 30) = 41$	$5 + 0.9 *$ (0.5*50+0.5*30) = 41	41	X_1, X2

$$f_{\tau}(1) = 76{,}55$$
$$f_{\tau}(2) = 56{,}55$$
$$f_{\tau}(3) = 76$$
$$f_{\tau}(4) = 41$$

Этап 1:

	$\varphi_j(X_i) = \nu_j(X_i) + \alpha \sum_{k=1}^{4} p_{jk}(X_i) F_2(k)$		$\max_G \varphi_j$	X_{*j}
	$i = 1$	$i = 2$		
S_1	$75{,}55 + 0.9 *$ (0.7*76+0.3*41) = 134,5	$76{,}55 + 0.9 *$ (0.7*76+0.3*41) = 134,5	134,5	X_1, X2
S_2	$56{,}55 + 0.9 *$ (0.7*76+0.3*41) = 115,5	$56{,}55 + 0.9 *$ (0.5*76+0.5*41) = 109,2	115,5	X_1

S_3	$76 + 0.9 * (0.5 * 76{,}55 + 0.5 * 30) = 129{,}275$	$76 + 0.9 *$ (0.5*76,55+0.5*30) = 129,275	129,275	X_1, X2
S_4	$41 + 0.9 * (0.5 * 76{,}55 + 0.5 * 56{,}55) = 100{,}895$	$41 + 0.9 *$ (0.5*76,55+0.5*56,55) = 100,895	100,895	X_1, X2

$$f_{\tau}(1) = 134{,}5$$
$$f_{\tau}(2) = 115{,}5$$
$$f_{\tau}(3) = 129{,}275$$
$$f_{\tau}(4) = 100{,}895$$

Уже на 3м этапе мы получаем одинаковые значения оптимальности для X1 и X2 для всех стратегий кроме 2й, где заметно преимущество первого решения. Такая степень аналогичности результатов, скорее всего, обусловлена случайным выбором численной модели данных.

Выводы

В ходе данной лабораторной работы нами были изучены марковские модели принятия решений, а также решение задач метод итераций при бесконечном горизонте планирования и при введении ограничения на N.

Использованная литература

1. Г.Вагнер «Основы исследования операций» Т.3, 1973г. – 493 стр.
2. И.М. Макаров, Т.М. Виноградская, А.А. Рубчинский, В.Б. Соколов «Теория выбора и принятия решений» М. – 1982 – 328 с.
3. Волков, Загоруйко «Исследование операций» - 2000 – 436 с.