Решение задачи динамического программирования

Ф.И.О., группа, поток

Клименков Владислав Максимович, К3341, МЕТОПТ 1.2

Предобработка условий задачи

Расчёт стоимости пакетов для покупки/продажи активов

В задании сказано, что покупать и продавать активы можно только пакетами, стоимость которых равна 25% от изначальной стоимости актива. Рассчитаем стоимость одного пакета для каждого актива:

- $\text{ЦБ1}: 100 \cdot 25\% = 25 \, (\text{д.е.})$
- $\text{ЦБ2}: 800 \cdot 25\% = 200 \, (\text{д.е.})$
- $\text{Деп.}: 400 \cdot 25\% = 100 \, (\text{д.е.})$

В дальнейшем будем использовать уже рассчитанные значения стоимости пакетов.

Описание задачи

У нас есть инвестиционный портфель, которым мы будем управлять на протяжении трёх этапов $k = 1$, $k = 2$ и $k = 3$, где $k = 1$ - первый этап, а $k = 3$ - последний этап.

На начало каждого этапа нам известно состояние инвестиционного портфеля S_k. Сам портфель может содержать четыре актива (в скобках указаны сокращения названий активов, а также индексы i, которые данные активы занимают в векторах и массивах):

- Ценная бумага 1 ($\text{ЦБ1}: i = 1$)

- Ценная бумага 2 ($\text{ЦБ2} : i = 2$)
- Депозиты ($\text{Деп.} : i = 3$)
- Свободные денежные средства ($\text{СДС} : i = 4$)

Стоимость активов измеряется в денежных единицах (д.е.).

Изначально в инвестиционном портфеле следующее распределение активов:

- $\text{ЦБ1} = 100\,\text{д.е.}$
- $\text{ЦБ2} = 800\,\text{д.е.}$
- $\text{Деп.} = 400\,\text{д.е.}$
- $\text{СДС} = 600\,\text{д.е.}$

Состояние S_k задаётся вектором из четырёх чисел, соответствующих стоимости активов в портфеле по принципу

$$\begin{aligned}
&S_k = (\\
&\quad \langle \text{стоимость ЦБ1} \rangle, \\
&\quad \langle \text{стоимость ЦБ2} \rangle, \\
&\quad \langle \text{стоимость Деп.} \rangle, \\
&\quad \langle \text{стоимость СДС} \rangle \\
&).
\end{aligned}$$

Соответственно, для первого этапа мы имеем $S_1 = (100, 800, 400, 600)$.

Также на каждом этапе мы можем совершить операцию x_k по разбалансировке (покупке и продаже) активов. Покупать и продавать активы можно только пакетами (лотами). Вот размеры пакетов для каждого актива:

- $1\,\text{пакет ЦБ1} = 25\,\text{д.е.}$
- $1\,\text{пакет ЦБ2} = 200\,\text{д.е.}$
- $1\,\text{пакет Деп.} = 100\,\text{д.е.}$

Для каждого этапа операция записывается в формате вектора из трёх чисел, где каждое число соответствует числу покупаемых или продаваемых пакетов активов

$$\begin{aligned}
&x_k = (\\
&\quad \langle \text{количество купленных/проданных пакетов ЦБ1} \rangle, \\
&\quad \langle \text{количество купленных/проданных пакетов ЦБ2} \rangle, \\
&\quad \langle \text{количество купленных/проданных пакетов Деп.} \rangle \\
&).
\end{aligned}$$

Например, если на первом этапе мы хотим купить два пакета ЦБ1 (на 50 д.е.) и продать один пакет Деп. (100 д.е.), не трогая ЦБ2, то это можно записать как $x_1 = (2, 0, -1)$.

Также нужно учитывать, что по условиям задачи за каждую операцию берётся комиссия. Вот какие комиссии для операций с активами:

- $\text{ЦБ1} : 4\%$
- $\text{ЦБ2} : 7\%$
- $\text{Деп.} : 5\%$

То есть, например, при покупке или продаже одного пакета Деп. стоимостью 100 д.е. придётся заплатить комиссию в 5%, в данном случае равную 5 д.е.

Все покупки активов и комиссии по операциям оплачиваются из СДС. Например, при покупке n пакетов актива со стоимостью пакета C и комиссией p суммарные затраты из СДС составят $n \cdot C \cdot (1 + p)$. При продаже n пакетов в СДС поступает $n \cdot C \cdot (1 - p)$.

После совершения операций x_k наступает стохастическое событие w_k (в данном случае под событием подразумевается полное вероятностное пространство) для соответствующего этапа k. В рамках события w_k может реализоваться один из трёх сценариев $w_{k,t}$, где t - реализовавшийся сценарий: позитивный сценарий ($t = 1$) $w_{k,1}$, нейтральный сценарий ($t = 2$) $w_{k,2}$ или негативный сценарий ($t = 3$) $w_{k,3}$. Ниже представлены таблицы с описанием событий w_k для всех трёх этапов. В таблицах указаны вероятности наступления того или иного сценария, а также коэффициент изменения стоимости актива при реализации данного сценария:

Этап №1:

Ситуация	Вероят.	ЦБ1	ЦБ2	Деп.
Благопр.	0.60	1.20	1.10	1.07
Нейтр.	0.30	1.05	1.02	1.03
Негатив.	0.10	0.80	0.95	1.00

Этап №2:

Ситуация	Вероят.	ЦБ1	ЦБ2	Деп.
Благопр.	0.30	1.4	1.15	1.01
Нейтр.	0.20	1.05	1.00	1.00
Негатив.	0.50	0.60	0.90	1.00

Этап №3:

Ситуация	Вероят.	ЦБ1	ЦБ2	Деп.
Благопр.	0.40	1.15	1.12	1.05
Нейтр.	0.40	1.05	1.01	1.01
Негатив.	0.20	0.70	0.94	1.00

То есть, например, если на втором этапе реализуется негативный сценарий (это произойдёт с вероятностью 0.50), то стоимость ЦБ2 будет домножена на коэффициент 0.90 (т.е. стоимость ЦБ2 снизится на 10%), если же на этом же этапе реализуется благоприятный сценарий (это произойдёт с вероятностью 0.30), то стоимость ЦБ2 будет домножена на коэффициент 1.15 (т.е. стоимость ЦБ2 вырастет на 15%). Аналогично и для других этапов, сценариев и активов.

Событие w_k можно представить в виде вектора из трёх сценариев $t = 1, 2, 3$:

$$w_k = (w_{k,1}, w_{k,2}, w_{k,3}),$$

где $w_{k,t}$ - конкретный сценарий.

Каждый сценарий $w_{k,t}$ задаётся вектором из трёх коэффициентов изменения стоимости активов ЦБ1, ЦБ2 и Деп. соответственно:

$$w_{k,t} = (w^1_{k,t}, w^2_{k,t}, w^3_{k,t}),$$

где $w^i_{k,t}$ - коэффициент изменения стоимости соответствующего актива i (см. соответствие индексов $i = 1, 2, 3, 4$ активам, указанное выше).

Стоимость СДС не изменяется в рамках события w_k, а потому для него не указывается коэффициент изменения стоимости.

Например, нейтральный сценарий второго этапа можно записать следующим образом:

$$w_{2,2} = (1.05, 1.01, 1.01)$$

Вероятность конкретного сценария задаётся функцией вероятности P_k:

$$P_k(t) = p_{k,t},$$

где $p_{k,t}$ - вероятность наступления сценария t на этапе k.

Например, для нейтрального сценария $t = 2$ первого этапа $k = 1$ имеем:

$P_1(2) = p_{1,2} = 0.30$

В результате мы получаем следующую последовательность:

$S_k \to x_k \to w_k \to S_{k+1}$

То есть сначала мы смотрим состояние на данном этапе, потом совершаем операции на данном этапе, потом происходит событие на данном этапе, после чего мы смотрим состояние уже на следующем этапе.

Также по заданию существуют ограничения, которые задают минимальную стоимость каждого актива в портфеле:

- ЦБ1 : $\text{Не менее } 30 \text{ д.е.}$
- ЦБ2 : $\text{Не менее } 150 \text{ д.е.}$
- Деп. : $\text{Не менее } 100 \text{ д.е.}$
- СДС : $\text{Не менее } 0 \text{ д.е.}$ (СДС не может быть отрицательным)

То есть, например, в любой момент стоимость ЦБ2 в портфеле не может опускаться ниже 150 д.е. (нельзя продать, чтобы стоимость стала меньше, и нельзя допустить, чтобы в рамках события стоимость ЦБ2 опустилась ниже данного значения). Аналогично и для других активов.

Цель задачи

Требуется разработать такой план управления покупками/продажами активов на этапах $k = 1, 2, 3$, чтобы максимизировать математическое ожидание общей стоимость инвестиционного портфеля в конце третьего этапа (S_4).

Математическая постановка задачи

В рамках задачи рассматриваются три этапа $k = 1, 2, 3$ управления инвестиционным портфелем, который содержит следующие активы (в скобках указаны индексы данных активов, согласно которым они будут располагаться в рамках массивов и векторов):

- ЦБ1 ($i = 1$),
- ЦБ2: ($i = 2$),
- Деп.: ($i = 3$),
- СДС: ($i = 4$).

Состояние S_k активов в портфеле в начале этапа k задаётся вектором из четырёх чисел:

$$S_k = (S_k^1, S_k^2, S_k^3, S_k^4),$$

где:

- S_k^1 - стоимость ЦБ1,
- S_k^2 - стоимость ЦБ2,
- S_k^3 - стоимость Деп.,
- S_k^4 - стоимость СДС

(позиции активов в векторе, как было сказано выше, соответствуют их индексам $i = 1, 2, 3, 4$, значения которых также были указаны выше).

Дано состояние портфеля на начало первого этапа:

$$S_1 = (100, 800, 400, 600)$$

На каждом из трёх этапов $k = 1, 2, 3$ совершается операция x_k ребалансировки портфеля, которая задаётся вектором из трёх чисел:

$$x_k = (x_k^1, x_k^2, x_k^3),$$

где x_k^i - количество купленных/проданных пакетов соответствующего актива ($x_k^i > 0$ - покупка, $x_k^i < 0$ - продажа, $x_k^i = 0$ - отсутствие операций купли/продажи).

Стоимость одного пакета для активов задаётся следующим массивом:

$$L = [25, 200, 100]$$

(позиции значений соответствуют индексам $i = 1, 2, 3$ активов).

Соответственно, стоимость актива i после применения операции x_k^i изменяется на следующую величину:

$\Delta_k^i = x_k^i \cdot L[i]$ для $i = 1, 2, 3$ (торгуемые активы: ЦБ1, ЦБ2, Деп.), где Δ_k^i - изменение стоимости актива i на этапе k после применения операции x_k^i.

В результате стоимость торгуемых активов изменяется следующим образом:

$(S'_k)^i = S_k^i + \Delta_k^i$ для $i = 1, 2, 3$

Также за каждую операцию купли/продажи приходится платить определённую комиссию. Комиссия для активов указаны в следующем массиве:

$c = [0.04, 0.07, 0.05]$

Суммарная стоимость всех операций (с учётом комиссии):

$$Cost(x_k) = \Sigma_{i=1}^{3}(\Delta_k^i + c[i] \cdot |\Delta_k^i|)$$

Здесь $|\Delta_k^i|$ - абсолютная стоимость операции, на которую начисляется комиссия.

Операции купли/продажи, а также взятие комиссии происходят за счёт СДС, в результате чего сразу после применения операции x_k значение СДС $(S_k')^4$ равняется:

$$(S_k')^4 = S_k^4 - Cost(x_k)$$

В результате, после применения операции x_k к портфелю с состоянием S_k мы получаем портфель с состоянием S_k':

$$S_k' = ((S_k')^1, (S_k')^2, (S_k')^3, (S_k')^4).$$

После применения операции x_k происходит стохастическое событие w_k (в данном случае под событием подразумевается также и полное вероятностное пространство), которое содержит три возможных сценария $w_{k,t}$:

- $w_{k,1}$ - позитивный сценарий ($t = 1$),
- $w_{k,2}$ - нейтральный сценарий ($t = 2$),
- $w_{k,3}$ - негативный сценарий ($t = 3$),

Соответственно, событие w_k задаётся множеством из трёх данных сценариев $w_{k,t}$ с $t = 1, 2, 3$:

$$w_k = w_{k,1}, w_{k,2}, w_{k,3}$$

В свою очередь, каждый сценарий $w_{k,t}$ задаётся вектором из трёх коэффициентов изменения стоимости активов ЦБ1, ЦБ2 и Деп. соответственно:

$$w_{k,t} = (w_{k,t}^1, w_{k,t}^2, w_{k,t}^3),$$

где $w_{k,t}^i$ - коэффициент изменения стоимости соответствующего актива i (стоимость СДС не изменяется в рамках события w_k, а потому для него не указывается коэффициент изменения стоимости).

Вероятность конкретного сценария задаётся функцией вероятности P_k:

$$P_k(w_k = w_{k,t}) = p_{k,t},$$

где $p_{k,t}$ - вероятность наступления сценария t на этапе k.

По условиям задачи, для всех сценариев на всех трёх этапах нам известны их вероятности и коэффициенты изменения стоимости активов:

Вероятности реализации сценариев:

- Этап 1:
    - $p_{1,1} = 0.60$
    - $p_{1,2} = 0.30$
    - $p_{1,3} = 0.10$
- Этап 2:
    - $p_{2,1} = 0.30$
    - $p_{2,2} = 0.20$
    - $p_{2,3} = 0.50$
- Этап 3:
    - $p_{3,1} = 0.40$
    - $p_{3,2} = 0.40$
    - $p_{3,3} = 0.20$

Коэффициенты изменения стоимости активов:

- Этап 1:
    - $w_{1,1} = (1.20, 1.10, 1.07)$
    - $w_{1,2} = (1.05, 1.02, 1.03)$
    - $w_{1,3} = (0.80, 0.95, 1.00)$
- Этап 2:
    - $w_{2,1} = (1.4, 1.15, 1.01)$
    - $w_{2,2} = (1.05, 1.00, 1.00)$
    - $w_{2,3} = (0.60, 0.90, 1.00)$
- Этап 3:
    - $w_{3,1} = (1.15, 1.12, 1.05)$
    - $w_{3,2} = (1.05, 1.01, 1.01)$
    - $w_{3,3} = (0.70, 0.94, 1.00)$

В результате наступления сценария $w_{k,t}$ стоимость актива i изменяется следующим образом:

$S^i_{k+1} = (S'_k)^i \cdot w^i_{k,t}$ для $i = 1, 2, 3,$

где S^i_{k+1} - стоимость актива i после реализации сценария $w_{k,t}$, а также стоимость актива i в начале этапа $k + 1$ (поле реализации сценария $w_{k,t}$ происходит переход от этапа k к этапу $k + 1$).

Соответственно, в рамках сценария $w_{k,t}$ происходит переход с этапа k к этапу $k + 1$ и состояние портфеля становится:

$S_{k+1} = (S^1_{k+1}, S^2_{k+1}, S^3_{k+1}, S^4_{k+1}),$

где $S^4_{k+1} = (S'_k)^4$ (стоимость СДС в рамках сценария $w_{k,t}$ не изменяется).

Также существует ограничения на минимально допустимую стоимость в любой момент времени для каждого актива. Ограничения задаются следующим массивом:

$S_{min} = [30, 150, 100, 0]$

(в рамках массива также учитывается, что СДС не может быть отрицательным, т.е. нельзя использовать кредитные средства).

В результате, на стоимость актива i действуют следующие ограничения:

- $S^i_k \geq S_{min}[i],$
- $(S'_k)^i \geq S_{min}[i],$
- $S^i_{k+1} \geq S_{min}[i].$

Видим, что в рамках задачи имеет место быть следующая последовательность (при реализации сценария $w_{k,t}$):

$S_k \rightarrow x_k \rightarrow S'_k \rightarrow w_{k,t} \rightarrow S_{k+1}$

Объединим шаги $x_k \rightarrow S'_k \rightarrow w_{k,t}$ в рамках функции перехода состояния T_k для этапа k и получим:

$S_{k+1} = T_k(S_k, x_k, w_{k,t}),$

где:

$S_{k+1} = (S^1_{k+1}, S^2_{k+1}, S^3_{k+1}, S^4_{k+1})$

для торгуемых активов $i = 1, 2, 3$:

$S_{k+1}^i = (S'_k)^i \cdot w_{k,t}^i$

$(S'_k)^i = S_k^i + \Delta_k^i$

$\Delta_k^i = x_k^i \cdot L[i]$

для СДС $i = 4$:

$(S'_k)^4 = S_k^4 - Cost(x_k)$

$Cost(x_k) = \Sigma_{i=1}^3 (\Delta_k^i + c[i] \cdot |\Delta_k^i|)$

Целевой функцией F назовём математическое ожидание суммарной стоимости всех активов в портфеле $E[sum(S_4)]$ после третьего этапа (критерий Байеса):

$F = E[sum(S_4)]$,

где:

$S_4 = (S_4^1, S_4^2, S_4^3, S_4^4)$ - состояние портфеля после третьего этапа (т.е. сразу после наступления события w_3).

В рамках данной задачи требуется максимизировать значение целевой функции:

$F = E[sum(S_4)] \to max$,

т.е. найти последовательность допустимых управлений (x_1, x_2, x_3), которая максимизирует целевую функцию $F = E[sum(S_4)]$ при заданном начальном состоянии S_1 портфеля и функциях перехода T_k.

Рекуррентное соотношение Беллмана

Определим функцию Беллмана $V_k(S_k)$ - максимальное ожидаемое значение целевой функции (матожидание суммарной стоимости портфеля на конец третьего этапа) при условии, что на этапе k мы имеем состояние портфеля S_k и применяем оптимальные управления на оставшихся этапах $k, k+1, ..., 3$.

1. Терминальное условие (этап $k = 4$)

После третьего этапа управления нет, ценность состояния - это просто сумма его компонент (общая стоимость портфеля).

$$V_4(S_4) = sum(S_4) = S_4^1 + S_4^2 + S_4^3 + S_4^4$$

2. Рекуррентное соотношение для этапов $k = 3, 2, 1$

Для этапа k состояние S_k известно. Мы выбираем управление x_k, которое приводит к промежуточному состоянию S'_k. Затем наступает случайное событие w_k с известным распределением вероятностей P_k, которое преобразует S'_k в следующее состояние S_{k+1}. Ценность выбранного управления x_k при данном S_k - это математическое ожидание ценности $V_{k+1}(S_{k+1})$ по всем возможным сценариям события w_k.

Оптимальное значение $V_k(S_k)$ - это максимум этого ожидания по всем допустимым управлениям x_k.

Формальная запись:

$$V_k(S_k) = max_{x_k \in D_k(S_k)} E_{w_k}[V_{k+1}(T_k(S_k, x_k, w_k))]$$

где:

- $D_k(S_k)$ – множество допустимых управлений x_k при данном состоянии S_k. Ограничения формируются из условий неотрицательности СДС после операции и минимальных значений активов:
    - $(S'_k)^i = S_k^i + x_k^i \cdot L[i] \geq S_{min}[i]$ для $i = 1, 2, 3$ (после ребалансировки).
    - $(S'_k)^4 = S_k^4 - Cost(x_k) \geq S_{min}[4] = 0$ (после ребалансировки).
    - Для всех $i = 1, 2, 3, 4$ должно также выполняться $S_{k+1}^i \geq S_{min}[i]$ после применения любого сценария события w_k.
- $E_{w_k}[...]$ – математическое ожидание по случайному событию w_k на этапе k. Поскольку распределение дискретное, его можно расписать в виде суммы:

  $$E_{w_k}[V_{k+1}(...)] = \Sigma_{t=1}^3 [p_{k,t} \cdot V_{k+1}(T_k(S_k, x_k, w_{k,t}))]$$

- $T_k(S_k, x_k, w_{k,t})$ – детерминированная функция перехода состояния, подробно описанная выше:

  $S_{k+1} = T_k(S_k, x_k, w_{k,t})$, где:

- Для торгуемых активов ($i = 1, 2, 3$):

  $$S_{k+1}^i = (S_k^i + x_k^i \cdot L[i]) \cdot w_{k,t}^i$$

- Для СДС ($i = 4$):

  $$Cost(x_k) = \Sigma_{i=1}^3 (x_k^i \cdot L[i] + c[i] \cdot |x_k^i \cdot L[i]|)$$

  $$S_{k+1}^4 = S_k^4 - Cost(x_k)$$

3. Итоговое рекуррентное уравнение

Объединяя все компоненты, получаем основное уравнение динамического программирования (обратная рекурсия):

Для $k = 3, 2, 1$:

$$V_k(S_k) = max_{x_k \in D_k(S_k)} (\Sigma_{t=1}^3 [p_{k,t} \cdot V_{k+1}(S_{k+1}(t))])$$

где $S_{k+1}(t)$ – результат детерминированного вычисления для конкретного сценария t:

$$\begin{aligned} &S_{k+1}(t) = (\\ &\quad (S_k^1 + x_k^1 \cdot L[1]) \cdot w_{k,t}^1, \\ &\quad (S_k^2 + x_k^2 \cdot L[2]) \cdot w_{k,t}^2, \\ &\quad (S_k^3 + x_k^3 \cdot L[3]) \cdot w_{k,t}^3, \\ &\quad S_k^4 - \Sigma_{i=1}^3 (x_k^i \cdot L[i] + c[i] \cdot |x_k^i \cdot L[i]|) \\ &) \end{aligned}$$

и $V_4(S_4) = S_4^1 + S_4^2 + S_4^3 + S_4^4$.

Цель задачи: Найти максимальное значение целевой функции $F = V_1(S_1)$ при заданном начальном состоянии $S_1 = (100, 800, 400, 600)$ и соответствующие ему оптимальные стратегии управления x_1, x_2, x_3.

Алгоритм решения задачи

Идея: Поскольку задача многоэтапная, стохастическая и имеет ограничения на управления и состояния, используем классический метод обратного хода. Из-за непрерывного пространства состояний применяется его дискретизация (сетка).

Шаг 1: Подготовка данных и параметров

Зафиксируем константы:

- $L = [25, 200, 100]$ (стоимость пакетов активов).
- $c = [0.04, 0.07, 0.05]$ (комиссии).
- $S_{min} = [30, 150, 100, 0]$ (минимальные стоимости).
- Начальное состояние $S_1 = [100, 800, 400, 600]$.
- Вероятности $p_{k,t}$ и коэффициенты $w^i_{k,t}$ для всех $k = 1, 2, 3$ и $t = 1, 2, 3$.

Шаг 2: Дискретизация пространства состояний

Пространство состояний $S_k = (S^1_k, S^2_k, S^3_k, S^4_k)$ 4-мерное и непрерывное. Для применения ДП необходимо перейти к конечному множеству точек.

1. Для каждого актива $i = 1, 2, 3, 4$ определить разумные границы изменения стоимости на каждом этапе. Например, можно взять от $S_{min}[i]$ до некоторой максимальной оценки, полученной из предположения о максимальном росте активов.

2. Разбить каждый диапазон на N_i точек (чем больше, тем точнее решение, но больше вычислений). Например, использовать равномерную сетку.

3. Получаем множество дискретных состояний G_k для каждого этапа k. В простейшем случае сетка одинакова для всех этапов.

Шаг 3: Вычисление функции ценности на терминальном этапе ($k = 4$)

Для каждого дискретного состояния S_4 из сетки G_4:

$V_4(S_4) = S^1_4 + S^2_4 + S^3_4 + S^4_4$ (просто сумма компонент).

Шаг 4: Обратный ход для $k = 3, 2, 1$

Для каждого этапа k, начиная с $k = 3$ до $k = 1$, и для каждого состояния S_k из сетки G_k выполнить:

Подшаг 4.1: Определение допустимой области управлений $D_k(S_k)$

1. Управление $x_k = (x^1_k, x^2_k, x^3_k)$ является целочисленным (т.к. торгуются только целые пакеты).

2. Заданы нижние границы для стоимости каждого торгового актива ($i = 1, 2, 3$):

$$S_k^i + x_k^i \cdot L[i] \geq S_{min}[i] \rightarrow x_k^i \geq (S_{min}[i] - S_k^i)/L[i]$$

3. Размер покупок не должен превышать стоимость СДС (СДС не может быть отрицательным):

$$Cost(x_k) = \Sigma_{i=1}^3 (x_k^i \cdot L[i] + c[i] \cdot |x_k^i \cdot L[i]|)$$

$$S_k^4 - Cost(x_k) \geq S_{min}[4] = 0$$

4. Ограничения на будущее состояние после сценариев:

   Для всех сценариев t должно быть $S_{k+1}^i(t) \geq S_{min}[i]$. Это накладывает дополнительные ограничения на x_k. Например, для худшего сценария ($w_{k,t}^i$ минимально) должно выполняться $(S_k^i + x_k^i \cdot L[i]) \cdot min(w_{k,t}^i) \geq S_{min}[i]$.

В результате для каждого S_k перебираем ограниченный диапазон целочисленных значений x_k^i (например, от -M до +M, где M - максимальное разумное число покупаемых/продаваемых пакетов) и отбрасываем те, что не удовлетворяют перечисленным условиям.

Подшаг 4.2: Вычисление ожидаемой ценности для каждого допустимого управления

Для каждого допустимого x_k из $D_k(S_k)$:

1. Для каждого сценария $t = 1, 2, 3$:

   - Рассчитать новое состояние $S_{k+1}(t)$ по формулам:

     $S_{k+1}^i(t) = (S_k^i + x_k^i \cdot L[i]) \cdot w_{k,t}^i$, для $i = 1, 2, 3$.

     $$S_{k+1}^4(t) = S_k^4 - Cost(x_k)$$

   - Полученное $S_{k+1}(t)$ скорее всего не попадет точно в узел дискретной сетки G_{k+1}, а значит для нахождения значения функции $V_{k+1}(S_{k+1}(t))$ необходимо прировнять $S_{k+1}(t)$ к ближайшему значению G_{k+1} (интерполяция по методу ближайшего соседа).

2. Вычислить ожидаемую ценность для данного x_k:

   $$Q_k(S_k, x_k) = \Sigma_{t=1}^3 [p_{k,t} * V_{k+1}(S_{k+1}(t))]$$

   (используем формулу математического ожидания).

Подшаг 4.3: Нахождение оптимального управления и значения функции Беллмана

Найти управление, максимизирующее ожидаемую ценность:

$$x_k^*(S_k) = argmax_{x_k \in D_k(S_k)} Q_k(S_k, x_k)$$

Сохранить это управление и соответствующее максимальное значение:

$$V_k(S_k) = Q_k(S_k, x_k^*(S_k))$$

Шаг 5: Прямой ход (Восстановление оптимальной траектории)

В результате предыдущего шага мы дойдём до этапа $k = 1$, где получим оптимальное управление $x_1^*(S_1)$ для уже известного нам начального состояния S_1. Далее остаётся только применить управление x_1^* к портфелю с состоянием S_1, получив состояние S_1'. После реализации стохастического события w_1 состояние портфеля изменится на S_2. Его нужно будет приробнять к ближайшему значению сетки G_2 для которого нам уже известно оптимальное управление $x_2^*(G_2)$. Аналогичным образом нужно будет пройти по всем оставшимся этапам до этапа $k = 3$, в результате чего будет получен набор оптимальных управлений (x_1^*, x_2^*, x_3^*), а также найдено максимальное ожидаемое значение целевой функции $F^* = V_1(S_1)$.

Стоит учитывать, что из-за стохастического характера событий, полное решение может состоять сразу из нескольких возможных траекторий, так как траектории для разных сценариев могут различаться. В крайнем случае, каждое событие может разбивать траекторию на три подтраектории, по одной подтраектории для каждого сценария.

Заключение

В рамках данной работы я разбирался с методом динамического программирования для решения задач оптимизации. Я проанализировал данную мне задачу (которая обладала весьма непростыми условиями), подробно расписал её в более свободной формулировке, а потом постарался максимально перевести её на математический язык. Также мной для данной задачи было составлено рекуррентное соотношение Беллмана, с помощью которого был реализован алгоритм решения задачи путём обратного прохода. В результате был получен опыт математической формулировки задач оптимизации и создание алгоритма их решения на основе динамического программирования.