НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені Ігоря СІКОРСЬКОГО»
Навчально-науковий фізико-технічний інститут
Кафедра математичних методів захисту інформації

«На правах рукопису»
УДК 519.21

«До захисту допущено»
В.о. завідувача кафедри
__________ Сергій ЯКОВЛЄВ
«__» __________ 2023 р.

Дипломна робота

на здобуття ступеня бакалавра
за освітньо-професійною програмою
«Математичні методи криптографічного захисту інформації»

зі спеціальності: 113 Прикладна математика
на тему: **«Оптимальне транспортування ймовірнісних розподілів на дискретних просторах»**

Виконав:
студент IV курсу, групи ФІ-93
Оржахівський Максим Олександрович __________

Керівник:
старший викладач кафедри ММЗІ, кандидат фіз. мат. наук
Рябов Георгій Валентинович __________

Рецензент:
кандидат фіз.-мат. наук, молодший науковий співробітник відділу теорії випадкових процесів Інституту математики НАН України
Вовчанський Микола Богданович __________

Засвідчую, що у цій дипломній роботі немає запозичень з праць інших авторів без відповідних посилань.

Студент __________

Київ — 2023

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені Ігоря СІКОРСЬКОГО»
Навчально-науковий фізико-технічний інститут
Кафедра математичних методів захисту інформації

Рівень вищої освіти — перший (бакалаврський)
Спеціальність — 113 Прикладна математика,
ОПП «Математичні методи криптографічного захисту інформації»

ЗАТВЕРДЖУЮ
В.о. завідувача кафедри
___________ Сергій ЯКОВЛЄВ
«__» ___________ 2023 р.

ЗАВДАННЯ
на дипломну роботу

Студент: Оржахівський Максим Олександрович

1. Тема роботи: «*Оптимальне транспортування ймовірнісних розподілів на дискретних просторах*», науковий керівник дисертації: старший викладач кафедри ММЗІ, кандидат фіз. мат. наук Рябов Георгій Валентинович,

затверджені наказом по університету №__ від «__» ___________ 2023 р.

2. Термін подання студентом роботи: «__» ___________ 2023 р.

3. Об’єкт дослідження: оптимальне транспортування ймовірнісних розподілів на дискретних просторах

4. Предмет дослідження: простір Хеммінга, оператор дискретної похідної, ймовірнісний розподіл

5. Перелік завдань: провести огляд опублікованих джерел за тематикою дослідження; розв’язати завдання оптимального транспортування на дво-точковому просторі; аналіз дуальності Канторовича на просторі Хеммінга; доведення дивергентного формулювання задачі оптимального транспортування.

6. Орієнтовний перелік графічного (ілюстративного) матеріалу: презентація доповіді

8. Дата видачі завдання: 10 вересня 2022 р.

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання	Примітка
1	Узгодження теми роботи із науковим керівником	01-15 вересня 2022 р.	Виконано
2	Огляд опублікованих джерел за тематикою дослідження	16-30 вересеня-жовтень 2022 р.	Виконано
3	Розв'язання завдання оптимального транспортування на дво-точковому просторі	Листопад-грудень 2022 р.	Виконано
4	Аналіз дуальності Канторовича на просторі Хеммінга	Січень-Березень 2023 р.	Виконано
5	Доведення дивергентного формулювання задачі оптимального транспортування	Квітень-тарвень 2023 р.	Виконано
6	Оформлення та захист дипломної роботи	Червень 2023 р.	Виконано

Студент ____________ Максим ОРЖАХІВСЬКИЙ

Керівник ____________ Георгій РЯБОВ

РЕФЕРАТ

Кваліфікаційна робота містить: 35 стор., 3 джерела.

У цій роботі було досліджено тему оптимального транспортування, а саме задача Канторовича. Було доведено дивергентне формулювання оптимального транспортування ймовірнісних розподілів на просторі Хеммінга.

В ході дослідження була перевірена коректність дивергентного формулювання задачі оптимального транспортування на двоточковому просторі, продемонстрована необхідність використання цього формулювання в наступному по складності випадку. Також було досліджене практичне застосування дуальності Канторовича, яка разом з потужними інструментами функціонального аналізу дали змогу довести поставлене дивергентне формулювання.

ОПТИМАЛЬНЕ ТРАНСПОРТУВАННЯ, ПРОСТІР ХЕММІНГА, ДИВЕРГЕНТНЕ ФОРМУЛЮВАННЯ, ДУАЛЬНІСТЬ КАНТОРОВИЧА

ABSTRACT

The qualification work contains: 35 pages, 3 references.

In this work the topic of optimal transportation was investigated, namely the Kantorovich problem. The divergent formulation of the optimal transportation of optimal transportation of probability distributions on the Hamming space Hamming space.

The study verified the correctness of the divergent formulation of the optimal transportation problem on a two-point space and demonstrated the need to use this formulation in the next most difficult case. The practical application of the Kantorovich duality was also investigated, which, together with powerful tools of functional analysis, made it possible to prove the divergent formulation.

OPTIMAL TRANSPORTATION, HAMMING SPACE, DIVERGENT FORMULATION, KANTOROVICH DUALITY

ЗМІСТ

ВСТУП

Актуальність дослідження. Актуальність даного дослідження полягає у тому, що на сьогоднішній день чисельні схеми на дискретних просторах викликають багато уваги та цікавості. Тому дослідження оптимального транспортування на дискретних просторах є актуальним.

Проблематикою мого дослідження є складність обчислювання задач оптимального перерозподілу ймовірнісних розподілів на дискретних просторах. Тому розробка математичних інструментів в цій темі має значення.

Нові дивергентні формулювання дають змогу наближено обчислювати складні задачі, які без такого інструменту практично неможливо обчислити, або дуже складно. Тому в цьому дослідженні буде доведене нове дивергентне формулювання для вирішення проблеми обчислення оптимального транспортування ймовірнісних розподілів на дискретних просторах.

Метою дослідження є доведення нового дивергентного формулювання в оптимальному транспортуванні. Для досягнення мети необхідно вирішити **задачу дослідження**, яка полягає у доведенні нового дивергентного формулювання. Для розв'язання задачі необхідно вирішити такі завдання:

1) провести огляд опублікованих джерел за тематикою дослідження;

2) розв'язати завдання оптимального транспортування на двоточковому просторі;

3) аналіз дуальності Канторовича на просторі Хеммінга;

4) доведення дивергентного формулювання задачі оптимального транспортування.

Об'єктом дослідження є оптимальне транспортування ймовірнісних розподілів на дискретних просторах.

Предметом дослідження є простір Хеммінга, оператор дискретної похідної, ймовірнісний розподіл.

При розв'язанні поставлених завдань використовувались такі *методи дослідження*: функціональний аналіз, теорія оптимального транспортування.

Наукова новизна отриманих результатів полягає у тому, що я доводжу нове дивергентне формулювання для оптимального транспортування.

Практичне значення результатів полягає у тому, що дивергентні формулювання є корисним апаратом, за допомогою якого, можна будувати числові схеми та знаходити оцінку вартості оптимального транспортування ймовірнісних розподілів на просторі Хеммінга.

1 ОГЛЯДОВА ЧАСТИНА

В даному розділи будуть викладені теоретичні відомості та означення оптимального транспортування, необхідні для доведення дивергентного формулювання.

1.1 Задача оптимального транспортування

Питання економії ресурсів є актуальним питанням з давніх часів, будь-яка «робота» вимагала витрачання певних ресурсів, чим менше ресурсів необхідно було витратити – тим вигіднішою була ця «робота». В деяких випадках оптимальний план роботи (тобто той, що є найменш витратним відносно того, чи іншого ресурсу) можна знайти дуже легко, а в деяких випадках це практично неможливо.

Задача оптимального транспортування виникла на підставі того, що при транспортуванні певної маси з точки A в точку B теж витрачаються ресурси і реалізувати транспортування можна різними способами, відповідно кожен спосіб має різну кінцеву вартість, яка вимірюється в тому чи іншому ресурсі, а можливо в сукупності витрачених ресурсів.

Оптимальне транспортування є математичною концепцією, яка вирішує проблему ефективного переміщення об'єктів або ресурсів з одного місця в інше. Розвиток теорії оптимального транспортування має багату історію та знайшло практичне застосування в багатьох галузях.

Походження оптимального транспортування можна прослідкувати від робіт французького математика Гаспара Монжа наприкінці 18-го століття. Монжем була запропонована ідея знаходження оптимального плану переміщення ґрунту з однієї ділянки на іншу для зміни ландшафту. Формулювання задачі оптимального транспортування Монжа полягає у знаходженні бієкції (відображення "один до одного") між множиною джерел та множиною пунктів призначення, яка мінімізує загальну вартість перерозподілу. Однак задача Монжа накладає додаткові обмеження на відображення, такі як збереження порядку точок і гарантування того, що кожній точці відправлення відповідав точно один пункт призначення.

Основна складність задачі Монжа полягає в тому, що пошук оптимального транспортування є обчислювально дорогим і складним. Зокрема, задача не піддається простим алгоритмам, а область пошуку може бути надзвичайно великою, особливо коли мова йдеться про об'ємні за розміром дані. Тому формулювання задачі потребувало спрощення, де на допомогу і прийшов Канторович, запропонувавши альтернативне формулювання задачі. Він послаблює деякі обмеження задачі Монжа і дозволяє перерозподіляти ресурси декількома шляхами, тобто ресурси з точки x могли мати декілька пунктів призначення в різні точки y. Задача Канторовича може бути розв'язана за допомогою методів лінійного програмування, що робить її легшою в плані обчислень порівняно з задачею Монжа. Надалі будемо розглядати саме задачу Канторовича.

1.2 Оптимальне транспортування у математиці

Поняття оптимального транспортування можна легко відобразити в математику. Математичне формулювання задачі оптимального транспортування полягає у визначенні найефективнішого, «найвигіднішого» способу переміщення певних ресурсів або певної маси з одного набору місць для цих самих мас, в інший, враховуючи різні обмеження та власне витрати. Основною ідеєю є пошук відображення, або відображень між двома наборами місць, яке мінімізує витрати.

1.3 Функція вартості

Для кращого та наочного розуміння теорії оптимального транспортування будемо наводити аналогію з задачею, де треба перерозподілити землю з однієї ділянки на іншу. Тоді нормалізуємо кількість землі на першій ділянці до одиниці.

Будемо моделювати і землю, і ділянку, яку необхідно заповнити, за допомогою ймовірнісних мір μ та ν, визначених відповідно на деяких вимірних просторах X та Y. Якщо A та B є вимірними підмножинами X та Y відповідно, то $\mu[A]$ дає міру того, скільки землі знаходиться всередині підділянки A; а $\nu[B]$ – скільки землі можна буде засипати в B.

Очевидно, що транспортування землі затратна операція по багатьох критеріях. Введемо вимірну функцію c на $X \times Y$, яка буде показувати витрати. Нехай потрібно перевезти масу землі з точки x у точку y, тоді функція $c(x,y)$ буде власне визначати кількість витрат на дане перевезення. Відповідно c – вимірна і невід'ємна величина, так як витрати не можуть бути менші за 0.

Означення 1.1. *Функція вартості* — це функція c (від слова «cost»), визначена на множині $X \times Y$, яка приймає два аргументи на вхід, один з множини X, а другий з Y, візьмемо x та y відповідно, а результатом функції є невід'ємне значення, яке показує «вартість» транспортування маси з точки x в точку y.

$$c(x,y) = v, v \in \mathbb{R}$$

1.4 План транспортування

Також нам слід розглянути поняття плану транспортування. По своїй суті план транспортування — це набір, послідовність дій (правил) транспортування, тобто дотримуючись певного плану транспортування ми виконаємо транспортування за цим планом.

Такий план має різні характеристики, наприклад ми можемо дізнатися скільки маси було перевезено при перевезенні з однієї локації в іншу.

Нам слід змоделювати ймовірнісну міру плану транспортування π на добуток просторів X та Y. Нехай $x \in X$, $y \in Y$, $d\pi(x,y)$ буде показувати кількість маси перенесеної з локації x в y, також ми не виключаємо можливості того, що деяка маса, яка знаходиться в x може бути розподілена між різними локаціями y.

Важливою умовою є те, що вся маса, яка транспортується з x та вся маса, яка була транспортована в y – збігається з $d\mu(x)$ та $d\nu(y)$ відповідно, для всіх $\pi \in P(X \times Y)$.

Означення 1.2. *План транспортування* — ймовірнісна міра $\pi \in P$ визначена на $X \times Y$, яка задовільняє наступній умові:

$$\pi[A \times Y] = \mu[A]$$

$$\pi[X \times B] = \nu[B],$$

де A та B підмножини X та Y відповідно, а μ та ν визначені в 1.3 ймовірнісні міри.

Множина всіх π, які задовільняють цю умову позначимо як Π.

$$\Pi(\mu,\nu) = \{\pi \in P(X\times Y), A \subseteq X, B \subseteq Y : \pi[A\times Y] = \mu[A], \pi[X\times B] = \nu[B]\}$$

1.5 Задача Канторовича

Канторович сформулював задачу оптимального транспортування, яка полягає в знаходженні такого плану транспортування π, для якого значення $I[\pi]$ буде мінімальним, іншими словами мінімізувати:

$$I[\pi] = \int_{X\times Y} c(x,y) d\pi(x,y),$$

по всім $\pi \in \Pi(\mu,\nu)$.

Тоді, для заданого плану транспортування π невід'ємну величину $I[\pi]$ будемо називати кінцева вартість транспортування, тобто $I[\pi]$ показує загальну вартість транспортування з планом π. Тоді сформулюємо означення оптимального плану транспортування між μ та ν.

Означення 1.3. *Оптимальний план транспортування*

$$T_c = \inf I[\pi], \forall \pi \in \Pi(\mu,\nu)$$

Тоді $\forall \pi : I[\pi] = T_c(\mu,\nu)$, якщо такі існують, будуть називатися *планами оптимального транспортування.*

1.6 Дуальність Канторовича[3]

В цьому підрозділі буде розглянуто дуже важливу теорему Канторовича-Рубінштейна, яка буде використовуватися в подальшому.

Кантрович розглянув частковий випадок, коли функція вартості 1.1 є функцією відстані, тобто коли

$$c(x,y) = d(x,y).$$

Але насправді його теорема є досить загальною і функція відстані теж може приймати різні види.

Теорема 1.1 (Дуальність Канторовича)**.** *Нехай X, Y – простори Поліша. $\mu \in P(X)$, $\nu \in P(Y)$. Нехай задана функція $c : X \times Y \to \mathbb{R}_+ \cup \{+\infty\}$ – функція вартості.*

Якщо $\pi \in P(x,y)$ та $(\varphi,\psi) \in L^1(d\mu) \times L^1(d\nu)$, тоді визначаємо

$$I[\pi] = \int_{X \times Y} c(x,y) d\pi(x,y),$$

$$J(\varphi,\psi) = \int_X \varphi d\mu + \int_Y \psi d\nu.$$

Визначимо Φ_c *як множину всіх вимірних функцій* $(\varphi,\psi) \in L^1(d\mu) \times L^1(d\nu)$, *які задовільняють наступну умову:*

$$\varphi(x) + \psi(y) \leq c(x,y),$$

тоді

$$\inf I[\pi] = \sup J(\varphi,\psi),$$

по всім $\pi \in \Pi(\mu,\nu)$ *та* $(\varphi,\psi) \in \Phi_c$.

Ця теорема є дуже потужним та корисним інструментом, який буде використано надалі для доведення дивергентного формулювання.

Такого роду дуальність формулювання задачі оптимального транспортування дає змогу зробити зручний перехід від пошуку інфемуму з планами транспортування — до пошуку супремуму з функція, заданими з особливою умовою.

1.7 Дивергентні формулювання

Задач оптимального транспортування досить багато, у всіх них є свої особливості, різні функції вартості та умови. Деякі задачі є практично тривіальними, деякі складними для обрахунків, а в деяких задачах практично неможливо знайти відповідь. В таких задачах на допомогу і приходять дивергентні формулювання. Приклади дивергентних формулювань в оптимальному транспортуванні розглянуті тут [3], та ось тут [2].

Розбіжна постановка задачі оптимального транспортування виникає тоді, коли функція витрат, що використовується, не є строго опуклою. У цьому випадку задача знаходження оптимального транспортування може мати декілька розв’язків або навіть не мати розв’язку взагалі. Для визначення функції вартості в таких задачах зазвичай використовують дивергенцію - функцію, яка кількісно виражає різницю між розподілами ймовірностей.

Розрахунки дивергенції можуть забезпечити додаткову гнучкість і дозволити більш детальне моделювання задач транспортування. Однак вони також створюють проблеми з точки зору математичних властивостей оптимізаційної задачі, таких як неопуклість, неунікальність розв'язку та обчислювальна складність. Як наслідок, для розв'язання дивергентних задач оптимального транспортування часто потрібні спеціалізовані алгоритми та чисельні методи.

Загалом, дивергентні формулювання в оптимальному транспортуванні забезпечують багатий математичний апарат для вирішення задач, і знайшли застосування в широкому спектрі дисциплін.

1.8 Постановка задачі

Дивергентні формулювання викликали особливий інтерес, в цьому дослідженні буде доведене одне з запропонованих формулювань. Воно має наступний вигляд:

Розглядається добуток двоточкових просторів $X = \{0,1\}^n$ з метрикою Хеммінга

$$d(x,y) = \sum_{k=1}^{n} I_{x_k \neq y_k}$$

(тобто d – кількість різних компонент в x та y).

На просторі X задано два розподіли: $(p_x)_{x \in X}$ та $(q_y)_{y \in X}$, тобто

$$p_x \geq 0, \sum_{x \in X} p_x = 1,$$

$$q_y \geq 0, \sum_{y \in X} q_y = 1.$$

Планом транспортування буде розподіл $(\pi_{xy})_{(x,y) \in X \times X}$ з маргінальними розподілами $(p_x)_{x\in X}$ та $(q_y)_{y\in Y}$, тобто

$$\forall (x,y) \in X \times X \, \pi_{xy} \geq 0, \sum_{(x,y)\in X\times X} \pi_{xy} = 1,$$

$$\forall x \in X \sum_{y\in X} \pi_{xy} = p_x, \forall y \in X \sum_{x\in X} \pi_{xy} = q_y.$$

Задача оптимального транспортування полягає в мінімізації функції вартості

$$I[\pi] = \sum_{(x,y)\in X\times X} \pi_{xy} d(x,y)$$

по всіх $(\pi_{xy})_{(x,y)\in X\times X}$ за вищевказаних умов.

Позначимо мінімальне значення через I_*.

Нехай перетворення $\epsilon_k : X \to X$ змінює k-у компоненту x. Наприклад,

$$\epsilon_1(0,1,1,0,0,0) = (1,1,1,0,0,0), \epsilon_3(1,0,1,0) = (1,0,0,0).$$

Для функції $u : X \to \mathbb{R}$ розглянемо різницеві оператори

$$D_k u(x) = u(\epsilon_k x) - u(x).$$

Нехай $u_1,...,u_n : X \to \mathbb{R}$ задовільняють рівняння

$$\sum_{k=1}^{n} D_k u_k(x) = (p_x - q_x). \tag{1.1}$$

Тоді рівність яку нам потрібно довести виглядає наступним чином

$$\inf \left(\sum_{x\in X} |u_1(x)| + ... + |u_n(x)| \right) = I_*. \tag{1.2}$$

Висновки до розділу 1

У цьому розділі було розглянуто теорію оптимального транспортування, було проаналізовано роботу Канторовича, перевага задачі, яку він сформулював над задачею Монжа, була розглянута важлива та корисна теорема, а саме теорема про дуальність формулювання задачі оптимального транспортування Канторовича. Всі ці інструменти необхідні для розуміння та доведення дивергентного формулювання для оптимального транспортування ймовірнісних розподілів на дискретних просторах, а саме на просторі Хеммінга, яке дасть змогу наближено обчислювати оптимальну вартість, і яке буде доведено в наступному розділі.

2 ПРАКТИЧНА ЧАСТИНА

У другому розділі буде розглянуто часткові випадки та доведення дивергентного формулювання 1.8. А саме розглянемо випадок коли $X = \{0{,}1\}$, тобто для $n = 1$, та випадок $n = 2$, які допоможуть зрозуміти важливість дивергентного формулювання в цілому та зручність використання теореми 1.1, а саме дуальність Канторовича та перехід від інфімуму з планами транспортування, до супремуму з 1-Ліпшицевими функціями.

2.1 Тривіальний випадок

У цьому підрозділі буде розглянуто тривіальний випадок у задачі дивергентного формулювання, який допоможе побачити різницю в складності при збільшенні n.

Розглянемо випадок, коли $n = 1$:

Маємо двоточковий простір $X = \{0{,}1\}$ з метрикою Хеммінга

$$d(0{,}0) = d(1{,}1) = 0,$$

$$d(0{,}1) = d(1{,}0).$$

На просторі X задано два розподіли μ та ν

$$\mu(1) = p, \mu(0) = 1 - p$$

$$\nu(1) = q, \nu(0) = 1 - q.$$

Розподіл π заданий на добутку $X \times X$, тобто зводиться до чотирьох значень π_{00}, π_{01}, π_{10}, π_{11}, таких що:

$$\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11} \geq 0, \pi_{00} + \pi_{01} + \pi_{10} + \pi_{11} = 1. \quad (2.1)$$

Крім того, проєкції π на першу та другу компоненти рівні μ та ν відповідно

$$\pi_{00} + \pi_{01} = 1 - p, \pi_{10} + \pi_{11} = p, \pi_{00} + \pi_{10} = 1 - q, \pi_{01} + \pi_{11} = q. \tag{2.2}$$

Тоді задача полягає у тому, щоб мінімізувати функцію вартості:

$$I[\pi] = \sum_{\pi_x y, x, y \in X} \pi_{xy} d(x,y),$$

по всім π_{00},π_{01},π_{10},π_{11} за умов 2.1 та 2.2.

Розглянемо детальніше $I[\pi]$, внаслідок простоти випадку ми можемо легко розписати суму з якою працюємо:

$$\sum_{\pi_x y, x, y \in X} \pi_{xy} d(x,y) = \pi_{00} d(0,0) + \pi_{01} d(0,1) + \pi_{10} d(1,0) + \pi_{11} d(1,1).$$

Маємо чотири доданки, але якщо уважніше поглянемо на деякі з них, а саме ті, які містять множники $d(0,0)$ та $d(1,1)$, то можна побачити, що використовуючи дану метрику вони будуть нульовими, тоді сума зведеться до двох доданків:

$$\sum_{\pi_x y, x, y \in X} \pi_{xy} d(x,y) = \pi_{01} d(0,1) + \pi_{10} d(1,0).$$

Тепер можна замінити решту метрик на їх числове значення, а саме $d(0,1) = d(1,0) = 1$, тоді сума матиме наступний вигляд:

$$\sum_{\pi_x y, x, y \in X} \pi_{xy} d(x,y) = \pi_{01} + \pi_{10},$$

отже завдання було перетворене на наступне: мінімізувати

$$I[\pi] = \pi_{01} + \pi_{10}. \tag{2.3}$$

Легко побачити, що задача звелась до досить простої. Тепер необхідно виразити один доданок через інший. Для цього розглянемо умову 2.2:

$$\pi_{00} + \pi_{01} = 1 - p,$$

$$\pi_{10} + \pi_{11} = p,$$

$$\pi_{00} + \pi_{10} = 1 - q,$$

$$\pi_{01} + \pi_{11} = q,$$

звернемо увагу на друге та четверте рівняння:

$$\pi_{10} + \pi_{11} = p,$$

$$\pi_{01} + \pi_{11} = q,$$

знайдемо різницю цих рівнянь:

$$\pi_{10} + \pi_{11} - \pi_{01} - \pi_{11} = p - q,$$

π_{11} скорочується і залишається залежність π_{10} та π_{01}, тоді слід виразити одне через інше, та отримаємо:

$$\pi_{10} - \pi_{01} = p - q,$$

$$\pi_{10} = \pi_{01} + p - q,$$

вважаємо, що $p \geq q$.

Підставимо виражене π_{10} в 2.3, отримаємо:

$$I[\pi] = \pi_{01} + \pi_{10},$$

$$I[\pi] = \pi_{01} + p - q + \pi_{01},$$

$$I[\pi] = 2\pi_{01} + p - q,$$

пам'ятаємо умову 2.1, та те, що $p \geq q$, маємо що обидва доданки $2\pi_{01}$ та

$(p-q) \geq 0$, тоді для мінімізації слід взяти $\pi_{01} = 0$, отримали:

$$I[\pi] = p - q,$$

при чому послідовно підставляючи в усі вище отримані рівняння 2.2 $\pi_{01} = 0$ маємо:

$$\pi_{10} = \pi_{01} + p - q,$$

$$\pi_{10} = p - q,$$

$$\pi_{01} + \pi_{11} = q,$$

$$\pi_{11} = q,$$

$$\pi_{00} + \pi_{01} = 1 - p,$$

$$\pi_{00} = 1 - p.$$

Повернемося до моменту де вважали, що $p \geq q$, тепер розглянемо протилежний випадок, суть не зміниться, але тоді ми будемо брати не різницю другого та четвертого, а четвертого та другого, тоді отримаємо:

$$\pi_{01} + \pi_{11} - \pi_{10} - \pi_{11} = p - q,$$

$$\pi_{01} = \pi_{10} + q - p,$$

підставляючи в 2.3 маємо:

$$I[\pi] = 2\pi_{10} + q - p,$$

приймаємо на цей раз $\pi_{10} = 0$ та маємо:

$$minI[\pi] = q - p,$$

$$\pi_{00} = 1 - q,$$

$$\pi_{01} = q - p,$$

$$\pi_{10} = 0,$$

$$\pi_{11} = p$$

.

Отже, можна прийти до висновку, що в такому випадку мінімальне значення 2.3, або ж $I_* = |p - q|$.

Тепер слід перевірити чи збігається знайдене мінімальне значення з тим, яке подається у дивергентному формулюванні для випадку $n = 1$.

2.2 Перевірка для тривіального випадку

В цьому підрозділі буде перевірено, чи виконується сформоване дивергентне формулювання для випадку $n = 1$.

Розглянемо праву частину сформульованої рівності, для неї розглянемо функцію $u : X \to \mathbb{R}$, маємо різницевий оператор D

$$Du(1) = u(0) - u(1),$$

$$Du(0) = u(1) - u(0),$$

як можемо побачити, це той самий оператор з 1.8,

$$D_k u(x) = u(\epsilon_k x) - u(x),$$

де просто підставили точки $\{1\}$ та $\{0\}$.

Якщо на просторі X задано рівномірний розподіл, то щільність μ відносно нього рівна:

$$\rho_\mu(0) = 1 - p,$$

$$\rho_\mu(1) = p,$$

а щільність ν відносно нього рівна:

$$\rho_\nu(0) = 1 - q,$$

$$\rho_\nu(1) = q.$$

Якщо прирівняти Du до різниці щільностей:

$$u(0)-u(1) = \rho_\mu(1)-\rho_\nu(1) = p-q, u(1)-u(0) = \rho_\mu(0)-\rho_\nu(0) = q-p. \quad (2.4)$$

Тепер, підставляючи $n = 1$ в праву частину (1.2) отримуємо:

$$\inf\left(\sum_{x\in X} |u_1(x)| + ... + |u_n(x)|\right) = \inf\Big(|u(0)| + |u(1)|\Big). \quad (2.5)$$

Щоб перевірити, чи дійсно виконується рівність (1.2) необхідно обчислити (2.5) (якщо це можливо).

Для початку спробуємо виразити $u(0)$ через $u(1)$, з першого рівняння умови (2.4) маємо:

$$u(0) - u(1) = (p - q),$$

виражаємо $u(0)$:

$$u(0) = (p - q) + u(1).$$

Нехай $p \geq q$, тоді різниця $(p - q) \geq 0$, розглянемо $u(1)$:
Якщо $u(1) \geq 0$, тоді очевидно, що $u(0) \geq 0$, отже (2.5), яке необхідно мінімізувати набуде вигляду:

$$u(0) + u(1),$$

ми позбулися модулів, через те, що кожен доданок ≥ 0, підставимо замість $u(0)$ його вираження через $u(1)$:

$$u(0) + u(1) = ((p - q) + u(1)) + u(1) =$$

$$= (p - q) + u(1)) + u(1) = (p - q) + 2u(1);$$

Враховуючи поставлені раніше умови, обидва доданки не від'ємні, а отже для отримання мінімального значення візьмемо $u(1) = 0$, тоді мінімальне значення буде дорівнювати $p-q$, при:

$$u(0) = (p-q)$$

$$u(1) = 0.$$

Розглянемо інший випадок, коли $-(p-q) \leq u(1) < 0$, тоді $u(0) \geq 0$, отримуємо, що $u(1)$ – від'ємне, $u(0)$ – не від'ємне, отже розкриваючи модулі маємо:

$$u(0) - u(1),$$

підставимо вираження $u(0)$ через $u(1)$:

$$u(0) - u(1) = (2(p-q) + u(1)) - u(1) = p - q,$$

отримали той самий результат, але вже для будь-якого $u(1) \in [-(p-q), 0)$, та $u(0) = (p-q) + u(1)$.

Тепер слід перевірити випадок коли $u(1) < -(p-q)$, тоді відповідно $u(0) < 0$, отже обидва модулі будуть розкриті зі знаком «$-$»:

$$-u(0) - u(1) = -((p-q) + u(1)) - u(1) =$$

$$= -(p-q) - 2u(1),$$

пам'ятаємо, що $u(1) < -(p-q)$, а отже вираз $-(p-q) - 2u(1) > (p-q)$, це значення більше, ніж те, що було знайдено раніше, а отже такий варіант не підходить.

Повернемось до порівняння p та q, ми розглянули один варіант, тепер слід розглянути другий, коли $p \leq q$. У цьому випадку будемо використовувати друге рівняння з умови (2.4) та виражати $u(1)$ через $u(0)$, а не навпаки. Маємо:

$$u(1) - u(0) = (q-p),$$

$$u(1) = u(0) + (q-p),$$

насправді можна використовувати і перше рівняння з (2.4), ми отримаємо те ж саме. Далі розглядаємо варіант, коли $u(0) \geq 0$, тоді $u(1) \geq 0$, через те, що $q - p \geq 0$, обидва модулі відкриваються зі знаком «+»:

$$u(0) + u(1) = u(0) + (u(0) + (q - p)) = 2u(0) + (q - p).$$

Пам'ятаємо, що $(q - p) \geq 0$, отже для мінімізації необхідно прийняти $u(0) = 0$, тоді мінімальне значення $(q - p)$, при

$$u(0) = 0, u(1) = (p - q).$$

Як і в випадку, коли $p \geq q$ розглянемо наступний варіант, а саме коли $-(q-p) \leq u(0) \leq 0$, тоді $u(1) \geq 0$ і модулі будуть відкриті наступним чином:

$$-u(0) + u(1) = -u(0) + (u(0) + (q - p)) = (q - p),$$

отже тут все аналогічно до минулого випадку, лише

$$-(q - p) \leq u(0) \leq 0, u(1) = u(0) + (q - p).$$

І останній аналогічний випадок, коли $u(0) < -(q - p)$, $u(1) < 0$, а отже обидва модулі будуть відкриті зі знаками «$-$»:

$$-u(0) - u(1) = -u(0) - (u(0) + 2(q - p)) = -2u(0) - (q - p),$$

а так як $u(0) < -(q - p)$, то і весь вираз буде

$$-2u(0) - (q - p) > (q - p).$$

Отже, для випадку $n = 1$ ліва частина (1.2), а точніше її мінімальне значення буде дорівнювати $|p - q|$. Як бачимо, результати співпали і для даного випадку дивергентне формулювання виконується і це досить легко перевірити. Але демонстрація одного часткового випадку не є гарантом того, що рівність буде виконуватися у загальному випадку. Наступним кроком буде дослідження поведінки цієї рівності для $n = 2$.

2.3 Складний частковий випадок

В підрозділі 2.2 був досліджений легкий частковий випадок, коли $n = 1$, було показано, що рівність дійсно виконується. Тепер слід дослідити поведінку цієї задачі для складнішого випадку, коли n буде дорівнювати хоча-б 2.

Дослідимо у що перетворюється рівність, яку нам потрібно довести. Розпочнемо з планів транспортування, маємо: двоточковий простір $X = \{0,1\}^2$, та ж сама метрика Хеммінга

$$d(x,y) = \sum_{k=1}^{2} I_{x_k \neq y_k},$$

$$I[\pi] = \sum_{(x,y) \in X \times X} \pi_{xy} d(x,y)$$

по всіх $(\pi_{xy})_{(x,y) \in X \times X}$, розпишемо цю суму для нашого випадку:

$$\begin{aligned} I[\pi] = {} & \pi_{00,00} d(00,00) + \pi_{00,01} d(00,01) + \pi_{00,10} d(00,10) + \pi_{00,11} d(00,11) + \\ & + \pi_{01,00} d(01,00) + \pi_{01,01} d(01,01) + \pi_{01,10} d(01,10) + \pi_{01,11} d(01,11) + \\ & + \pi_{10,00} d(10,00) + \pi_{10,01} d(10,01) + \pi_{10,10} d(10,10) + \pi_{10,11} d(10,11) + \\ & + \pi_{11,00} d(11,00) + \pi_{11,01} d(11,01) + \pi_{11,10} d(11,10) + \pi_{11,11} d(11,11) \end{aligned} \tag{2.6}$$

Як і у минулому випадку, деякі доданки можна скоротити, а саме це чотири доданки, де $x = y, d(x,y) = 0$. Залишається все ще велика кількість доданків, і навіть використовуючи умови планів транспортування, які можна тут записати, виразити все вже не вийде. Бачимо, що вже для, здавалось невеликого n, задача набуває значних обертів у складності обрахунку, і використання планів транспортування вже не є таким зручним і простим. Тому на допомогу і приходить дивергентне формулювання, яке дасть змогу знаходити наближені числові значення для будь-якого n.

2.4 Використання дуальності Канторовича у загальному випадку задачі

Як ми побачили в (2.6) використання планів транспортування не завжди легко дає результат, спробуємо використати теорему Канторовича про дуальність формулювання, тоді будемо працювати з функціями 1-Ліпшиця і доведемо наше дивергентне формулювання (1.2).

Згадуючи дуальність Канторовича (1.1), маємо:

$$\inf I[\pi] = \sup \sum_{x \in X} f(x)(p_x - q_x), \tag{2.7}$$

при чому $f(x)$ – 1-Ліпшицева функція, а це означає, що:

$$|f(x) - f(y)| \leq d(x,y),$$

тобто модуль різниці функції f від x та y дорівнює власне відстані між x та y.

Розглянемо тепер умову (1.1), а саме:

$$\sum_{k=1}^{n} D_k u_k(x) = (p_x - q_x).$$

Було б добре виразити звідси $(p_x - q_x)$, яке ми використовуємо в (2.7), тоді:

$$(p_x - q_x) = \sum_{k=1}^{n} D_k u_k(x). \tag{2.8}$$

Підставляючи вираз з (2.8) в (2.7) маємо:

$$\inf I[\pi] = \sup \sum_{x \in X} f(x)(p_x - q_x) = \sup \sum_{x \in X} f(x) \sum_{k=1}^{n} D_k u_k(x) =$$

$$= \sup \sum_{x \in X} \sum_{k=1}^{n} f(x) D_k u_k(x).$$

Розглянемо тепер $\sum_{k=1}^{n} D_k u_k(x)$:

$$D_k u_k(x) = u_k(\epsilon_k x) - u_k(x),$$

$$\sum_{k=1}^{n} D_k u_k(x) = D_1 u_1(x) + D_2 u_2(x) + ... + D_n u_n(x),$$

розкриваючи D_k :

$$\sum_{k=1}^{n} D_k u_k(x) = u_1(\epsilon_1 x) - u_1(x) + u_2(\epsilon_2 x) - u_2(x) + ... + u_n(\epsilon_n x) - u_n(x),$$

розіб'ємо цей вираз на різницю двох сум:

$$\sum_{k=1}^{n} D_k u_k(x) = \big(u_1(\epsilon_1 x) + u_2(\epsilon_2 x) + ... + u_n(\epsilon_n x)\big) - \big(u_1(x) + u_2(x) + ... + u_n(x)\big),$$

$$\sum_{k=1}^{n} D_k u_k(x) = \sum_{k=1}^{n} u_k(\epsilon_k x) - \sum_{k=1}^{n} u_k(x),$$

підставляючи в суму замість $\sum_{k=1}^{n} D_k u_k(x)$ маємо:

$$\sup \sum_{x \in X} f(x) \Big(\sum_{k=1}^{n} u_k(\epsilon_k x) - \sum_{k=1}^{n} u_k(x) \Big),$$

$$\sup \sum_{x \in X} f(x) \Big(\sum_{k=1}^{n} u_k(\epsilon_k x) \Big) - f(x) \Big(\sum_{k=1}^{n} u_k(x) \Big).$$

Розглянемо $\epsilon_k : \forall x \in X : \big(\epsilon_k(\epsilon_k x)\big) = x$, тоді для деякого $u_k(x)$ завжди знайдеться така $f(\epsilon_k x)$, що буде існувати різниця:

$$f(\epsilon_k x) u_k(x) - f(x) u_k\big(\epsilon_k(\epsilon_k x)\big) = u_k(x)\big(f(\epsilon_k x) - f(x)\big),$$

як бачимо, в початковій сумі ми можемо перекинути D_k з $u_k(x)$ на $f(x)$.

$$\sup \sum_{x\in X}\sum_{k=1}^{n} D_k f(x) u_k(x),$$

тепер, розглядаючи $D_k f(x)$ можна скористатися тим, що $f(x)$ – 1-Ліпшицева функція, а отже:

$$|f(x) - f(y)| \leq d(x,y),$$

$$|f(\epsilon_k x) - f(x)| \leq 1,$$

так як між $(\epsilon_k x)$ та x різниця лише в одну компоненту.

Заміняючи $D_k f(x)$ на найкращий випадок, тобто в залежності від знака на 1 або -1 отримаємо наступне:

$$\sup \sum_{x\in X}\sum_{k=1}^{n} D_k f(x) u_k(x) \leq \sum_{x\in X}\sum_{k=1}^{n} |u_k(x)|,$$

звідки й отримуємо:

$$\inf I[\pi] \leq \sum_{x\in X}\sum_{k=1}^{n} |u_k(x)|.$$

У цьому підрозділі ми отримали нерівність для нашого дивергентного формулювання (1.2), тепер для того, щоб довести рівність, необхідно показати що нерівність буде виконуватися і в іншу сторону.

2.5 Доведення рівності у дивергентному формулюванні

У підрозділи 2.4 було доведено нерівність в одну сторону, тепер необхідно довести її в іншу, що в кінцевому результаті отримати бажану рівність.

Згадаємо, що собою являє відстань між розподілами у нашому випадку:

$$W_1(p,q) = \sup \sum_{x \in X} f(x)(p_x - q_x).$$

Повернемось до виразу $D_k f(x)$. Нехай маємо векторний простір $(D_k f(x))_{1 \leq k \leq n, x \in X} \in \mathbb{R}^{n2^n}$. Користуючись методами, опублікованими в статті [2], розглянемо такий лінійний функціонал L на цьому просторі, що:

$$L(Df) = \sum_{x \in X} f(x)(p_x - q_x). \tag{2.9}$$

Щоб переконатися, чи лінійний функціонал є коректно визначеним, потрібно перевірити дві умови: лінійність та однозначність.

Лінійність:

$$L(Df + Dg) = L(D(f + g)) = \sum_{x \in X} (f(x) + g(x))(p_x - q_x) =$$

$$= \sum_{x \in X} f(x)(p_x - q_x) + \sum_{x \in X} g(x)(p_x - q_x) = L(Df) + L(Dg),$$

$$L(\alpha Df) = L(D\alpha f) = \sum_{x \in X} \alpha f(x)(p_x - q_x) =$$

$$= \alpha \sum_{x \in X} f(x)(p_x - q_x) = \alpha L(Df),$$

отже лінійність виконується.

Лема 2.1. *Нехай $Df(x) = Dg(x)$, тоді з означення D маємо:*

$$f(\epsilon_k x) - f(x) = g(\epsilon_k x) - g(x),$$

для всіх $x \in X$ та $1 \leq k \leq n$,

$$\forall x \in X, k \in [1,n] : f(\epsilon_k x) = g(\epsilon_k x) + c, c = const$$

Доведення. Розглянемо, що відбувається для $x = (0,...,0,...,0)$, тобто x – вектор нулів, підставимо:

$$f(0,...,1,...,0) = g(0,...,1,...,0) + (f(0,...,0,...,0) - g(0,...,0,...,0)),$$

останній доданок $(f(0,...,0,...,0) - g(0,...,0,...,0)) = c$, де c – $const$. Отже:

$$f(0,...,1,...,0) = g(0,...,1,...,0) + c.$$

Тепер розглянемо що відбудеться, якщо почати додавати кількість одиниць у вектор x, нехай x – вектор нулів з однією одиницею:

$$f(0,...,1,...,1,...,0) = g(0,...,1,...,1,...,0) + (f(0,...,1,...,0) - g(0,...,1,...,0)),$$

але ми перед цим дізналися, що $f(0,...,1,...,0) = g(0,...,1,...,0) + const$, а отже $f(0,...,1,...,0) - g(0,...,1,...,0) = c$, підставляючи отримаємо:

$$f(0,...,1,...,1,...,0) = g(0,...,1,...,1,...,0) + c,$$

звідки зрозуміло що $\forall x \in X, k \in [1,n] : f(\epsilon_k x) = g(\epsilon_k x) + c, c = const.$ □

Однозначність:

З леми 2.1 випливає, що якщо $Df(x) = Dg(x)$, то $L(Df) = L(Dg)$.

Було показано, що функціонал $L(Df)$ дійсно є коректно визначеним лінійним функціоналом. Отже, розглянемо його норму:

$$\|L\| = \sup_{|D_k f(x)| \le 1} \sum_{x \in X} f(x)(p_x - q_x),$$

як бачимо, норма лінійного функціоналу співпала з відстанню між p та q.

З теореми Хана-Банаха [1] існують такі $u_1(x),...,u_n(x)$, що:

$$\|L\| = \sum_{k=1}^{n} \sum_{x \in X} |u_k(x)|, \tag{2.10}$$

$$L(Df) = \sum_{k=1}^{n} \sum_{x \in X} D_k f(x) u_k(x), \tag{2.11}$$

Підставимо в ліву частину (2.11) рівність з (2.9), та отримаємо:

$$L(Df) = \sum_{k=1}^{n}\sum_{x\in X} D_k f(x) u_k(x),$$

$$L(Df) = \sum_{x\in X} f(x)(p_x - q_x),$$

$$\sum_{k=1}^{n}\sum_{x\in X} D_k f(x) u_k(x) = \sum_{x\in X} f(x)(p_x - q_x),$$

перекинемо D_k на $u_k(x)$:

$$\sum_{k=1}^{n}\sum_{x\in X} f(x) D_k u_k(x) = \sum_{x\in X} f(x)(p_x - q_x),$$

звідки можемо отримати:

$$\sum_{k=1}^{n}\sum_{x\in X} D_k u_k(x) = p - q, \tag{2.12}$$

отже, ми отримали, для $u(x) : \sum_{k=1}^{n}\sum_{x\in X} D_k u_k(x) = p - q$ буде виконуватися нерівність:

$$\inf_{u(x):\sum_{k=1}^{n}\sum_{x\in X} D_k u_k(x) = p-q} \sum_{k=1}^{n}\sum_{x\in X} |u_k(x)| \leq \sum_{k=1}^{n}\sum_{x\in X} |u_k(x)|,$$

а

$$\sum_{k=1}^{n}\sum_{x\in X} |u_k(x)| = \|L\| = I_*,$$

отже отримали, що для деяких $u(x)$:

$$I_* \geq \sum_{k=1}^{n}\sum_{x\in X} |u_k(x)|.$$

Але з іншого боку у 2.4 було доведено, що

$$I_* \leq \inf \sum_{k=1}^{n}\sum_{x\in X} |u_k(x)|$$

, отже отримуємо рівність:

$$I_* = \sum_{k=1}^{n} \sum_{x \in X} |u_k(x)|,$$

тобто те, що ми і доводимо (1.2).

Використовуючи теорему Канторовича про дуальність та функціональний аналіз, ми довели дивергентне формулювання (1.2).

Висновки до розділу 2

У цьому розділі був розглянутий приклад, який використовується з одного боку для наочної перевірки дивергентного формулювання, а з іншого боку, для демонстрації різниці складності підрахунку на крок складнішого прикладу. Після порівняння складності та демонстрації необхідності дивергентного формулювання, з використанням дуальності Канторовича, а також потужних інструментів функціонального аналізу було доведено дивергентне формулювання в оптимальному транспортуванні ймовірнісних розподілів на просторі Хеммінга. Дане формулювання дасть змогу досліджувати чисельно оптимальну вартість такого перерозподілу.

ВИСНОВКИ

У ході даної роботи було проведено аналіз теорії на тему оптимального транспортування, особлива увага надавалась роботам Канторовича, були проаналізовані його теореми та розглядався простір Хеммінга. Теорема про дуальність Канторовича стала потужним інструментом у ході вирішення задачі дослідження, а саме доведення дивергентного формулювання.

Був наведений приклад часткового випадку, на якому було перевірено коректність доведеного дивергентного формулювання, а наступний приклад продемонстрував своєю складністю, особливо порівняно з першим прикладом, необхідність використання даного формулювання. Також був використаний функціональний аналіз, який на пару з дуальністю Канторовича є фундаментом доведення дивергентного формулювання оптимального транспортування ймовірнісних розподілів на просторі Хемннімга.

Доведене мною дивергентне формулювання стане чудовим інструментом для побудови чисельних схем наближених значень оптимальної вартості транспортування ймовірнісних розподілів на просторі Хеммінга. Розвиток досліджень у цій темі може дати результати для доведень нових дивергентних формулювань, але вже на модифікованих, наприклад, граничних просторах, гаусівських.

ПЕРЕЛІК ПОСИЛАНЬ

[1] Kolmogorov A. N. та Fomin S. V. *Elements of The Theory of Functions and Functional Analysis*. Англ. Т. 1. GRAYLOCK PRESS, 1963.

[2] Riabov G. V. *A representation for the Kantorovich–Rubinstein distance defined by the Cameron–Martin norm of a Gaussian measure on a Banach space // Theory of Stochastic Processes*. Англ. Т. 21. 2016, с. 84—90. URL: http://tsp.imath.kiev.ua/files/2120/art2120_07.pdf.

[3] Villani C. *Topics in optimal transportation*. Англ. Т. 58. American Mathematical Society, 2003.