КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

Факультет комп'ютерних наук та кібернетики

Кафедра обчислювальної математики

Кваліфікаційна робота

на здобуття ступеня магістра

за спеціальністю 113 Прикладна математика

на тему:

Задача оптимального транспорту та рівноваги Неша

Виконав студент 2-го курсу
Олянін Денис Валерійович

_______________ (підпис)

Науковий керівник:
професор, доктор фіз.-мат. наук
Семенов Володимир Вікторович

_______________ (підпис)

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

Студент _______________ (підпис)

Роботу розглянуто й допущено до захисту на засіданні кафедри обчислювальної математики
«05» травня 2023_ р.,
протокол № 7___
Завідувач кафедри
Ляшко Сергій Іванович _______________ (підпис)

Київ – 2023

РЕФЕРАТ

Кількість сторінок – 44, кількість ілюстрацій – 14, кількість використаних джерел – 10.

Перелік ключових слів – ОПТИМАЛЬНИЙ ТРАНСПОРТ, РІВНОВАГА НЕША, ДИСКРЕТНИЙ ОПТИМАЛЬНИЙ ТРАНСПОРТ, ТРАНСПОРТНА ЗАДАЧА, МЕТОД МІНІМІЗАЦІЇ ПОТУЖНОСТЕЙ SINKHORN.

Об'єкт дослідження – задача оптимального транспорту, застосування рівноваги Неша для розв'язування задач оптимального транспорту.

Мета роботи – демонстрація алгоритмів розв'язування задачі оптимального транспорту при дискретних та неперервних розподілах джерел та споживачів продукції, дослідження єдиності розв'язків при використанні рівноваги Неша.

Методи та інструменти – теорія оптимального транспортування, алгоритми для розв'язку задач оптимального транспорту, градієнтний метод для оптимізації; безкоштовне, вільно поширюване інтегроване середовище розробки Visual Studio Code; Мова програмування Python

Результати та їх новизна – для вирішення даної задачі раніше не використовувалась рівновага Неша у поєднанні з оптимальним транспорту.

Взаємозв'язок з іншими роботами – розглянуто модель з джерела [1], застосовано методи з джерел [2], [4].

Зміст

ВСТУП

Оптимальний транспорт є математичною теорією, яка активно розвивається, та знайшла застосування в багатьох галузях, таких як транспортна логістика, економіка, статистика, фізика, геометрія та інші.

Останні роки були відзначені значним розвитком теорії оптимального транспорту, включаючи розв'язання нових задач та розробку нових методів розв'язання існуючих. Одним із сучасних напрямів досліджень в цій області є застосування глибинного навчання для розв'язання задач оптимального транспорту, що відкриває нові можливості в цій області.

Щодо використання рівноваги Неша у теорії оптимального транспорту, то це концепція, яка дає змогу моделювати рівновагу між постачальниками та споживачами, які мають різні потреби та вимоги. Застосування рівноваги Неша в оптимальному транспорті може допомогти знайти оптимальні рішення, які забезпечують мінімальні витрати на транспортування товарів або ресурсів, з урахуванням вимог та потреб різних учасників. Це може бути корисним для вирішення проблем транспортної логістики, зменшення витрат на перевезення та збільшення ефективності.

До інших сучасних напрямів досліджень у теорії оптимального транспорту можна віднести розвиток теорії маршрутів та суперпозиції, дослідження дискретних моделей, застосування геометричних методів та інші. В цілому, теорія оптимального транспорту є досить активним та перспективним напрямком, що відкриває багато можливостей для вирішення складних задач транспортної логістики та оптимізації витрат на транспортування.

Метою роботи є виявлення єдиних розв’язків у вигляді рівноваги Неша в задачах з використанням оптимального транспорту.

Об’єктом дослідження є застосування рівноваги Неша в умовах оптимального транспорту в економічних моделях, існування та єдиність розв’язку. Досягнення розв’язку задач оптимального транспорту зі складними обмеженнями відбувається за допомогою різних оптимізаційних та варіаційних методів, зокрема симплексний метод, метод мінімізації потужності «Sinkhorn», проекційний градієнтний метод.

Даний перелік методів був розроблений мовою програмування Python, використовуючи бібліотеки numpy, torch, POT(Python optimal transport), matplotlib.

1 ТЕОРЕТИЧНА ЧАСТИНА

1.1 Формулювання задачі

Нехай X - простір типів гравців, породжений ймовірності мірою $\mu \in P(X)$, Y - простір дій та функція вартості транспортування $\Phi: X \times Y \times P(Y) \to \mathbb{R}$. Гравці типу x приймають дії у платячи вартість $\Phi(x, y, \nu)$ де $\nu \in P(Y)$ – являє собою розподіл дій. Функція вартості залежить від дій інших гравців тільки через розподіл дій, визначає, що гра є анонімною, тобто не важливо, що за гравець приймає участь у грі.

Визначення. Маржиналами будемо вважати такі функції щільності, які отримані внаслідок інтегрування спільного розподілу γ по X для отримання першої маржинали, та відповідно по Y для отримання другої маржинали.

$$\mu = \int_X \gamma \mathrm{d}x$$

$$\nu = \int_Y \gamma \mathrm{d}y$$

Згідно літератури [1] ми можемо визначити рівновагу Неша через спільний розподіл ймовірностей $\gamma \in P(X \times Y)$ з його першою маржиналою μ такою, що

$$\gamma(\{(x, y) \in X \times Y: \Phi(x, y, z) = \min_{z \in Y} \Phi(x, y, \nu)\}) = 1 \qquad (1)$$

де ν – друга маржинала γ.

У загальному $\gamma(A \times B)$ інтерпретує ймовірність того, що гравець має свій тип A та виконує дію типу B. Рівновага буде називатись чистою, якщо усі гравці використовують чисту стратегію майже всюди по відношенню до μ. Формула (1) стверджує, що усі гравці обирають стратегію для мінімізації значень їх функції вартості транспортування, враховуючи їх типи та ν такі, що ν виступає другою маржиналою розподілу γ. Цю умову можна вважати умовою повноти.

Для підходу оптимального транспорту, що розглядається у даній дипломній роботі ми обмежимось адитивно сепарабельним випадком, де

$$\Phi(x, y, \nu) = c(x, y) + V[\nu](y). \tag{2}$$

Таким чином через сепарабельність зв'язок з оптимальним транспортом виражається у тому, що якщо γ – рівновага Неша, то середнє $c(x, y)$ мінімізується серед ймовірнісних мір μ та ν, які невідомі, як маржинали, тобто розв'язується задача оптимального транспорту:

$$W_c(\mu, \nu) \coloneqq \inf_{\gamma \in \Pi(\mu,\nu)} \iint_{X \times Y} c(x, y) d\gamma(x, y),$$

де $\Pi(\mu, \nu)$ – множина спільних розподілів μ та ν як маржинал. У евклідовому підході є умови на с та μ, які гарантують, що існує така γ яка є чистою, незважаючи який є ν.

Також припускаючи, що функція *V[ν]* є похідною якогось функціоналу $\mathrm{E}(\nu)$, тобто якоюсь варіацією, якщо ν є точкою мінімуму $W_c(\mu, \nu) + \mathrm{E}(\nu)$ та γ розв'язує формулу 2, можна використовувати варіаційний підхід для знаходження рівноваги Неша.

Тоді, мінімізуючи функцію $W_c(\mu, \nu) + \mathrm{E}(\nu)$, ми знайдемо мінімум ν, завдяки якому потім, розв'язавши задачу оптимального транспорту, знайдемо рівновагу Неша γ.

1.2 Дискретне формулювання оптимального транспорту

У даному розділі ми розглядаємо дискретне формулювання задачі оптимального транспорту, яке стосується випадків, коли обидві міри, що розглядаються, дискретні. Таке формулювання дозволяє ефективно застосовувати методи лінійного програмування для вирішення задачі і має практичне застосування в ряді випадків, коли простір та кількість точок обмежені.

Дискретна міра у оптимальному транспорті з вагами $\boldsymbol{a}$ та локаціями $x_1, x_2, \dots, x_n \in X$ визначається як

$$\alpha = \sum_{i=1}^{n} \boldsymbol{a}_i \delta_{x_i}$$

де δ_x - це дельта-функція Дірака на позиції x, де знаходиться одиниця маси нескінченно сконцентрована в локації x.

Зручна особливість оптимального транспорту є в тому, що оптимальний транспорт може мати справу з мірами, які є або дискретними, або неперервними, або обома одночасно в межах однієї задачі. Для цього використовується множина Радонових мір $M(X)$ на просторі X. Формальне визначення цього набору вимагає, щоб X було містив відстань, яку позначимо d, оскільки до міри можна дістатися лише шляхом інтегрування неперервних функцій, позначених як $f \in C(X)$. Інтегрування даних функцій за дискретною мірою α дає нам суму

$$\int_X f(x) d\alpha(x) = \sum_{i=1}^{n} \mathbf{a}_i f(x_i).$$

Радонова міра – це міра на сигма-алгебрі множин на Гаусдорфовому топологічному просторі, яка є локально скінченною та внутрішньо регулярною. Як прикладом міри Радона слугують ймовірнісні міри на сигма-алгебрі Борелівських множин будь-якого повного сепарабельного метричного простору, які якраз і використовуються у даній дипломній роботі.

Для більш загальних мір, наприклад на просторі $X = \mathbb{R}^d$, де $d \in \mathbb{N} -$ це розмірність простору, ми маємо щільність $d\alpha(x) = \rho_\alpha(x)dx$, яка вимірна за Лебегом, і позначимо як $\rho_\alpha = \frac{d\alpha}{dx}$, тобто

$$\forall h \in C(\mathbb{R}^d), \qquad \int_{\mathbb{R}^d} h(x)d\alpha(x) = \int_{\mathbb{R}^d} h(x)\rho_\alpha(x)d(x)$$

Інтегрування будь-якої неперервної функцією $f \in C(X)$ за довільною мірою $\alpha \in M(X)$, яка може і не мати щільності чи бути сумою мір Дірака, даватиме

$$\int_X f(x)\mathrm{d}\alpha(x) \in \mathbb{R}.$$

Якщо X не є компактним, то будемо вимагати, що f має компактний носій, або матиме границю 0 на нескінченності.

1.2.1 Стандартна задача Монжа в оптимальному транспорті

Задача Монжа про оптимальний транспорт бере свій початок у кінці XVIII століття, коли французький математик Гаспар Монж розглянув проблему мінімізації витрат на перевезення певної кількості вантажів з одного місця до іншого. Він запропонував оптимізаційну модель для розподілу вантажів між різними джерелами та призначеннями таким чином, щоб мінімізувати загальні витрати на транспортування.

Дано дві міри маси α та β, визначені на компактних множинах X та Y відповідно, що знаходяться в метричному просторі з відстанню $d(x, y)$. Функція вартості $\mathrm{C}(x, y)$ задає вартість перевезення одиниці маси з точки x до точки y.

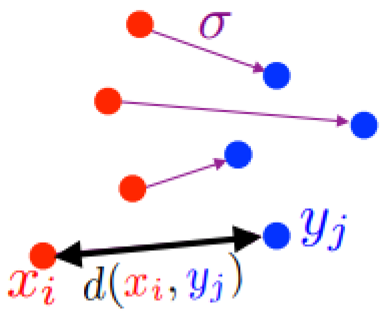

Рисунок 1.1 Задача Монжа про оптимальний транспорт

Покладемо матрицю витрат $C_{i,j}$, де $i \in 1..n, j \in 1..m$, де $n, m \in \mathbb{N}$, також вважатимемо, що $n = m$. Розв'язок цієї задачі полягає у пошуку бієкції σ в множині $Perm(n)$ усіх перестановок n елементів, розв'язуючи наступну задачу мінімізації:

$$\min_{\sigma \in Perm(n)} \frac{1}{n} \sum_{i=1}^{n} C_{i,\sigma(i)}.$$

Для дискретних мір

$$\alpha = \sum_{i=1}^{n} \boldsymbol{a}_i \delta_{x_i} \text{ та } \beta = \sum_{j=1}^{m} \boldsymbol{b}_j \delta_{y_j},$$

розв'язок задачі Монже є таке відображення $T: \{x_1, \dots, x_n\} \to \{y_1, \dots, y_m\}$, що кожній точці x_i відповідає єдина точка y_j та переміщує масу α до маси β, тобто виконується умова

$$\forall j \in 1..m, \quad b_j = \sum_{i:T(x_i)=y_j} a_i.$$

Розв'язок цієї задачі зображений на рисунку 1.2.

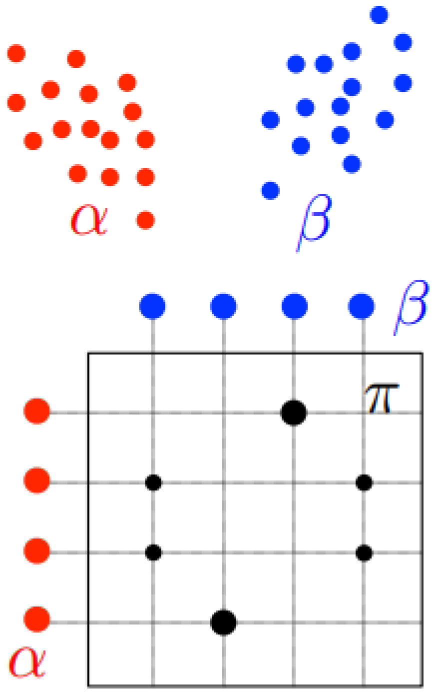

Рисунок 1.2 Дискретний оптимальний транспорт

Для подальшої зручності введемо оператор образу міри. Для будь-якого неперервного відображення $T: X \to Y$ ми введемо оператор образу міри

$$T_{\#}: M(X) \to M(Y).$$

Для дискретних мір це відображення полягає у переміщенні позицій усіх точок, що належать області визначення міри.

$$T_{\#}\alpha = \sum_{i} \mathbf{a}_i \delta_{T(x_i)}.$$

Приклад такого відображення зображений на рисунку 1.3.

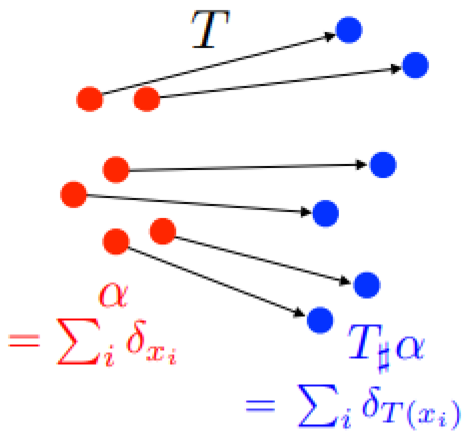

Рисунок 1.3 Оператор образу міри $T_\#$

Наведемо визначення цього ж оператора для більш узагальнених мір:

Для $T: X \to Y$ образ міри під дією цього відображення $\beta = T_\# \alpha \in M(Y)$ для якогось $\alpha \in M(X)$ задовольняє умову

$$\forall h \in C(Y), \qquad \int_Y h(y) \mathrm{d}\beta(y) = \int_X h\big(T(x)\big) \mathrm{d}\alpha(x)$$

Еквівалентно для будь якої вимірної множини $B \subset Y$ виконується:

$$\beta(B) = \alpha(\{x \in X: T(x) \in B\}) = \alpha(T^{-1}(B)).$$

Задача Монжа може бути розширена до випадку, де дві довільні міри (α, β) розташовані на просторах X, Y можуть бути пов'язані відображенням $T: X \to Y$, що мінімізує

$$\min_T \left\{ \int_X c\big(x, T(x)\big) \mathrm{d}\alpha(x): T_\# \alpha = \beta \right\}.$$

Де $T_\# \alpha = \beta$ означає, що T переміщує масу з α до β використовуючи оператор образу міри, що визначили раніше.

Радонові міри можна сприймати як зображення розподілів випадкових величин. Випадкова величина X на просторі X це відображення з ймовірнісного простору $(\Omega, \mathbb{P})$ на простір X і її розподіл α є Радоновою мірою.

1.2.2 Формулювання Канторовича

Леонід Канторович у своїй праці[10] запропонував ідею, що переміщення маси має бути менш детермінованою, тобто маса, яка має бути переміщена з будь-якої точки x_i може транспортуватися до багатьох точок призначення. У такому формулюванні присутня гнучкість у тому, що замість тепер перестановок σ чи відображення T використовується матриця зчеплень $\mathbf{P} \in \mathbb{R}_+^{n\times m}$, де $\mathbf{P}_{i,j}$ описує кількість переміщеної маси з локації i до локації j, або ж з x_i до y_j. Завдяки цьому ми можемо ввести умову збереження маси простішу, ніж відображення Монжа:

$$\mathbf{U}(\mathbf{a},\mathbf{b}) = \left\{\mathbf{P} \in \mathbb{R}_+^{n\times m}: \left(\sum_j \mathbf{P}_{i,j}\right)_i = a, \left(\sum_i \mathbf{P}_{i,j}\right)_j = b\right\}.$$

Таким чином формулювання оптимального транспорту Канторовича можна записати наступним чином:

$$\mathcal{L}_{\mathrm{C}}(\alpha,\beta) = \mathrm{L}_{\mathrm{C}}(\mathbf{a},\mathbf{b}) = \min_{\mathbf{P}\in\mathbf{U}(\mathbf{a},\mathbf{b})} \langle \mathbf{C}, \mathbf{P}\rangle = \sum_{i,j} \mathbf{C}_{i,j}\mathbf{P}_{i,j},$$

де **a** та **b** – відповідні ваги для дискретних мір α та β, а $\mathbf{C}_{i,j} = c(x_i, y_j)$ – функція витрат.

Це формулювання має узагальнений вигляд для неперервних мір $\alpha \in X$ та $\beta \in Y$, де X та Y є компактиними просторами. Тоді умова збереження маси прийматиме наступний вигляд:

$$U(\alpha,\beta) = \left\{\pi \in M_+^1(X\times Y): P_{X\#}\pi = \alpha, P_{Y\#}\pi = \beta\right\},$$

де π $\in M_{+}^{1}(X \times Y)$ – зчеплення двох Радонових мір, що є спільним розподілом на добутку просторів, $P_{X\#}$ та $P_{Y\#}$ - оператори образу міри, проекцій $P_X(x,y) = x$ та $P_Y(x,y) = y$.

Отже, узагальнений вигляд задачі оптимального транспорту у формулюванні Канторовича такий:

$$\mathcal{L}_{\mathrm{C}}(\alpha,\beta) = \min_{\pi \in U(\alpha,\beta)} \int_{X \times Y} c(x,y)\mathrm{d}\pi(x,y), \tag{3}$$

де π є розв’язком задачі оптимального транспорту, зображений на рисунку 2.4.

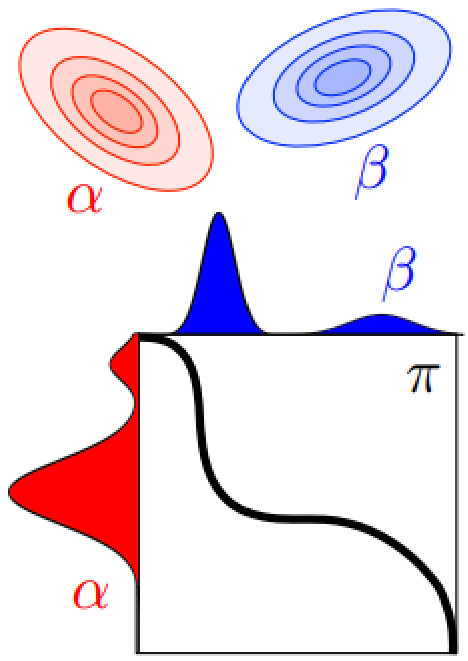

Рисунок 2.4 Оптимальний транспорт при неперервних мірах

1.3. Динамічне формулювання оптимального транспорту

У цілому, з точки зору динаміки оптимального транспорту вартість вимірюється у квадраті геодезичної відстані. Цей підхід описує оптимальне транспортування між двома мірами як криву в просторі мір, яка мінімізує сумарну довжину. У цьому розділі використовується позначення (α_0, α_1) замість (α, β) згідно з концепцією, що ми рухаємося від часу $t = 0$ з однієї міри до іншої в часі $t \ = 1$.

У випадку $X = Y = \mathbb{R}^d$ та $c(x, y) = \|x - y\|^2$ оптимальна транспортна відстань $W_2^2(\alpha, \beta)$, визначена у формулі 3 формулюванні Канторовича, може бути обчислена, шукаючи мінімальний шлях довжини $(\alpha_t)_{t=0}^1$ між цими двома мірами. Цей шлях описується перенесенням міри за допомогою векторного поля v_t, визначеного на кожному моменті часу. Векторне поле v_t та шлях α_t повинні задовольняти умові збереження маси, що призводить до

$$\frac{\partial \alpha_t}{\partial t} + div(\alpha_t v_t) = 0 \text{ та } \alpha_{t=0} = \alpha_0, \qquad \alpha_{t=1} = \alpha_1. \tag{4}$$

Саме векторне поле v_t в таких задачах зазвичай є потенційною течією рідини чи газу у просторі.

Нескінченно мала довжина такого векторного поля вимірюється за допомогою L^2 норми, пов'язаної з мірою α_t, як описано нижче:

$$\|v_t\|_{L^2(\alpha_t)} = \left(\int_{R^d} v_t(x) \mathrm{d}\alpha_t(x) \right)^{\frac{1}{2}}$$

Таким чином посилаючись на дослідження Benamou and Brenier [2] визначається формулювання мінімальної довжини W_2 з такими $(\alpha_t, v_t)_t$, що задовільняють умову 4:

$$W_2^2(\alpha_0, \alpha_1) = \min_{(\alpha_t, v_t)_t} \int_0^1 \int_{R^d} v_t(x) \mathrm{d}\alpha_t(x)\, \mathrm{d}t,$$

де α_t – скалярна міра, та $v_t -$ векторна міра.

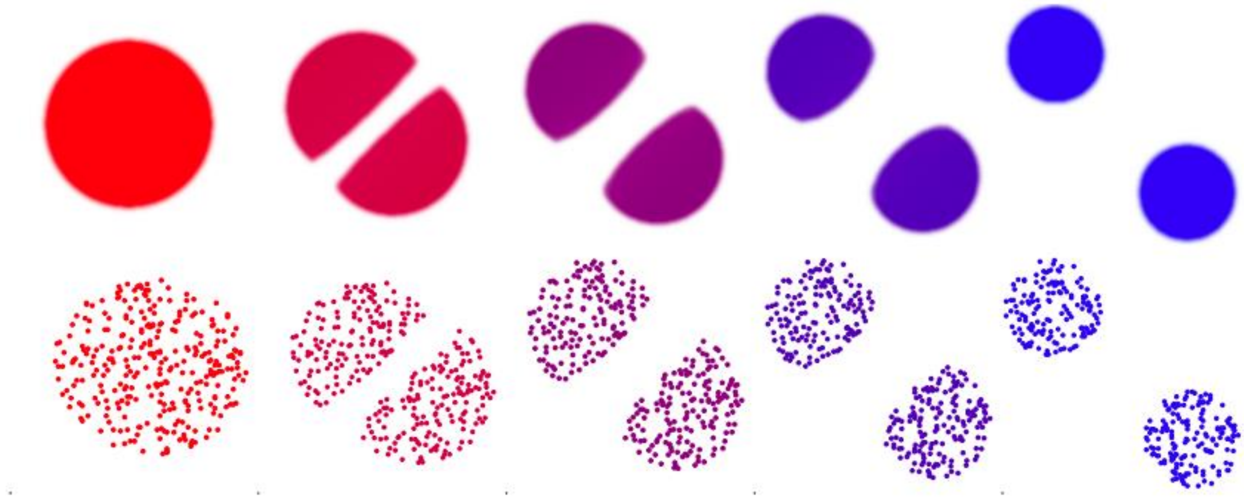

Рисунок 3 Приклад таких шляхів мір, де знизу дискретні міри, а зверху міри з щільністю лебегової міри нуль

Дане формулювання є неопуклим у змінних $(\alpha_t, v_t)_t$ через обмеження з їх добутком в умові 4. У роботі [2] вводять таку векторну міру, що часто називають моментом

$$J_t = \alpha_t v_t,$$

яка натомість дає цьому формулюванні опуклість у змінних$(\alpha_t, J_t)_t$, тоді

$$W_2^2(\alpha_0, \alpha_1) = \min_{(\alpha_t, J_t)_t \in C(\alpha_0, \alpha_1)} \int_0^1 \int_{R^d} \theta\big(\alpha_t(x), J_t(\mathrm{x})\big) \mathrm{d}x \mathrm{d}t,$$

Де $C(\alpha_0, \alpha_1) = \left\{(\alpha_t, J_t)_t : \frac{\partial \alpha}{\partial t} + div(J_t) = 0, \alpha_{t=0} = \alpha_0,\ \alpha_{t=1} = \alpha_1\right\}$ та функція $\theta : \to R^+ \cup \{+\infty\}$, яка має наступний вигляд:

$$\forall (a, b) \in R_+ \times R^d, \qquad \theta(a, b) = \begin{cases} \frac{\|b\|^2}{a} & \text{якщо } a > 0, \\ 0 \ \text{якщо } (a, b) = 0, \\ +\infty & \text{інакше.} \end{cases}$$

2 РІВНОВАГА НЕША

2.1 Модель рівноваги

У моделі введемо компактні метричні простори типів гравців X та дій Y, а також ймовірнісні міри μ та ν для опису розподілу типів та розподілу дій. Взаємодію між агентами описується за допомогою функції витрат та додаткової функції $V[\nu]$, яка залежить від розподілу дій ν. Модель припускає наявність невідомого ймовірнісного розподілу $\gamma \in P(X \times Y)$ в просторі типів та дій, згідно з яким агенти приймають дії. Такий γ породжує розподіл дій ν, його другу маржинальну міру, яку ми позначаємо як $\nu = \pi_{Y\#}\gamma$. За конструкцією, перша маржинальна міра γ, $\pi_{X\#}\gamma$, повинна бути рівною μ. Цей розподіл γ повинен задовольняти певні обмеження, щоб забезпечити належність до деякої області D, а саме:

$$D = \{\nu \in \mathcal{L}^1(m_0): V[\nu] \in \mathcal{L}^1(\nu)\} = \left\{\nu \in \mathcal{L}^1(m_0): \int_Y |V[\nu]|\, d\nu \ < +\infty\right\}.$$

Ці обмеження забезпечують, що функція $V[\nu]$ є обмеженою. Модель також вимагає, щоб соціальні витрати на рівень дій агентів були скінченні та можна їх було обчислити. Соціальні витрати складаються з суми витрат на дії та додаткового члена $V[\nu]$. Для розподілу дій ν, який виникає з розподілу γ, вимагається, щоб його додатковий член $V[\nu]$ також був скінченим.

Розподіл ймовірностей $\gamma \in P(X \times Y)$ є рівновагою Неша, якщо його перша маржинальна міра дорівнює μ, друга маржинальна міра ν належить до D і існує функція $\phi \in C(X)$ така, що

$$c(x, y) + V[\nu] \geq \phi(x) \quad \forall x \in X \text{ відносно міри } m_0 \text{ м. н. для } y, \text{з рівністю } \gamma - \text{м. н.}$$

Рівновага Неша, яка задовольняє ці умови, називається чистою, якщо вона задається графіком, тобто має вигляд $\gamma = (id, T)_{\#}\mu$ для деякого Борелівського відображення $T: X \to Y$.

Як можливий приклад моделі, розглянемо множину агентів, які розподілені відносно певної ймовірнісної міри $\mu \in P(X)$, де $X \subset \mathbb{R}^2$. Ці агенти зацікавлені у надані послуг споживачам, які складають підмножину з простору Y, тобто перевезення продукту. Вартість переміщення продуктів x до споживача y складає $c(x, y)$. В додаток до вартості переміщення також береться до уваги витрати, внаслідок взаємодії між агентами $\nu \to V[\nu]$. Тоді рівновагою ми будемо вважати такий розподіл дій ν, що для розподілу агентів μ вартість переміщення є найменшою.

2.2 Зв'язок з оптимальним транспортом

Для довільного розподілу ймовірностей ν на просторі дій Y, множина $\Pi(\mu, \nu)$ складається з усіх розподілів ймовірностей на $X \times Y$, які мають μ і ν як маржинали. Натомість, $W_c(\mu, \nu)$ є мінімальною вартістю перевезення маси з розподілу μ до розподілу ν за витратами $c(x, y)$ і визначається як значення задачі оптимального транспортування Монжа-Канторовича.

$$\mathrm{W_c}(\mu, \nu) := \min_{\gamma \in \Pi(\mu,\nu)} \iint_{\mathrm{X \times Y}} c(x, y) \mathrm{d}\gamma(x, y)$$

Також введемо непорожню множину планів оптимальних транспортувань $\Pi_0(\mu, \nu)$, який відповідає відповідній умові:

$$\Pi_0(\mu, \nu) := \left\{ \gamma \in \Pi(\mu, \nu) \colon \iint_{\mathrm{X \times Y}} c(x, y) \mathrm{d}\gamma(x, y) = \; Wc(\mu, \nu) \right\}.$$

Першим зв'язком між рівновагою Неша та оптимальним транспортом є наступне спостереження.

Лема 1. Якщо γ – рівновага Неша, та ν – це його друга маржинальна міра, то $\gamma \in \Pi_0(\mu, \nu)$.

Доведення

Покладемо функцію $\phi \in C(X)$ таку, що виконується умова (та умова, де є обмеження на гамму і рівновагу Неша) та нехай існує $\eta \in \Pi(\mu, \nu)$, тоді маємо

$$\iint_{X\times Y} c(x,y)\, d\eta(x,y) \geq \iint_{X\times Y} (\phi(x) - V[\nu](y))\, d\eta(x,y)$$

$$= \iint_{X} \phi(x)\, d\mu(x) - \iint_{Y} V[\nu](y)\, d\nu(y) = \iint_{X\times Y} c(x,y)\, d\gamma(x,y)$$

Отже, $\gamma \in \Pi_0(\mu, \nu)$, що і треба було довести.

Доведення цієї леми також показує, що $\phi(x)$ розв'язує дуальність $W_c(\mu, \nu)$, тобто максимізує функціонал

$$\int_X \phi(x) d\mu(x) + \int_Y \phi^c(y) d\nu(y),$$

де ϕ^c – c-трансформація функції ϕ, тобто

$$\phi^c(y) = \min_{x\in X}\{c(x,y) - \phi(x)\}.$$

У роботі МакКенна та Гангбо [2] було наведено зручний наслідок, що гарантує що оптимальний γ буде обов'язково чистим, незалежно від ν, при абсолютній неперервності μ відносно міри Лебега, а $c(x,y)$ є гладкою та строго опуклою функцією $x - y$:

Наслідок 1. Припустимо, що $X = \Omega$, де Ω - це відкрита, зв'язна та обмежена підмножина R^d з мінімальною межею,, нехай μ є абсолютно неперервною за міри Лебега, функція c диференційована по своєму першому аргументу, градієнт $\nabla_x c$ неперервний на $R^d \times Y$, і він задовольняє узагальнену умову Спенса-Міррлеса[3]: для кожного $x \in X$, відображення $y \in Y \to \nabla_x c(x, y)$ є ін'єктивним. Тоді для кожного $\nu \in P(Y)$, $\Pi_0(\mu, \nu)$ містить лише один елемент, і цей елемент має вигляд $\gamma = (id, T)_{\#}\mu$. Отже, усі рівноваги Неша є чистими.

2.3 Варіаційний підхід

Основним припущенням для варіаційного підходу є те, що відображення $\nu \to V[\nu]$ є диференціалом за визначенням 1. У цьому випадку, варіаційний підхід базується на спостереженні, що умова рівноваги є умовою оптимальності першого порядку для мінімізації $Wc(\mu, \nu) + E[\nu]$.

Визначення 1 Нехай D визначена як у першому розділі. Відображення $\nu \in D \to V[\nu]$ є диференціалом на D, якщо D є опуклою і існує функція E: $D \to R$, така що для кожної пари $(\rho, \nu) \in D^2$, $V[\nu] \in \mathcal{L}^1(\rho)$ і

$$\lim_{\varepsilon \to 0^+} \frac{E[(1-\varepsilon)\nu + \varepsilon\rho] - E[\nu]}{\varepsilon} = \int_Y V[\nu] d(\rho - \nu),$$

тобто V[ν] є першою варіацією E, яку ми позначаємо $V[\nu] = \frac{\delta E}{\delta \nu}$.

Розглянемо локальний випадок, де m_0 є нашою мірою, і для $\nu \in D = P(Y) \cap \mathcal{L}^1(m_0)$ та m_0 майже всюди на y для деякої неперервної f:

$$V[\nu] = f\big(y, \nu(y)\big).$$

По-перше, покладемо, що f обмежена, і визначена, для всіх $\nu \in D$ таких, що:

$$F(y,\nu) = \int_0^{\nu} f(y,s)ds\,, \qquad E[\nu] = \int_Y F\big(y,\nu(y)\big)dm_0(y) \tag{5}$$

Оскільки F є Ліпшицевою по ν, так і по y, то з теореми Лебега про мажоровану збіжність випливає, що $V[\nu]$ є диференціалом E на D. Припустимо, що f задовольняє умову зростання для деяких $a \ \geq \ 0,\ b \ > \ 0,\ \alpha \ > \ 0$, що для кожного $\nu \ \geq \ 0$ і майже скрізь для y:

$$a(\nu^{\alpha} - 1) \leq f(y,\nu) \leq b(\nu^{\alpha} + 1).$$

Для $p = \alpha + 1$ відповідний функціонал E визначений для всіх $\nu \in D = P(Y) \cap \mathcal{L}^p(m_0)$ як у формулі 5. Завдяки умові зростання та теоремі Лебега про мажоровану збіжність, V є диференціалом E на D, $V[\nu] \in \mathcal{L}^{p'}(m_0)$, де p' - спряжений показник до p, тобто $p' = \frac{p}{p-1}$, якщо $\nu \in D$. Окрім умови росту, яка дозволяє застосовувати теорему Лебега та гарантує, що $V[\nu]\nu$ є інтегровним, ми бачимо, що локальні V є диференціалами.

Розглянемо тепер випадок взаємодії, де $V[\nu]$ визначений для деякого $\varphi \in C(Y \times Y)$ як

$$V[\nu](y) = \int_Y \varphi(y,z)d\nu(z).$$

Тоді визначимо квадратичний функціонал

$$E[\nu] = \frac{1}{2}\iint_{Y\times Y} \varphi(y,z)\, d\nu(y)\, d\nu(z)\ .$$

Розкриваючи по ε, $E[\nu \ + \ \varepsilon(\rho \ - \ \nu)]$, легко обчислити його диференціал

$$\lim_{\varepsilon\to 0}\frac{E[\nu + \varepsilon(\rho - \nu)] - E[\nu]}{\varepsilon}$$
$$= \frac{1}{2}\iint \varphi(y,z)[d\nu(y)\, d(\rho - \nu)(z) + d\nu(z)\, d(\rho - \nu)(y)]$$
$$= \frac{1}{2}\iint [\varphi(y,z) + \varphi(z,y)]\, d\nu(z)\, d(\rho - \nu)(y).$$

Тому

$$\frac{\delta E}{\delta \nu}(y) = \int\limits_{Y} \varphi^{sym}(y,z)d\nu(z) \quad : \varphi^{sym}(y,z) = \frac{\varphi(y,z) + \varphi(z,y)}{2}.$$

Отже, V є диференціалом E на $P(Y)$ в тому разі, якщо φ є симетричною, тобто $\varphi(y,z) = \varphi(z,y)$, що, наприклад, має місце, якщо φ є функцією відстані між y та z. Безсумнівно, ми можемо розглядати більш загальні випадки, де функціонал V є сумою симетричного взаємодійного члена та локального члена. Такі потенціали V все ще зберігають диференціальну структуру

Нехай функціонал V на D приймає наступний вигляд:

$$V[\nu] = \frac{\delta E}{\delta \nu}. \tag{6}$$

Тоді розглянемо варіаційну задачу

$$\inf_{\nu\in D} J_{\mu}[\nu] \tag{7}$$

де $J_{\mu}[\nu] := Wc(\mu,\nu) + E[\nu]$.

Щоб довести, що точки мінімуму формули (7) є рівновагою, нам потрібно мати можливість диференціювати вираз $Wc(\mu,\nu)$ по ν. Для цього потрібна певна структура на *X*, μ та *c*: зокрема, $X \subset R^d$, c диференційована відносно x, а μ - еквівалентна мірі Лебега.

Лема 2. Нехай $X = \overline{\Omega}$, де $\Omega \subset \mathbb{R}^d$ – відкрита, обмежена та зв'язна з межею, якою можна нехтувати, що μ еквівалентний Лебеговій мірі на X. Тоді для будь-якого $y \in Y$, $c(.,y)$ є диференційовний по x та $\nabla_x c$ є обмеженим на $X \times Y$. Нехай $\nu \in P(Y)$, тоді існує єдиний потенціал Канторовича ϕ: $\phi^c = \min_{x \in X}\{c(x,y) - \phi(x)\}$ $\forall y \in Y$ між μ та ν, що $\forall p \in P(Y)$

$$\lim_{\varepsilon \to 0} \frac{W_c\big(\mu, \nu + \varepsilon(p - \nu)\big) - W_c(\mu, \nu)}{\varepsilon} = \int_Y \phi^c \mathrm{d}(p - \nu).$$

Теорема 1 Точки мінімуму є рівновагами:

Нехай $V[\nu]$ є диференціалом, згідно формули (6) із $D = P(Y) \cap \mathcal{L}^p(m_0)$ для деякого $p \in [1, +\infty]$. Використовуючи попередню лему, якщо ν розв'язує задачу з формули (7), а $\gamma \in \Pi o(\mu, \nu)$, то γ є рівновагою Неша.

Проте зазначимо, що умова оптимальності для формули 7 полягає в наступному: існує стала M майже скрізь на ν, що

$$\begin{cases} \phi^c + V[\nu] \geq M, \\ \phi^c + V[\nu] = M, \end{cases}$$

де ϕ^c - це c-трансформація ϕ, $\phi^c = \min_{x \in X}\{c(x,y) - \phi(x)\}$ $\forall y \in Y$.

Щоб отримати результат існування з Теореми 1, припустимо, що *V[ν]* визначено для $\nu \in P(Y) \cap \mathcal{L}^1(m_0)$ за формулою:

$$V[\nu] = f\big(y, \nu(y)\big) + \int_Y \varphi(y,z) d\nu(z), \tag{8}$$

де $\varphi \in C(Y \times Y)$ є симетричною, f є неперервною та монотонно зростає з другим аргументом та задовольняє умову росту для деяких $a > 0$, $b > 0$ та $\alpha > 0$. Для

$p = \alpha + 1$, відповідний енергетичний функціонал визначений для всіх $\nu \in D = P(Y) \cap \mathcal{L}^p(m_0)$ за формулою:

$$E[\nu] = \int_Y F\big(y, \nu(y)\big) \mathrm{d}m_0(y) \ + \frac{1}{2} \ \varphi(y,z) \ d\nu(y) \ d\nu(z),$$

де F визначена за формулою (5). Функціонал $F(y, .)$ є опуклим та задовольняє умову росту $a(p^{-1} \nu^p - \nu) \leq F(y, \nu) \leq b(p^{-1} \nu^p + \nu)$.

Таким чином, $V[\nu] \in \mathcal{L}^{p'}(m_0)$ як тільки, коли $\nu \in D$, і, за нерівністю Гьольдера,

$$V[\nu]\rho \in \mathcal{L}^1(m_0) \quad \forall \rho, \nu \in D.$$

Наслідок 2 (Існування рівноваги за допомогою мінімізації): Припустимо, що виконуються передумови Леми A.1[1], $\nu \to V[\nu]$ має форму (8), де f та φ задовольняють вищезгадані умови, тоді (7) має мінімуми в $P(Y) \cap \mathcal{L}^p(m_0)$, звідси і з'являється існування рівноваги Неша.

Оскільки ϕ є c-перетворенням, вона є неперервною, тому обмеженою на Y, а інтегральний член є обмеженим, оскільки φ обмежена. Таким чином,отримаємо $\nu \in L^\infty(m_0)$.

У варіаційному підході дуже важлива опуклість: якщо E є опуклою, то пошук рівноваг та мінімізація J_μ еквівалентні, тобто, якщо виконуються умови Теореми 3.2, якщо E є опуклою на D, то наступні твердження еквівалентні:

— ν є розв'язком (3.5), а $\gamma \in \Pi_o(\mu, \nu)$,

— γ є рівновагою, а $\nu = \Pi Y \# \gamma$.

Якщо, крім того, E є строго опуклою, то виконується наступний результат про єдиність:

Наслідок 3 (Єдиність розв'язку в строго опуклому випадку) Припустимо, що виконуються умови Теореми 1. Якщо E є строго опуклою, то всі рівноваги мають спільний другий маржинальний елемент ν. Якщо, крім того, виконуються умови наслідку 1, то існує не більше однієї рівноваги Неша.

2.4 Опуклість

Досі наш підхід, заснований на варіаціях, дозволив нам довести існування рівноваги через задачу мінімізації (7). Попередні результати не є повністю задовільними через можливість неоптимальних рівноваг та незручність умов оптимальності. За строгих умов опуклості ми досягаємо єдиності оптимальних рівноваг. Однак, враховуючи конкуренцію між опуклістю та неопуклістю в $E[\nu]$, у випадку

$$E[\nu] = \int_Y F\big(y, \nu(y)\big) dm_0(y) + \frac{1}{2} \iint_{Y\times Y} \varphi(y,z) d\nu(y) d\nu(z)\,,$$

ми можемо використовувати спеціальну структуру опуклості для отримання нових результатів щодо єдиності та характеристик.

Розглянемо одновимірний випадок, де функціонал J_μ не є опуклим відносно ν, але є опуклим щодо T, відображення оптимального транспортування від μ до ν. Покладемо $X = Y = [0,1]$, m_0 - це міра Лебега на $[0,1]$, μ - абсолютно неперервна відносно міри Лебега, і покладемо $V[\nu]$:

$$V[\nu](y) = f\big(\nu(y)\big) + \nu(y) + \int_{[0,1]} \varphi(y,z)\, d\nu(z)$$

та вартість перенесення c має вигляд:

$$c(x,y) = C(x-y),$$

де C є строго опуклою та диференційованою;

f є зростаючою;

ν є опуклою на $[0,1]$;

φ опукла, симетрична, диференційована і має локальний Ліпшицевий градієнт.

У загальному випадку вартість

$$E[\nu] = \int_{[0,1]} F\big(\nu(y)\big)dy + \frac{1}{2}\iint_{[0,1]^2} \varphi(y,z)d\nu(y)d\nu(z) + \int_{[0,1]} \nu(y)d\nu(y),$$

та функціонал $J_{\mu} = W_c(\mu,.) + E$ не є опуклими. Однак, J_{μ} має властивості опуклості, при наступній інтерполяції.

Нехай $(\rho,\nu) \in P([0,1])^2$, тоді існують єдині відображення оптимального транспорту T_0 від μ до ν, та T_1, відповідно від μ до ρ, для вартості c і вони є неспадними[5]. Для $t \in [0,1]$ визначимо: $\nu_t := T_t\#\mu$, де $T_t = ((1-t)T_0 + tT_1)$, тоді за конструкцією крива $t \to \nu_t$ з'єднує $\nu_0 = \nu$ та $\nu_1 = \rho$.

Визначення 2 Функціонал $J: P(Y) \to R \cup \{+\infty\}$ називається опуклим за переміщенням (Displacement convex), якщо $t \in [0,1] \to J[\nu_t]$ є опуклим (для кожного вибору кінцевих точок ν та ρ). Він називається строго опуклим за переміщенням, коли, крім того, $J[\nu_t] < (1-t)J[\nu] + tJ[\rho]$, коли $t \in (0,1)$ та $\rho \neq \mu$.

Нехай (ν,ρ) - дві ймовірнісні міри в області визначення *E* (яка є опуклою за опуклістю F), визначимо ν_t як вище, та розглянемо чотири складові J_{μ} окремо.

За визначенням W_c, ν_t та внаслідок строгої опуклості *C* маємо:

$$\begin{aligned} W_c(\mu,\nu_t) &\leq \int_0^1 C\Big(x - \big((1-t)T_0(x) + tT_1(x)\big)\Big)d\mu(x) \\ &\leq (1-t)\int_0^1 C\big(x - T_0(x)\big)d\mu(x) + t\int_0^1 C\big(x - T_1(x)\big)d\mu(x) \\ &= (1-t)W_c(\mu,\nu) + tW_c(\mu,\rho), \end{aligned}$$

при $t \in (0,1)$ та $\nu \neq \rho$.

Конструкцією

$$\int_0^1 v \, d\nu_{\mathrm{t}} = \int_0^1 v(T_{\mathrm{t}}(x)) \, d\mu(x),$$

є опуклою по відношенню до t, завдяки опуклості v.

Аналогічно

$$\iint_{[0,1]^2} \varphi(y,z) \, d\nu_{\mathrm{t}}(y) \, d\nu_{\mathrm{t}}(z) = \iint_{[0,1]^2} \varphi(T_{\mathrm{t}}(x), T_{\mathrm{t}}(y)) \, d\mu(x) \, d\mu(y)$$

є опуклою по відношенню до t, за опуклістю φ.

Опуклість останнього доданка є складнішою. Оскільки $\nu_{\mathrm{t}} = T_{\mathrm{t}}\#\mu$ і T_{t} є неспадним, то принаймні формально ми маємо $\nu_{\mathrm{t}}(T_{\mathrm{t}}(x))T'_{\mathrm{t}}(x) = \mu(x)$. За формулою заміни змінних маємо:

$$\int_0^1 F(\nu_{\mathrm{t}}(y)) \, dy = \int_0^1 F(\nu_{\mathrm{t}}(T_{\mathrm{t}}(x)))T'_{\mathrm{t}}(x) \, dx = \int_0^1 F\left(\frac{\mu(x)}{T'_{\mathrm{t}}(x)}\right)T'_{\mathrm{t}}(x) dx.$$

Закінчимо спостереженням, що $\alpha \to F(\mu(x)\alpha^{-1})\alpha$ є опуклою і що $T_t'(x)$ є лінійною по відношенню до t.

Отже, J_μ є строго опуклим за переміщенням, і приймає не більше одного мінімуму. Таким чином було досягнуто результат про єдиність рівноваги Неша:

Теорема 2 про єдиність рівноваги за допомогою опуклості зміщення у одновимірному випадку: За наведеними вище припущеннями, еквівалентність ν є точкою мінімуму функціоналу в задачі (7) та $\gamma \in \Pi_o(\mu,\nu) \leftrightarrow \gamma$ є рівновагою. Оскільки J_μ є строго опуклим за переміщенням, існує єдина рівновага, яка обов'язково є чистою.

Дане доведення можна узагальнити на більшу розмірність, коли оптимальний транспорт є квадратичним.

Покладемо $X = Y = \overline{\Omega}$, $\Omega \subset \mathbb{R}^d$- відкрита, опукла та обмежена, μ є неперервною за Лебегом міра та має додатньовизначену щільність на Ω. Вартість транспортування у нас квадратична, тобто

$$c(x, y) = \frac{1}{2}|x - y|^2.$$

Тоді функціонал V приймає форму:

$$V[\nu](y) = f\big(\nu(y)\big) + \nu(y) + \int_Y \phi(y, z)\mathrm{d}\nu(z),$$

де ν- опукла, $\phi \in C(\mathbb{R}^d \times \mathbb{R}^d)$ є симетрична і належить $C^{1,1}_{loc}$, тобто C^1 з локальним Ліпшивецим градієнтом та f задовольняє наступні умови:

$$\lim_{\nu \to 0^+} f(\nu) = -\infty \text{ та } \lim_{\nu \to +\infty} f(\nu) = +\infty.$$

Відповідно, позначивши через F первісну f функціонал E матиме вигляд:

$$E[\nu] = \int_Y F\big(\nu(y)\big)\mathrm{d}y + \int_Y \nu(y)\mathrm{d}\nu(y) + \frac{1}{2}\iint_{Y^2} \phi(y, z)\mathrm{d}\nu(z)\mathrm{d}\nu(y).$$

Оскільки вартість транспортування у нас квадратична, тоді використовуючи теорему 2 ми отримаємо єдиність та чистоту оптимальних шляхів γ між μ та довільному ν. Тоді наша варіаційна задача приймає форму

$$\inf_{\nu \in P(\overline{\Omega})} J_\mu[\nu],$$

де $J_\mu[\nu] = \frac{1}{2}W_2^2(\mu, \nu) + E[\nu]$, а $W_2^2(\mu, \nu)$ – відстань Вассерштайна між μ та ν:

$$W_2^2(\mu, \nu) = \inf_{\gamma \in \Pi(\mu, \nu)} \int_{X^2} |x - y|^2 \mathrm{d}\gamma(x, y)$$

3 МЕТОДИ ОБЧИСЛЕННЯ РОЗВ’ЯЗКУ

Розглянемо чисельні методи для розв'язання задачі оптимального транспорту у двох варіантах: дискретному та неперервному.

3.1 Методи лінійного програмування

У дискретному випадку задача оптимального транспорту полягає в знаходженні оптимального плану переміщення маси між двома скінченими наборами точок з вагами. Тут ми маємо справу з задачею лінійного програмування, яка може бути вирішена за допомогою різних чисельних методів:

3.1.1 Транспортна задача лінійного програмування

У розділі 1 було наведено задачу оптимального транспорту за формулюванням Канторовича:

$$\mathrm{L}_{\mathbf{C}}(\mathbf{a}, \mathbf{b}) = \min_{\mathbf{P} \in \mathbf{U}(\mathbf{a}, \mathbf{b})} \langle \mathbf{C}, \mathbf{P} \rangle \ = \sum_{i,j} \mathbf{C}_{i,j} \mathbf{P}_{i,j}$$

Дану рівняння можна перетворити у стандартну форму задачі лінійного програмування, з лінійною цільовою функцією, обмеженнями рівності, що визначаються матрицею, та вектором констант, обмеженнями невід’ємності на змінних.

3.1.2 Симплексний метод

Цей класичний метод лінійного програмування вирішує задачу, переміщуючись вздовж крайніх точок допустимої множини. Це ітераційний процес, який продовжується до знаходження оптимального розв'язку, що має поліноміальну складність.

3.1.3 Методи внутрішньої точки

Ці методи працюють на основі опуклої оптимізації, відштовхуючись від внутрішньої точки допустимої множини та знаходячи оптимальний розв'язок, переміщуючись в середині множини. Вони забезпечують кращу швидкість збіжності порівняно з симплексним методом та ефективніше масштабуються для великих задач.

Мінімізувати $\sum_{i,j} c_{i,j} x_{i,j} - \mu * \sum_{i,j} log\left(x_{i,j}\right)$ з обмеженнями:

— $\sum_j x_{i,j} = a_i$ для всіх i,
— $\sum_i x_{i,j} = b_j$ для всіх j,

де $\mu > 0$ - це параметр бар'єрної функції, $x_{i,j}$ - це кількість ресурсу, який транспортується від джерела i до цілі j, $c_{i,j}$ - це вартість транспортування одиниці ресурсу від i до j, a_i - це загальна кількість ресурсу в джерелі i, і b_j - це необхідна кількість ресурсу в цілі j.

Далі задача може бути вирішена за допомогою методів оптимізації, таких як метод Ньютона. По мірі того, як рішення зближується до оптимального, параметр μ зменшується, що дозволяє наблизитися до області, де обмеження є активними.

3.1.4 Угорський алгоритм

Це спеціальний метод для розв'язання задачі оптимального транспорту з квадратною матрицею вартості. Він забезпечує збіжність до оптимального розв'язку за поліноміальний час, шукаючи відповідність між відправниками та отримувачами таким чином, щоб мінімізувати загальну вартість транспортування.

Припустимо, що у нас є квадратна матриця вартостей C, де $c_{i,j}$ є вартістю перевезення від і-гої точки відправлення до j-гої точки призначення.

Алгоритм містить наступні кроки:

— Ініціалізація. Для кожного рядка матриці вартостей віднімаємо найменше значення цього рядка від усіх елементів цього рядка. Також це робимо для кожного стовпця. Тобто, формуємо нову матрицю C', де $c'_{i,j} = c_{i,j} - min(c_{i,k}) - min(c_{k,j})$ для всіх k.
— Використовуємо мінімальну кількість горизонтальних або вертикальних ліній, щоб покрити всі нулі в матриці. Якщо число ліній дорівнює розміру матриці, то ми знайшли оптимальне рішення.
— Якщо поточне покриття не є оптимальним, обираємо найменший непокритий елемент матриці, позначимо його через h. Віднімаємо h від усіх непокритих елементів матриці та додаємо h до елементів, що лежать на перетині ліній покриття.
— Повертаємось до кроку 2 і повторюємо процес, доки не буде знайдено оптимальне рішення.

3.2 Методи неперервної оптимізації

У неперервному випадку задача оптимального транспорту стосується двох неперервних мір розподілу. Чисельні методи для таких задач включають:

3.2.1 Методи градієнтного спуску

Ці методи полягають у мінімізації функціоналу Канторовича за допомогою ітеративного оновлення плану транспортування, використовуючи градієнтний спуск або інші методи оптимізації. Можуть бути застосовані різні варіанти градієнтного спуску, такі як стохастичний градієнтний спуск або градієнтний спуск з адаптивним розміром кроку.

3.2.2 Методи мінімізації потужності (Sinkhorn)

Алгоритм Sinkhorn - це ще один важливий метод для вирішення задачі оптимального транспорту. Він особливо корисний для обробки великих даних завдяки своїм ефективним обчислювальним властивостям. При неперервних розподілах спершу їх потрібно дискретизувати, аби можна було використовувати алгоритм Sinkhorn.

Нехай маємо два дискретні розподіли, репрезентовані векторами $a \in \mathbb{R}_+^n$ та $b \in \mathbb{R}_+^m$, які сумуються до одиниці. Також маємо матрицю вартостей $C \in \mathbb{R}^{n\times m}$.

Цільова функція з регуляризацією ентропією має вигляд:

$$L_\varepsilon(P) = \langle P, C\rangle_F + \varepsilon H(P),$$

де $\langle P, C\rangle_F = \sum_{i,j} P_{i,j} C_{i,j}$ – скалярний добуток Фробеніуса,

$H(P) = -\sum_{i,j} P_{i,j} C_{i,j}$ – це ентропія матриці P,

а $\varepsilon > 0$ – параметр регуляризації.

Ми шукаємо матрицю P таку, що мінімізує $L_\varepsilon(P)$, за умови, що маржинали матриці P дорівнюють a і b.

Сам алгоритм:

1. Ініціалізуємо $u_0 = 0, v_0 = 0$
2. Повторюємо для кожного k=1,2,… до збіжності:
   a. $u_k = \frac{a}{Kv_{k-1}}$
   b. $v_k = b/(K^{\mathrm{T}} u_k)$

3. Результатом є матриця $P = diag(u_k)Kdiag(v_k)$, де $K = e^{-\frac{C}{\varepsilon}}, diag(x)$ – діагональна матриця з вектором x на діагоналі, а операція ділення та множення виконуються поелементно.

3.2.3 Диференціальні рівняння з частинними похідними

Для деяких задач оптимального транспорту можна отримати аналітичний розв'язок або зробити спрощення, що дозволяє переформулювати проблему як систему диференціальних рівнянь з частинними похідними. Чисельні методи для розв'язання цих рівнянь, такі як методи скінченних різниць або методи скінченних елементів, можуть бути застосовані для знаходження наближених розв'язків. У динамічному формулюванні оптимального транспорту, ціль полягає у мінімізації витрат, пов'язаних з переміщенням розподілу маси від одного часу до іншого, які можуть бути модельовані як рух частинок в потоці.

Оптимальний транспорт розв’язується за допомогою варіаційних методів (продовження) мінімізації та методи прямо-двоїстого відображення. Варіаційні методи дозволяють врахувати різні обмеження та регуляризаційні члени, роблячи їх гнучкими та широко застосовними для різних задач оптимального транспорту.

4 РЕЗУЛЬТАТИ ОБЧИСЛЕНЬ

У даному розділі представленні результати обчислення наведених методів на різних випадках: дискретному та неперервному.

Розглянемо прикладну задачу дискретному випадку, де у нас є множини пекарень та кафе, які мають своє розташування у двовимірному просторі. Це доволі інтуітивна задача, де пекарні повині надати свої продукти кафе, розв'язки якої є оптимальне транспортування, яке ми обчислимо, використовуючи метод мінімізації потужності та лінійне програмування для формулювання Канторовича.

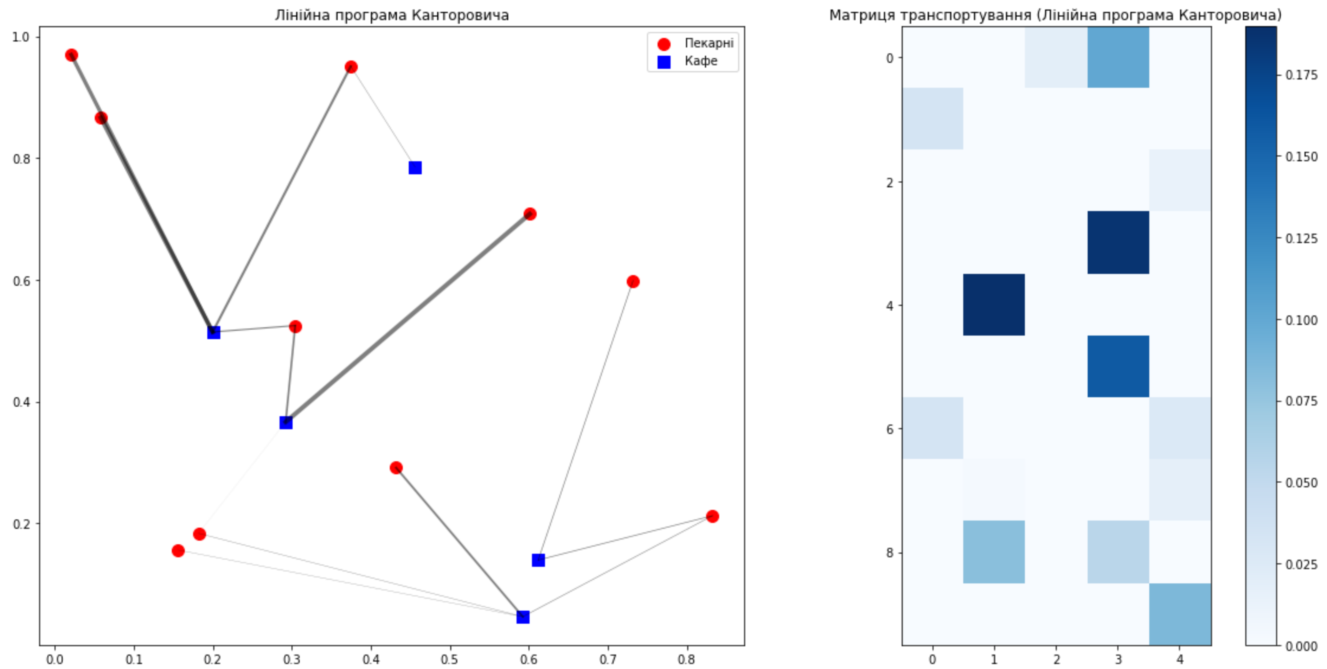

Рисунок 4.1 Результати лінійного програмування

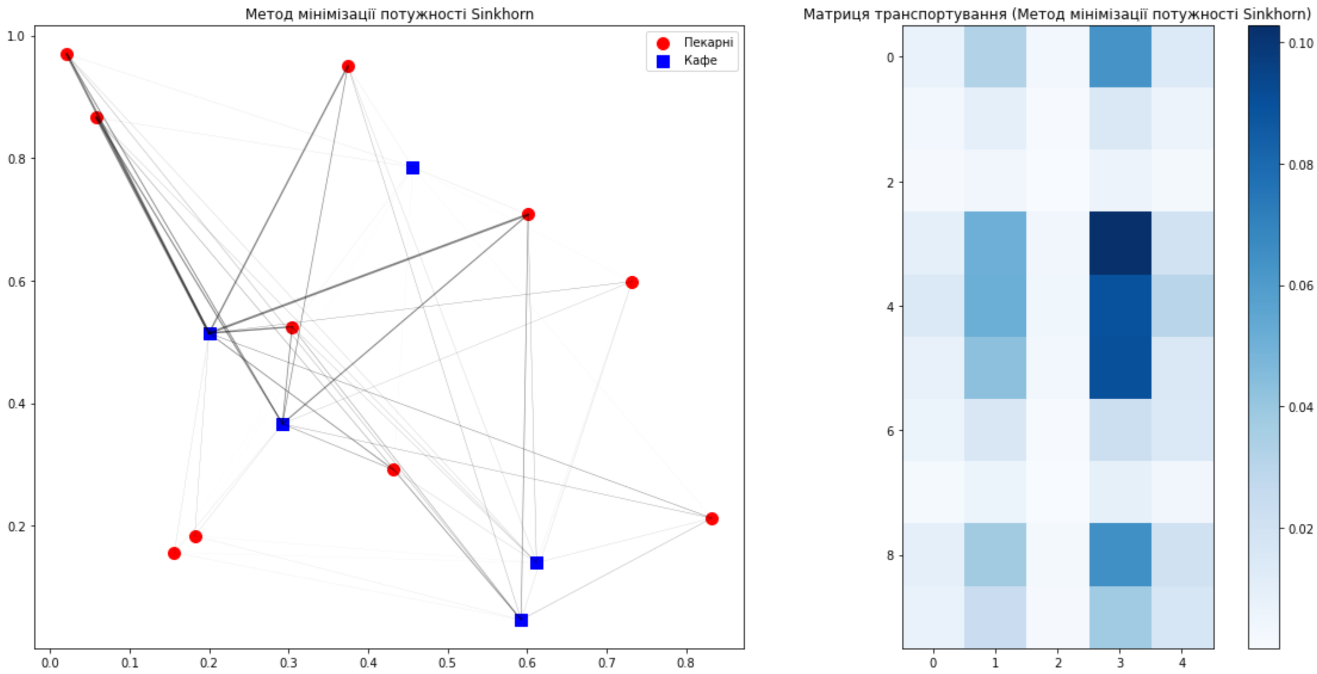

Рисунок 4.2 Метод мінімізації потужності Sinkhorn

У неперервному випадку нехай у нас розподіл джерел продукції рівномірно належить розподілу, що рівний сумі трьох нормальних розподілів $\mu \sim N(5,5) + N(35,15) + N(20,40)$, а призначення для транспортування рівномірно належить розподілу, вигляд якого є сума двох нормальних розподілів $\nu \sim N(20,10) + N(80,10)$, зображено на рисунку 4.3.

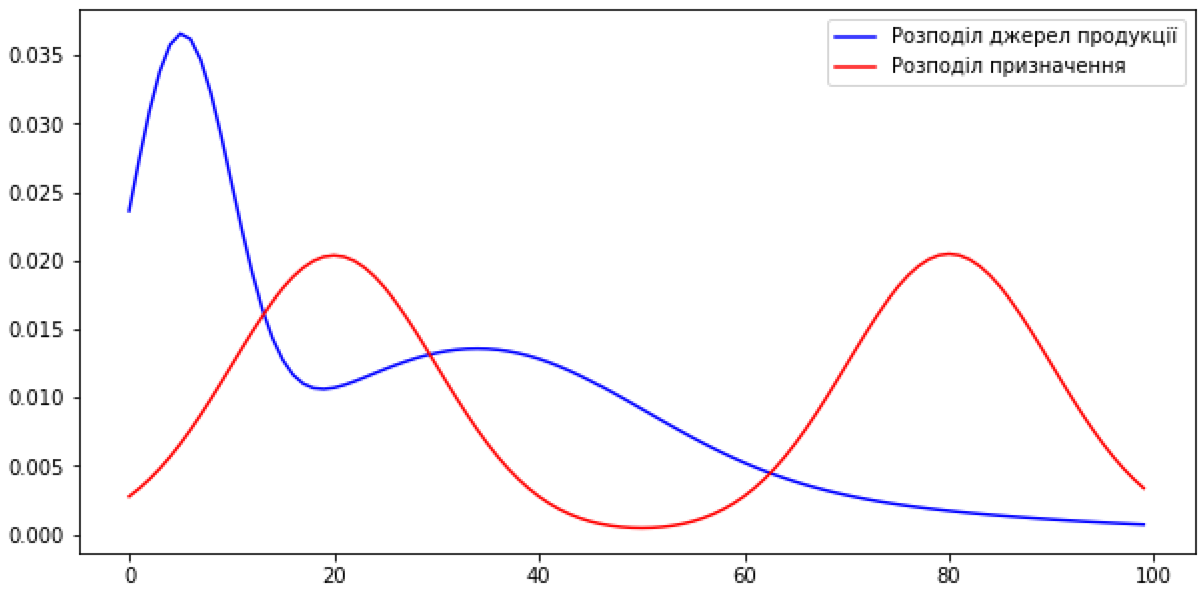

Рисунок 4.3 Розподіли джерел та призначення

При такому формулюванні для розв'язку задачі буде використовуватись дискретизація на 100 точок, тоді матриця вартостей транспортування буде відстанню між локаціями джерел та призначення, зображена на рисунку 4.4. Елементи матриці в даному випадку є евклідовою відстанню між точками розподілів.

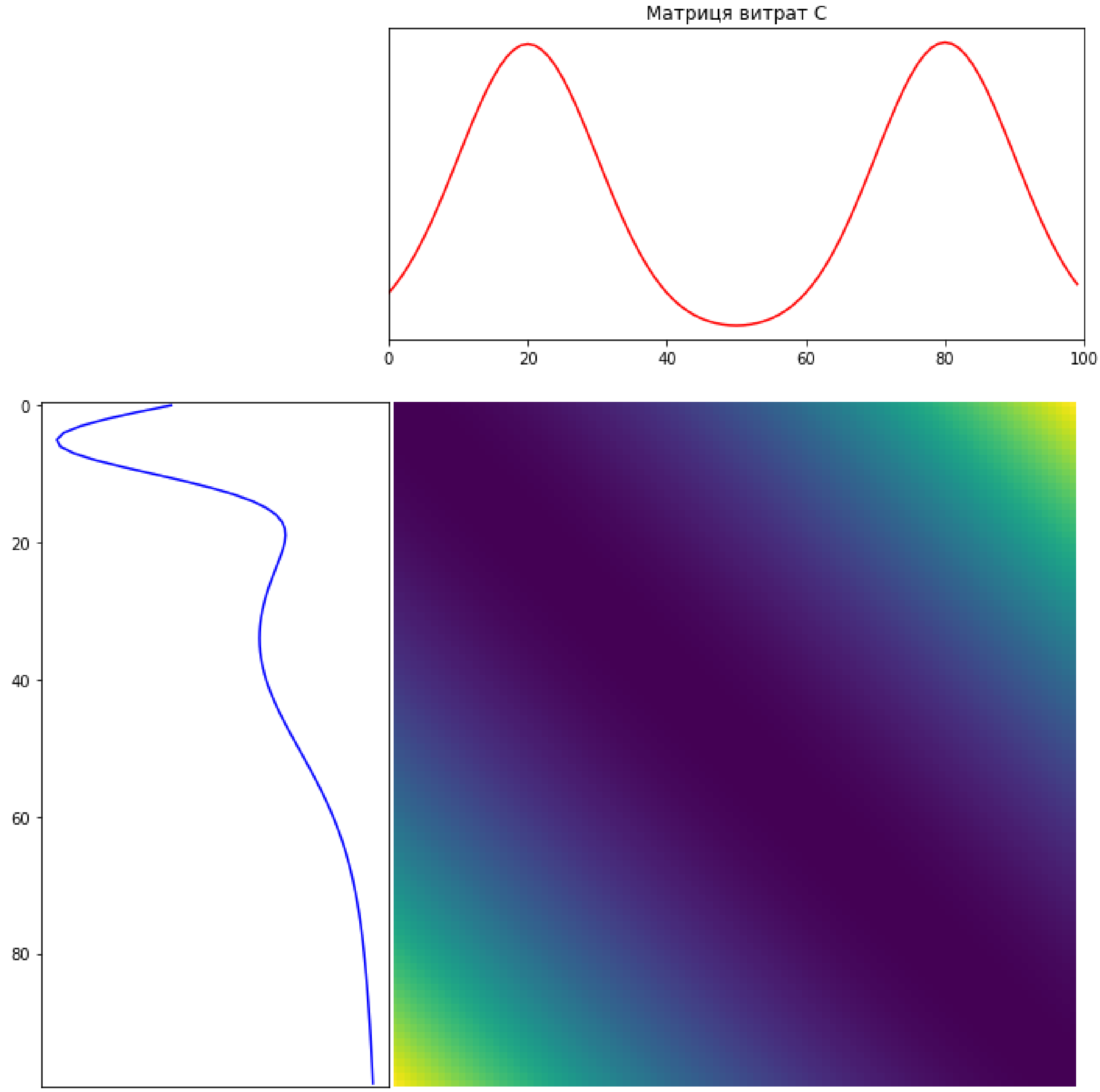

Рисунок 4.4 Матриця витрат

Тепер знайдемо оптимальне транспортування за допомогою методу мінімізації потужності Sinkhorn з параметром регуляризації 0.001.

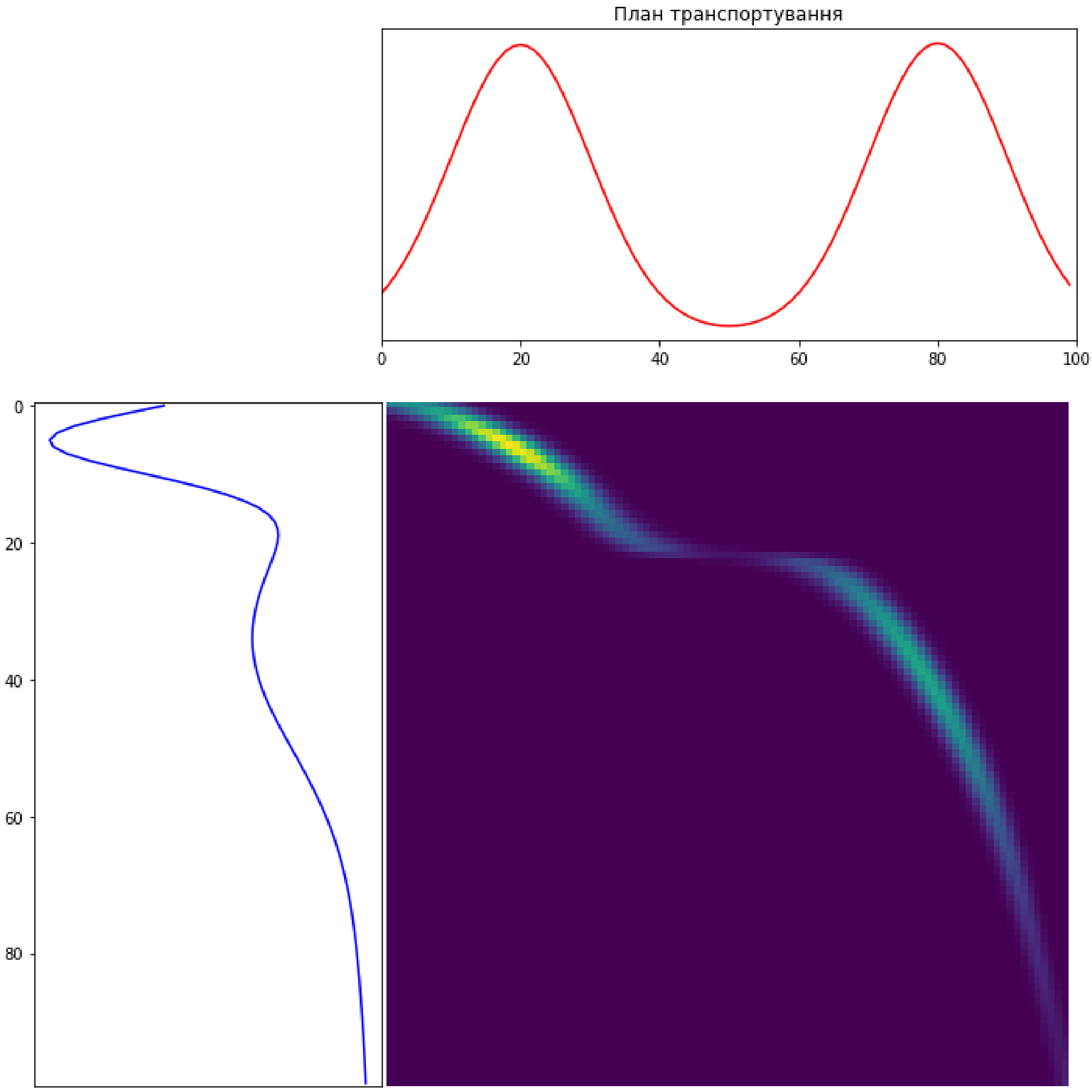

Рисунок 4.5 Оптимальне транспортування методом Sinkhorn, з коефіцієнтом регуляризації 0.001

Також розглянемо випадок, коли розподіл джерел продукції рівномірно належить розподілу, що рівний сумі трьох нормальних розподілів $\mu \sim N(10,10) + N(50,10) + N(90,10)$, а призначення для транспортування рівномірно належить розподілу, вигляд якого є сума двох нормальних розподілів $\nu \sim N(25,10) + N(75,10)$, зображено на рисунку 4.6.

Рисунок 4.6 Розподіли джерел та призначення

Матриця витрат така ж як на рисунку 4.4. Тоді оптимальний план з регуляризацією з коефіцієнтом в одну тисячну матиме вигляд як на рисунку 4.7.

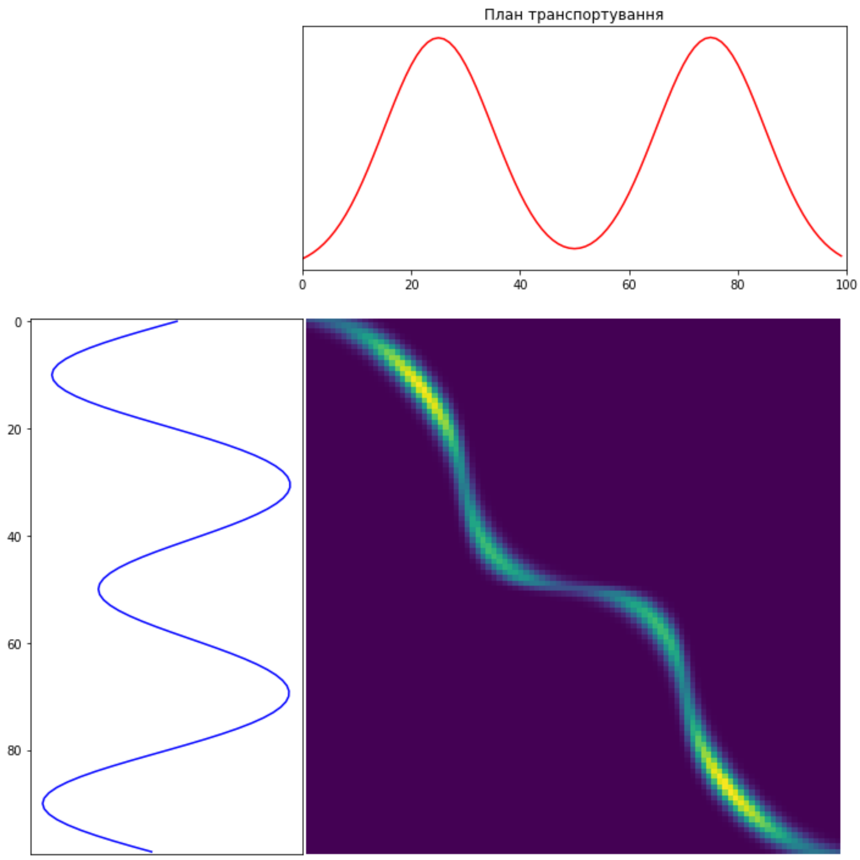

Рисунок 4.7 Оптимальний план транспортування

Нехай $\Omega = (0, 1)$, m_0 - це міра Лебега на (0, 1), μ абсолютно неперервна за m_0 з додатньовизначеною щільністю, яка позначається μ. Розв'яжемо варіаційну задачу

$$\inf_{\nu} J_\mu[\nu],$$

де $J_\mu[\nu]$відповідно рівний:

$$J_\mu[\nu] = \frac{1}{2}\, W_2^2(\mu, \nu) + \int_0^1 \nu f(\nu) \; + \frac{1}{2} \iint_{[0,1]^2} \varphi(y, z) \mathrm{d}\nu(y) \mathrm{d}\nu(z),$$

де $\varphi \in C_{loc}^{1,1}$ є опуклою та симетричною.

Шуканий нами ν є точкою мінімуму функціоналу $J_\mu[\nu]$. Для мінімізації даного функціоналу використовуватимемо проекційний градієнтний метод.

Покладемо $\mu \sim N(20{,}5) + N(60{,}20)$, $\phi(z) = 10^{-4}|z|^2$, $f = x^8$. Відслідкуємо похибку та результат мінімізаці на рисунках 4.8 та 4.9

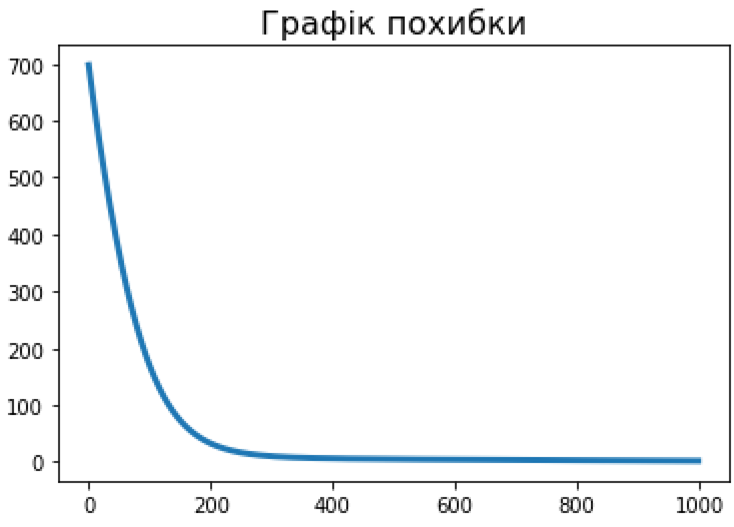

Рисунок 4.8 Графік похибки відносно кількості ітерацій проекційного градієнтного методу.

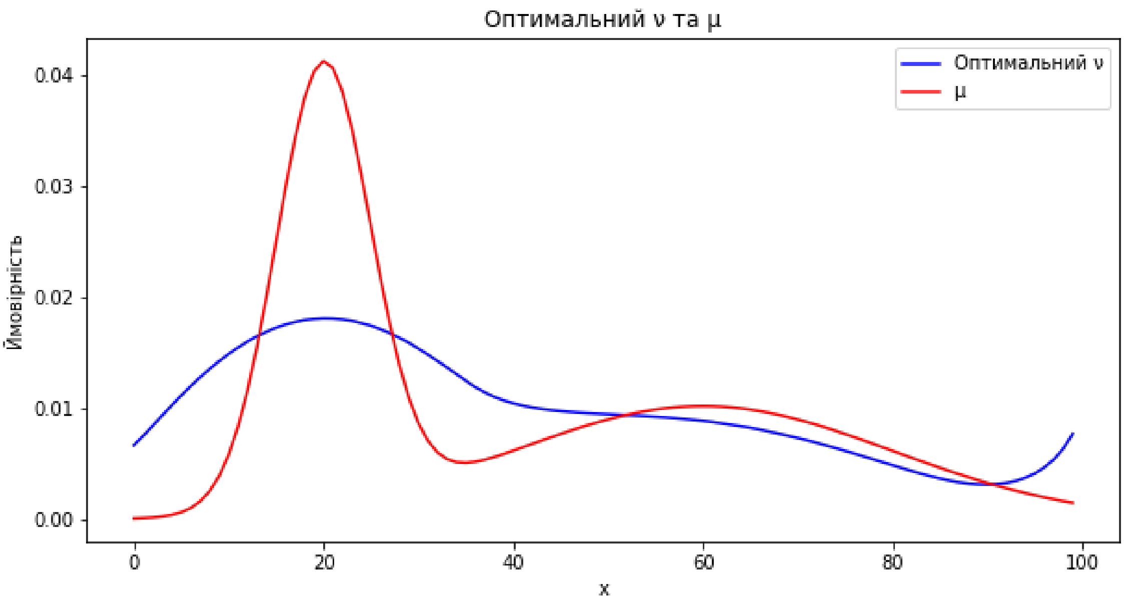

Рисунок 4.9 Графік оптимального ν, що мінімізує функціонал J_{μ}

ВИСНОВКИ

У цій роботі було розглянуто розв’язок оптимального транспорту як задачу для пошуку рівноваги Неша у некооперативних іграх, а саме виявлено та представлено опуклі структури моделі. Було показано існування та єдиність розв’язків. Для розв’язання поставленої задачі було використано проекційний градієнтний метод для пошуку мінімуму опуклого функціоналу. Також було представлено алгоритми для пошуку оптимального транспорту у дискретному випадку формулювання задачі так і неперервному випадку, такі як лінійна програма Канторовича, метод внутрішньої точки та угорський алгоритм, метод мінімізації потужності Sinkhorn.

При виконані роботи було розроблено мовою програмування Python алгоритми, що розв’язують задачі оптимального транспорту, як і у дискретному так і у неперервному випадках.

Задача оптимального транспорту може бути використана для планування та оптимізації транспортних маршрутів і витрат на перевезення. Також і у сфері обробки зображень в алгоритмах для вирівнювання кольорів між зображеннями, видалення шуму з зображень. Ще один напрям впровадження оптимального транспорту є статистика та аналіз даних, наприклад, порівняння розподілів даних, побудови гістограм, оцінки відстані між розподілами. Рівновага Неша є основоположним інструментом в теорії ігор для моделювання стратегічного взаємодії між різними гравцями. Це може бути використано в багатьох економічних ситуаціях, включаючи аукціони, біржі, відносини роботодавців і працівників, і так далі. Рівновага Неша і задача оптимального транспорту використовуються в багатьох алгоритмах машинного навчання. Наприклад, вони можуть бути використані для створення більш ефективних методів класифікації, кластеризації, регресії та генеративних моделей.

Напрямок використання оптимального транспорту та рівноваги Неша з відкритими в них опуклими структурами є перспективним для подальших досліджень, адже кількість сфер застосувань дуже широка і важлива для оптимізації витрат для різних бізнесів. Сам напрям є молодим та має багато відкритих дверей для продовження досліджень.

СПИСОК ЛІТЕРАТУРИ

1. Blanchet A., Carlier G. Optimal transport and cournot-nash equilibria. Mathematics of operations research. 2016. Vol. 41, № 1. P. 125–145. URL: https://doi.org/10.1287/moor.2015.0719 (дата звернення: 11.05.2023).
2. Benamou J.-D., Brenier Y. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numerische mathematik. 2000. Vol. 84, № 3. P. 375–393. URL: https://doi.org/10.1007/s002110050002 (дата звернення: 13.05.2023).
3. Acciaio B., Veraguas J. B., Jia J. Cournot--Nash equilibrium and optimal transport in a dynamic setting. SIAM journal on control and optimization. 2021. Vol. 59, № 3. P. 2273–2300. URL: https://doi.org/10.1137/20m1321462 (дата звернення: 13.05.2023).
4. Blanchet A., Carlier G., Nenna L. Computation of cournot–nash equilibria by entropic regularization. Vietnam journal of mathematics. 2017. Vol. 46, № 1. P. 15–31. URL: https://doi.org/10.1007/s10013-017-0255-x (дата звернення: 13.05.2023).
5. Carlier G. Duality and existence for a class of mass transportation problems and economic applications. Advances in mathematical economics. Tokyo, 2003. P. 1–21. URL: https://doi.org/10.1007/978-4-431-53979-7_1 (дата звернення: 13.05.2023).
6. Gangbo W., McCann R. J. The geometry of optimal transportation. Acta mathematica. 1996. Vol. 177, № 2. P. 113–161. URL: https://doi.org/10.1007/bf02392620 (дата звернення: 13.05.2023).
7. 16. Numerical Methods for Large-Scale Optimal Transport / N. Tupitsa. ArXiv.org. URL: https://doi.org/10.48550/arXiv.2210.11368 (дата звернення: 11.05.2023).
8. Peyré G., Cuturi M. Computational optimal transport: with applications to data science. Foundations and trends® in machine learning. 2019. Vol. 11, № 5-6.

P. 355–607. URL: https://doi.org/10.1561/2200000073 (дата звернення: 13.05.2023).

9. Villani C. Topics in optimal transportation. Providence, RI : American Mathematical Society, 2003. 370 p.
10. Л. В. Канторович, "О перемещении масс", Теория представлений, динамические системы. XI, Специальный выпуск, Зап. научн. сем. ПОМИ, 312, ПОМИ, СПб., 2004, 11–14; J. Math. Sci. (N. Y.), 133:4 (2006), 1381–1382