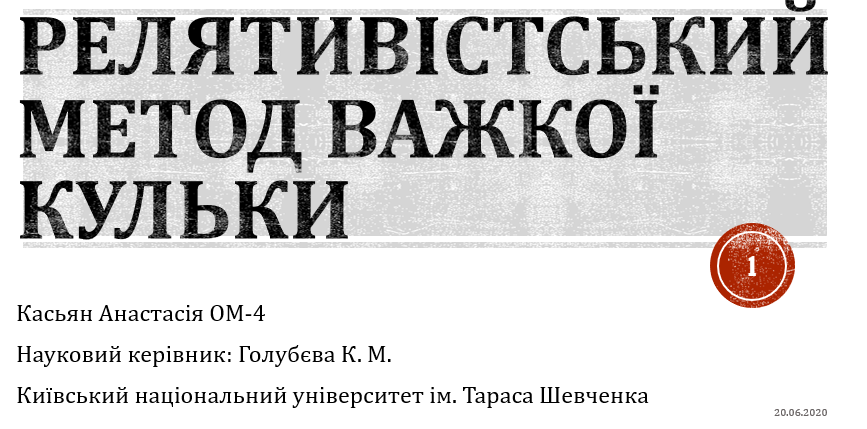
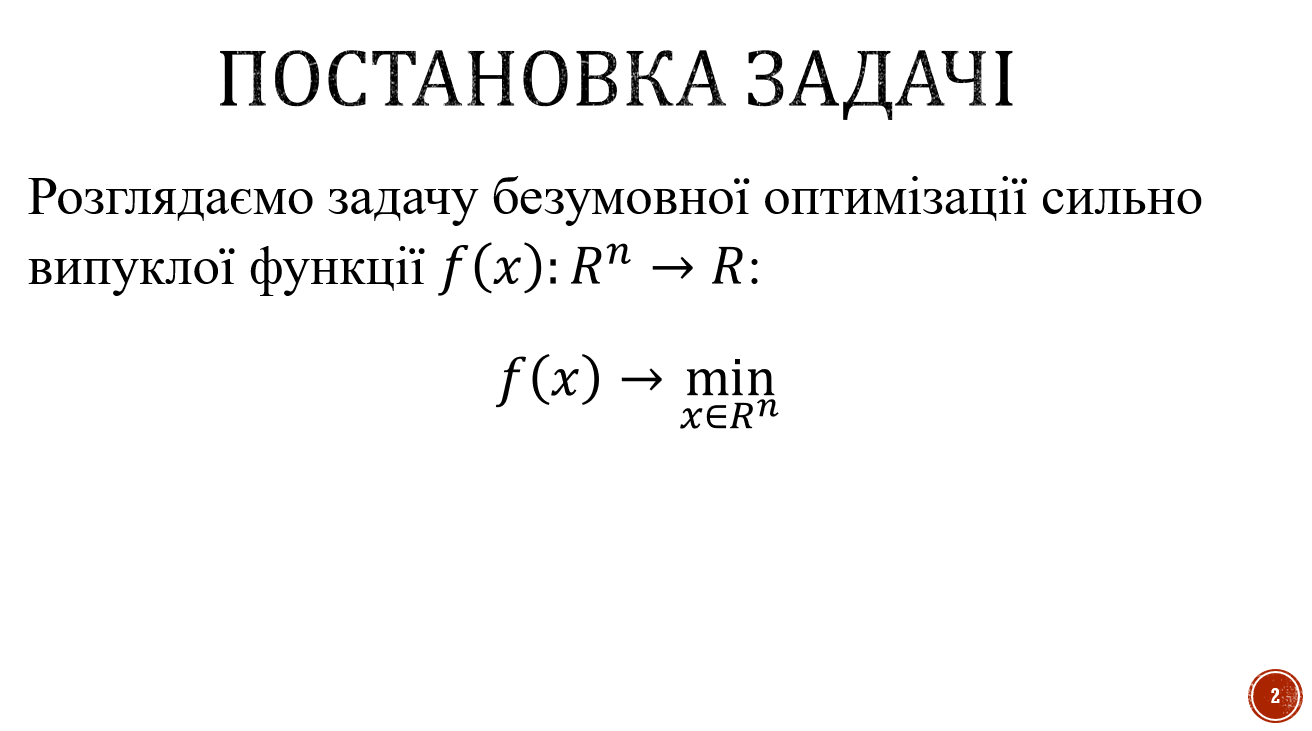
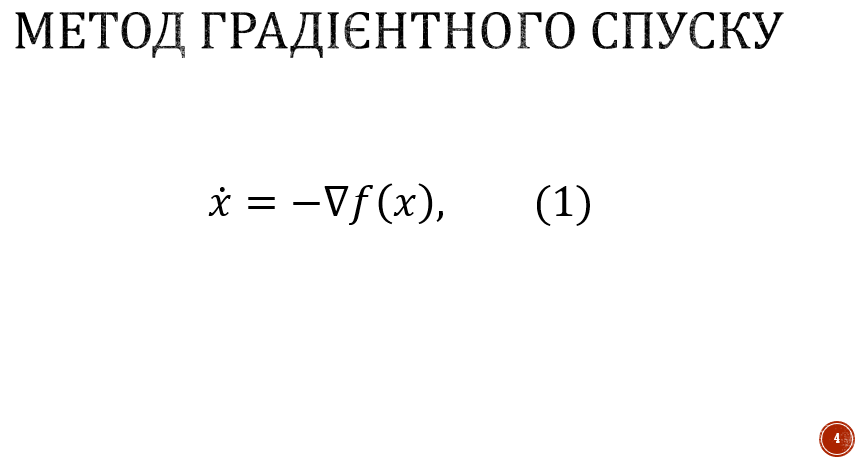
1. ****

Кваліфікаційна робота на тему

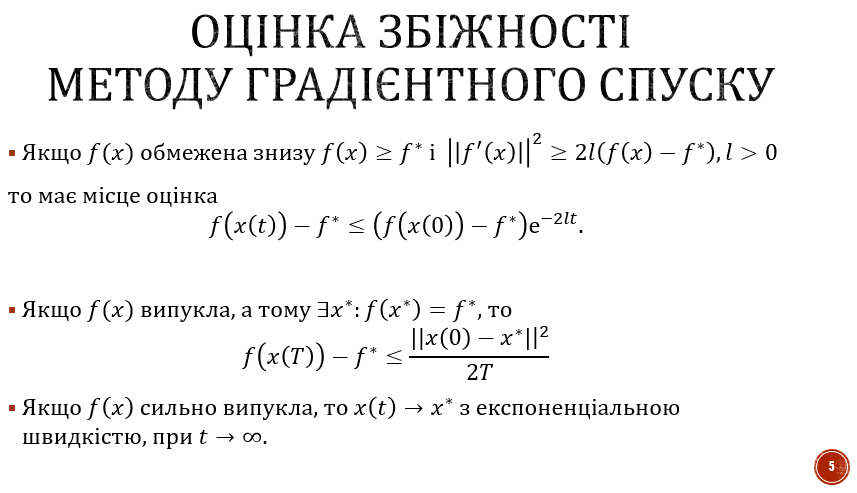
1. 

Я розглядала задачу мінімізації дійсної функції багатьох змінних.

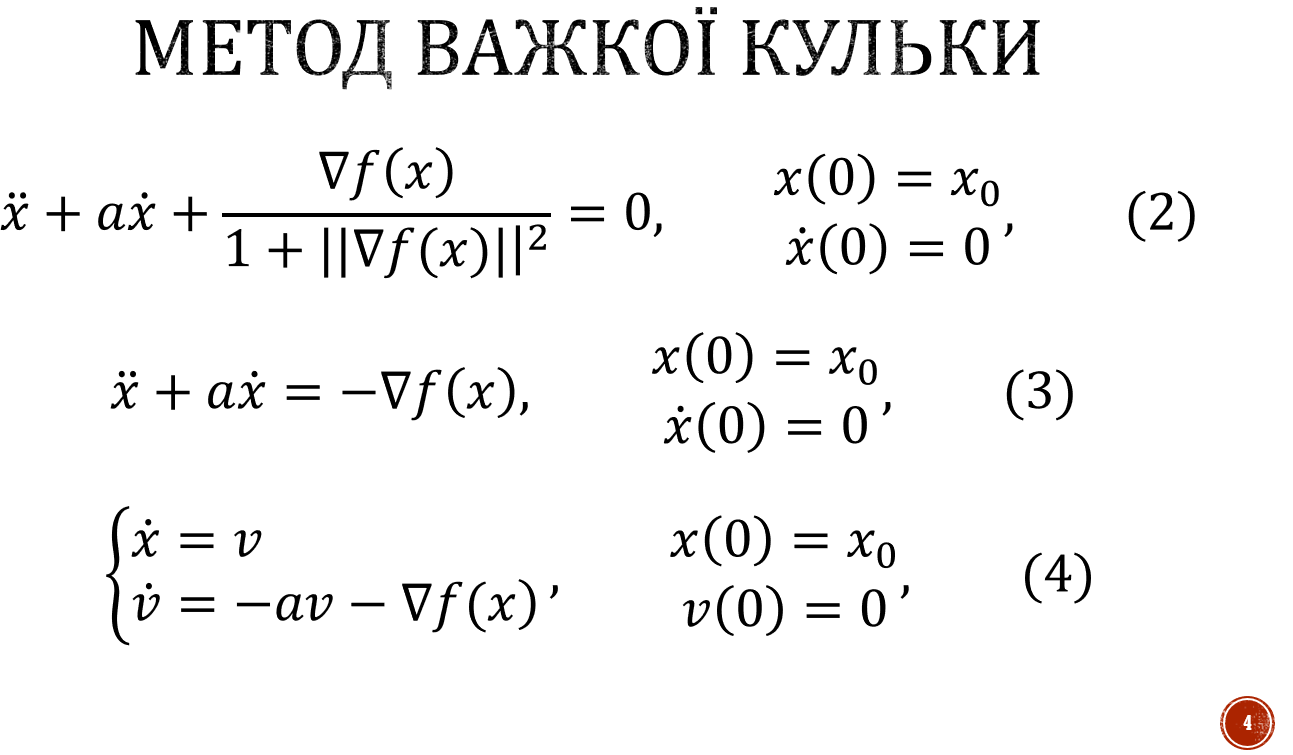
З теоретичної точки зору лише у випадку сильно випуклих функцій є можливість доведення існування розв’язку задачі мінімізації.

1. ****

Метод градієнтного спуску це найпростіший з наближених методів пошуку мінімуму функції. Найбільш відомим є ітераційний варіант цього методу, але в цій роботі розглядається неперевний варіант, який бачимо на екрані.

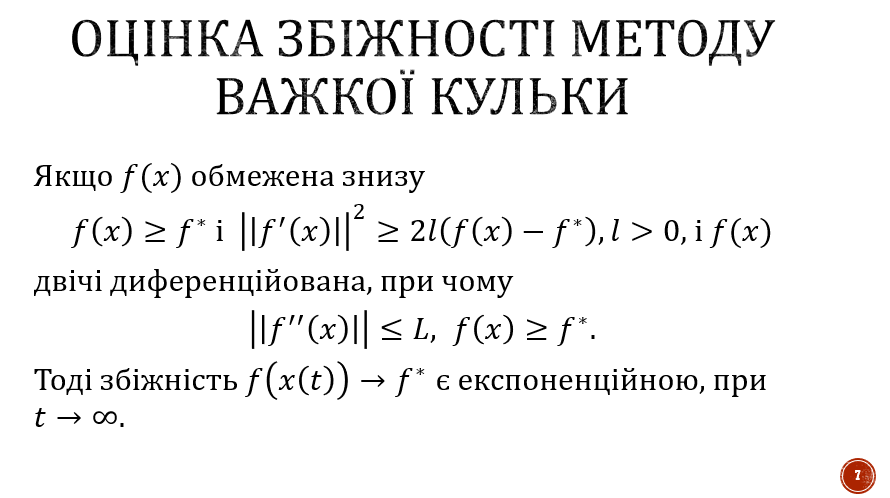
1. ****

Дати оцінку збіжності методу градієнтного спуску можливо лише для класу сильно опуклих функцій. У цьому випадку градієнтний метод збігається до точки мінімуму з експоненційною швидкістю (при певних обмеженнях на цільову функцію).

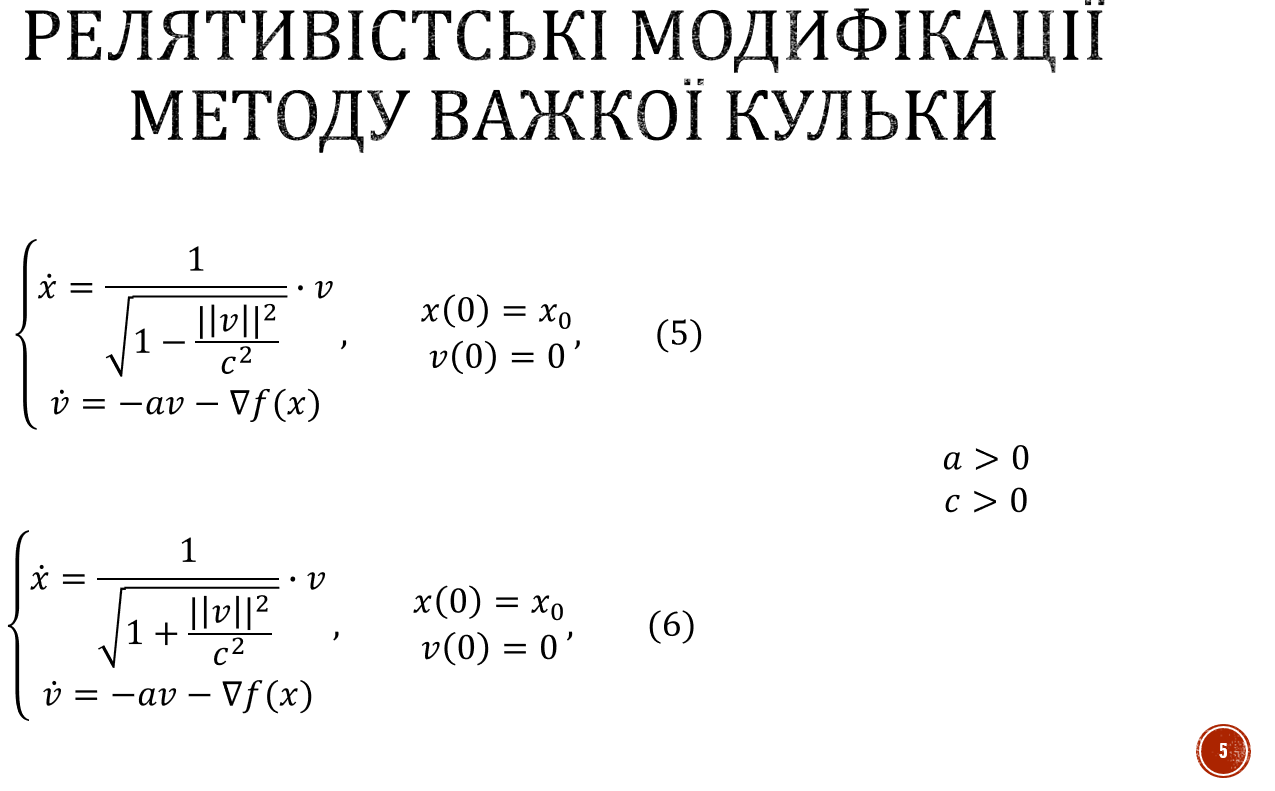
1. 

Рівняння методу важкої кульки (2) може бути отримане після деяких спрощень з рівняння руху матеріальної точки по поверхні заданій функцією . Поле тяжіння діє у від’ємному напрямку осі . При цьому вважається, що точка не може відірватися від поверхні. Також наявний супротив руху (тертя або супротив повітря), який пропорційний швидкості. Звідси походить назва методу.

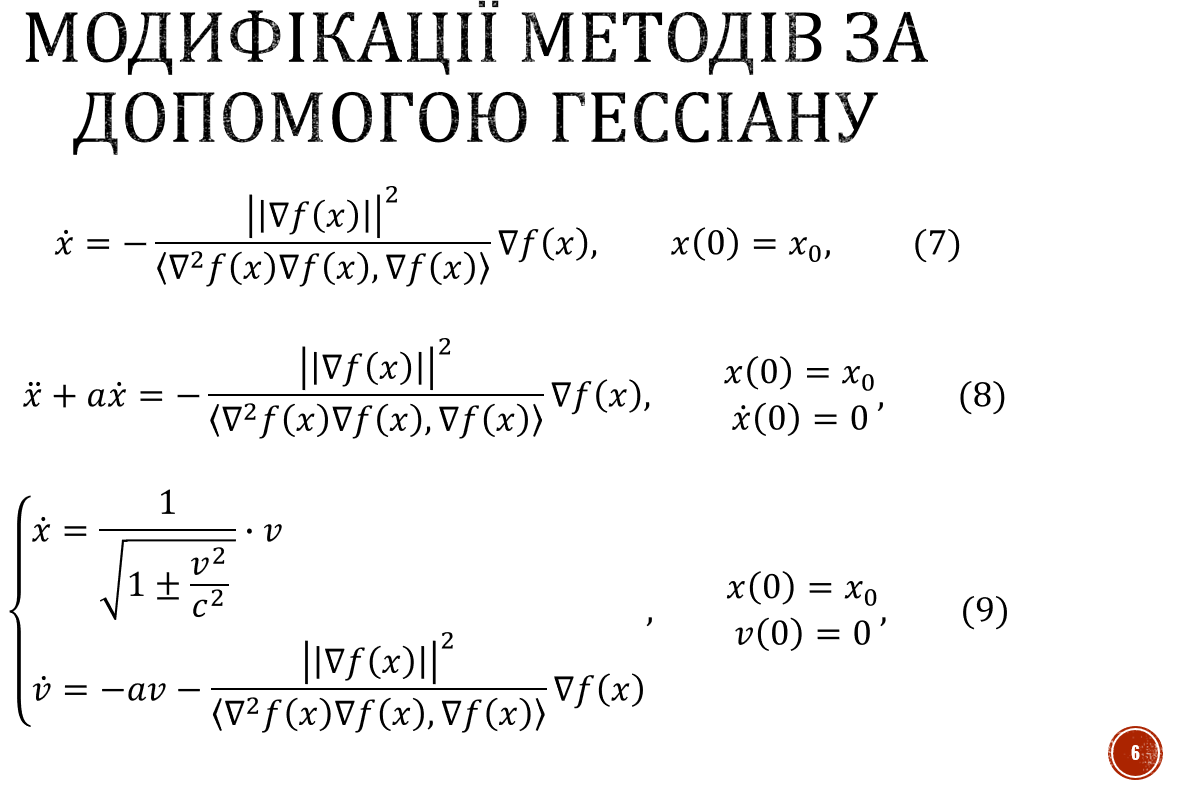
Якщо ввести окрему змінну для першої похідної , то ми отримаємо задачу Коші для системи з двох рівнянь першого порядку , до якої вже можна застосовувати наближені методи знаходження розв’язку.

1. ****

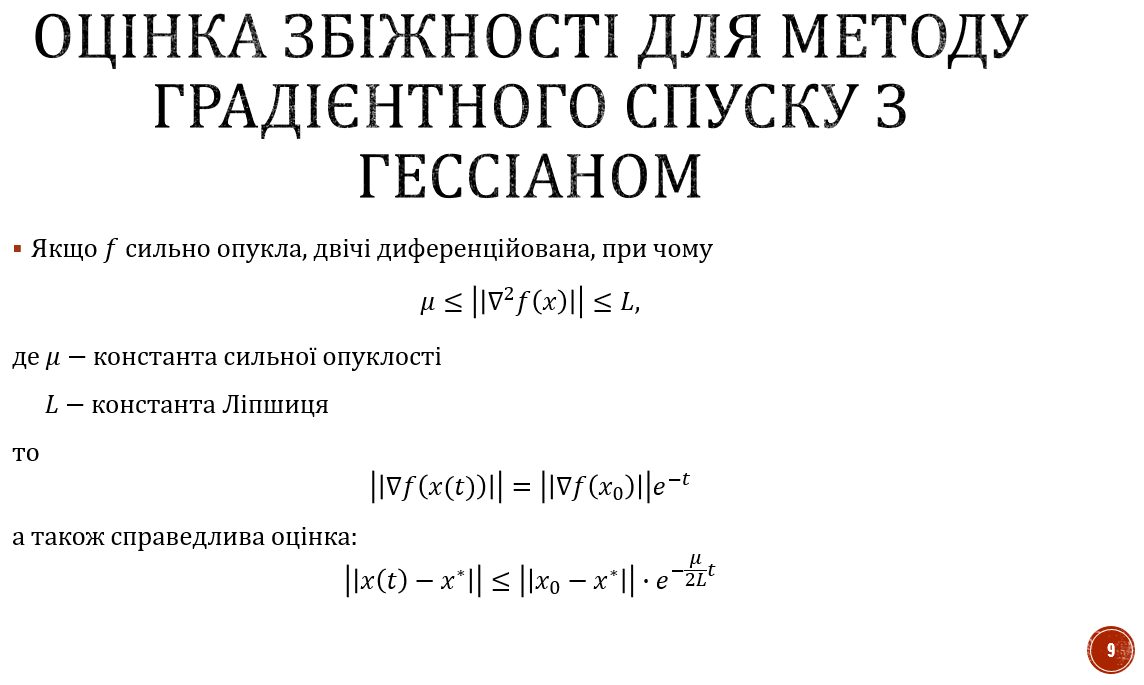
Як і у випадку градієнтного спуску для методу важкої кульки може бути показано, що для сильно опуклих функцій метод збігається до точки мінімуму з експоненційною швидкістю. (як і для градієнтного методу при певних обмеженнях на цільову функцію)

1. 

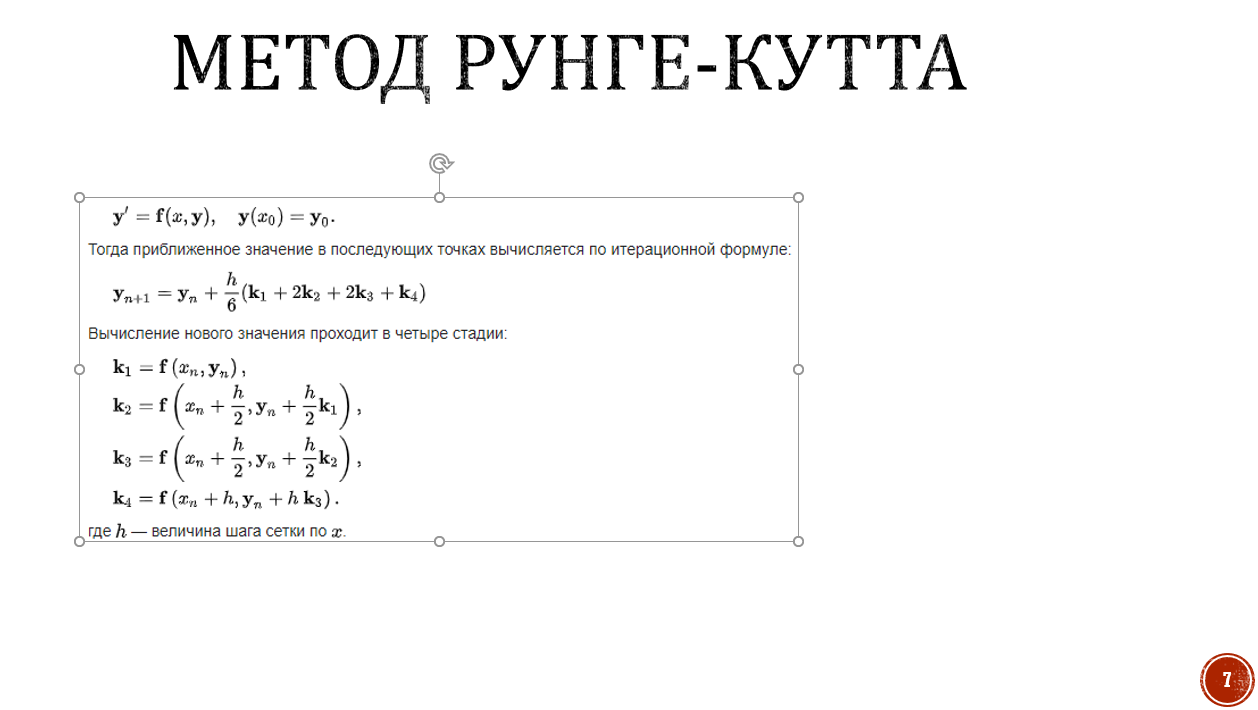
В якості завдання для кваліфікаційної роботи мені було запропоновано дослідження методу релятивістської важкої кульки. Як і сам метод важкої кульки, ідея методу релятивістської важкої кульки запозичена з механіки, а саме з релятивістської теорії. У вираз для швидкості додається фактор Лоренца, як це зазвичай робиться у теорії відносності з імпульсом.

1. 

Для порівняння ефективності методів, в якості референсного використовується разом зі звичайним градієнтним спуском також більше ефективна його модифікація за допомогою матриці Гессе. Витоки цього методу губляться десь у статтях та семінарах сучасної теорії оптимізації, але як показали обчислювальні експерименти – практично метод є досить ефективним.

1. ****

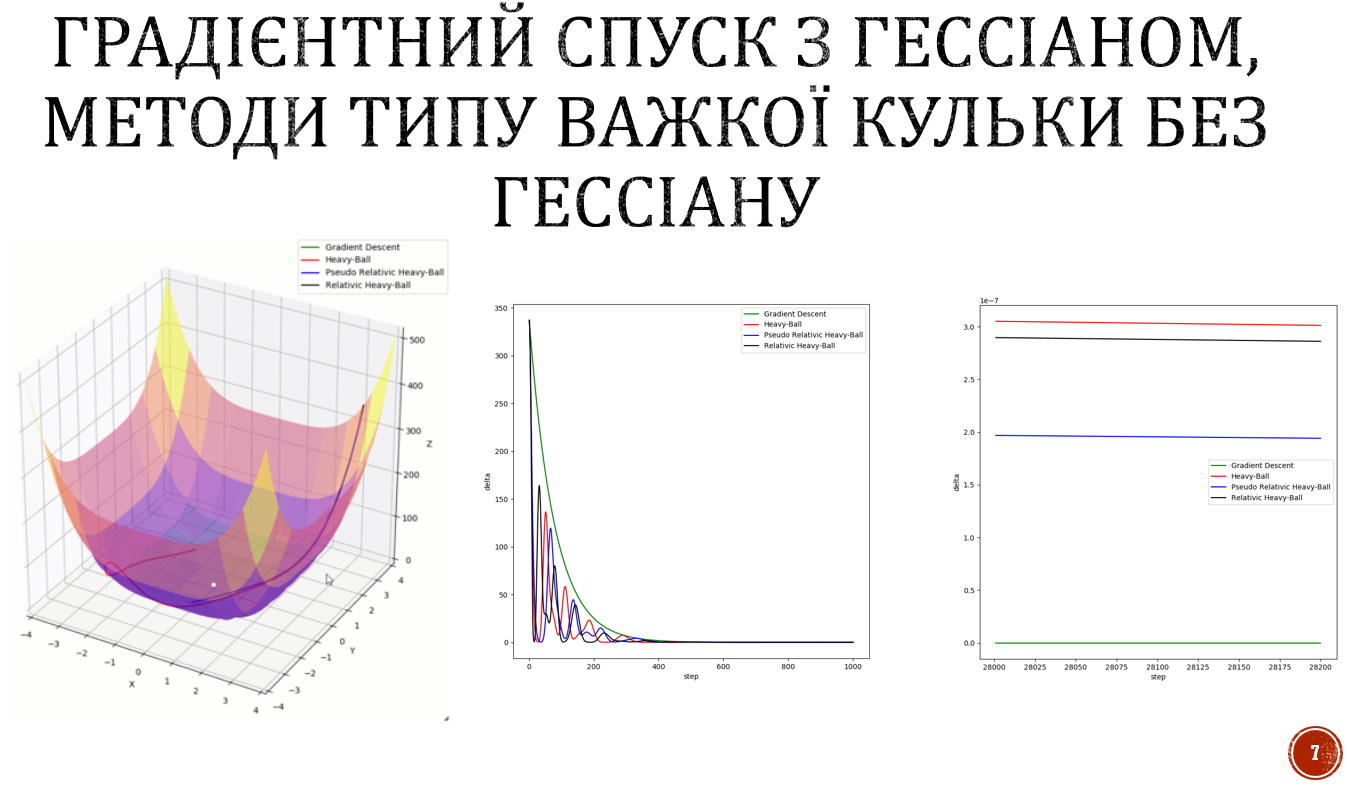
Аналогічно методам, що були розглянуті раніше, у цьому випадку для сильно опуклих функцій може бути показана експоненційна швидкість збіжності до точки мінімуму.

1. 

З точки зору практичної реалізації всі неперервні методи, як градієнтні так і важкої кульки, є задачами Коші для системи диференційних рівнянь.

Для такої задачі Коші наближений розв’язок може бути знайдений методом Рунге-Кутта. В цьому методі значення функції розв’язку в наступних точках рахується за ітераційною формулою. Знаходження нового значення проходить в 4 стадії. Тут величина кроку сітки по .

Цей метод має четвертий порядок точності. Це означає, що помилка на одному кроці має порядок , а сумарна помилка на кінцевому інтервалі інтегрування має порядок .

1. ****

Умовно експерименти які я проводила можна розділити на три частини:

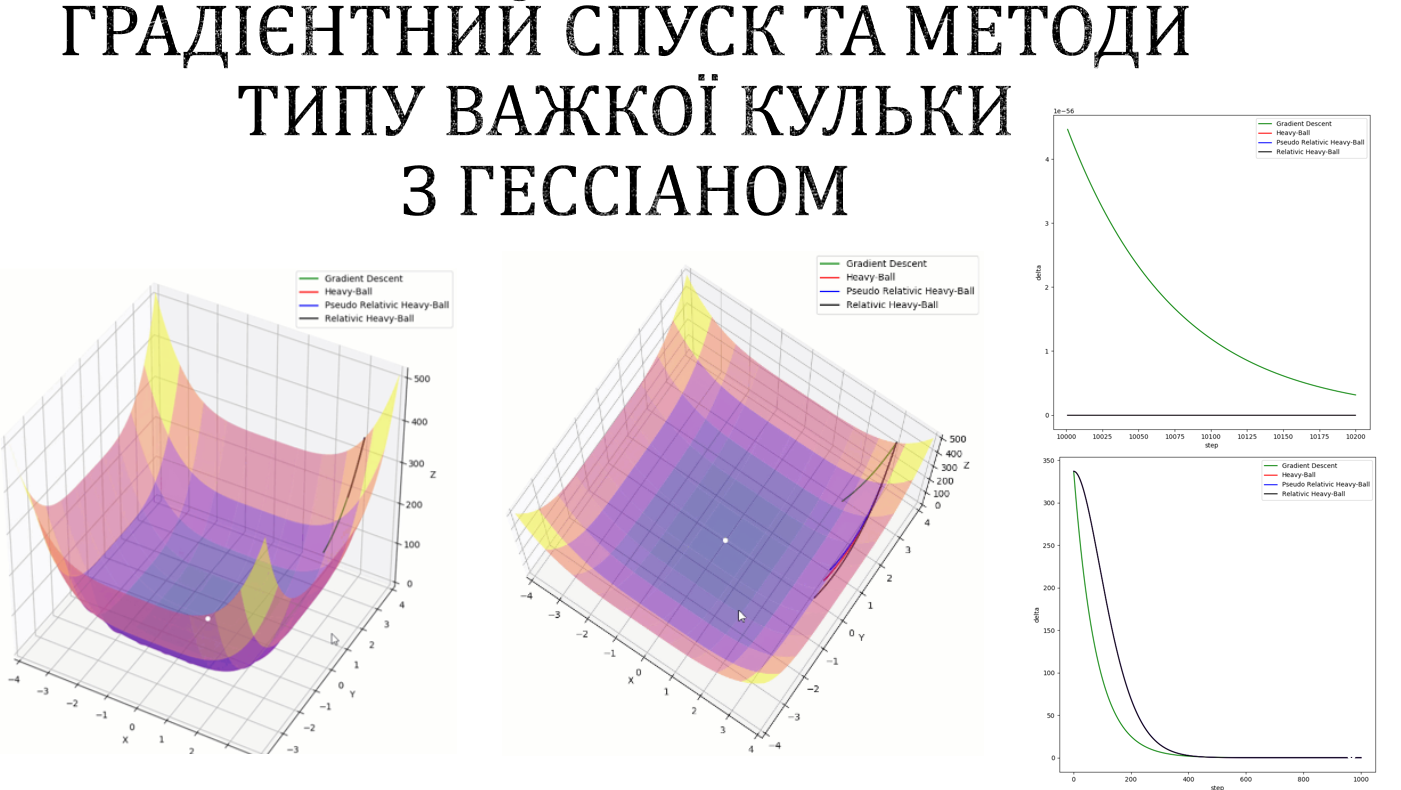
1. Порівняння методів важкої кульки з градієнтним методом без застосування Гессіану
2. Порівняння градієнтного методу з Гессіаном із методами важкої кульки без Гессіану
3. А також порівняння методів важкої кульки з градієнтним методом із Гессіаном

Для порівняння були обрані наступні модельні функції: квадратичні форми, сума четвертих ступенів і логістична регресія.

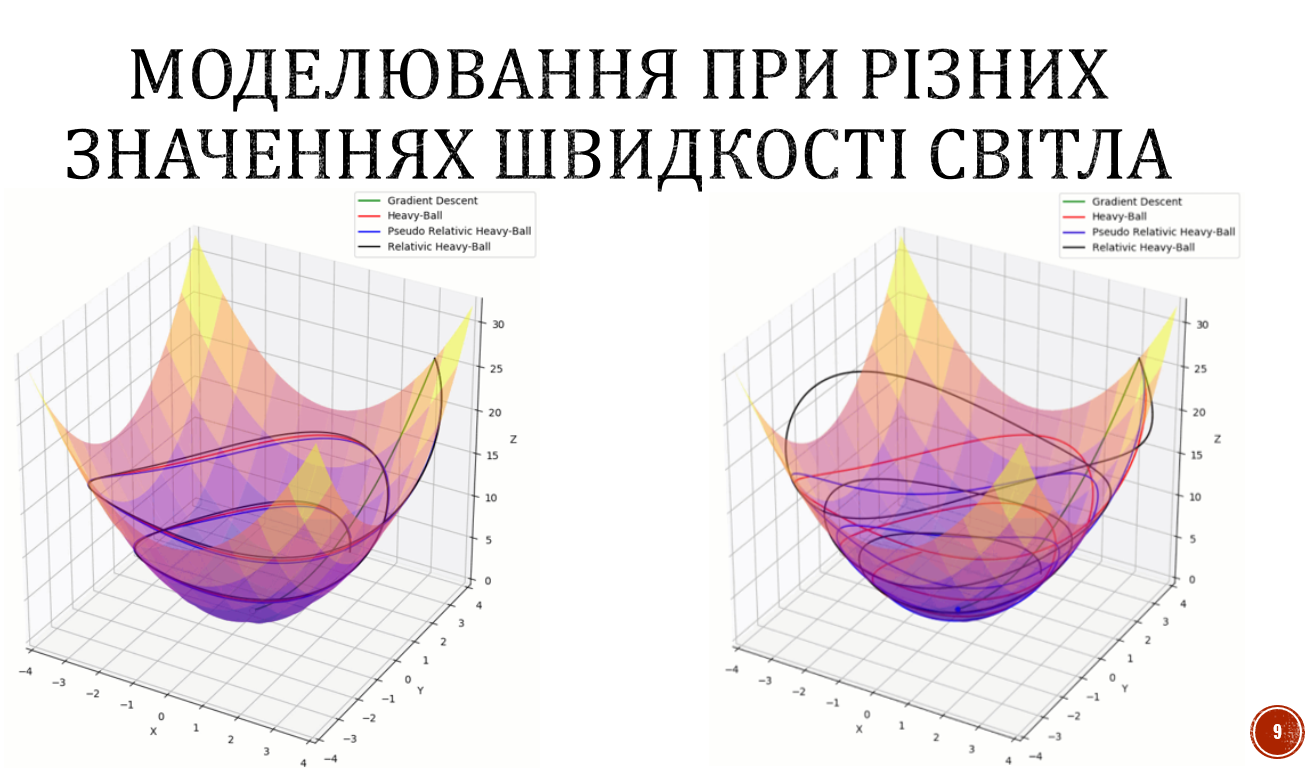
У випадку квадратичних форм серед іншого розглядався варіант з великою різницею коефіцієнтів за напрямком (графік з вузькою долиною).

На даних графіках видно, що градієнтний спуск модифікований Гессіаном виявляється швидшим ніж методи важкої кульки без такої модифікації.

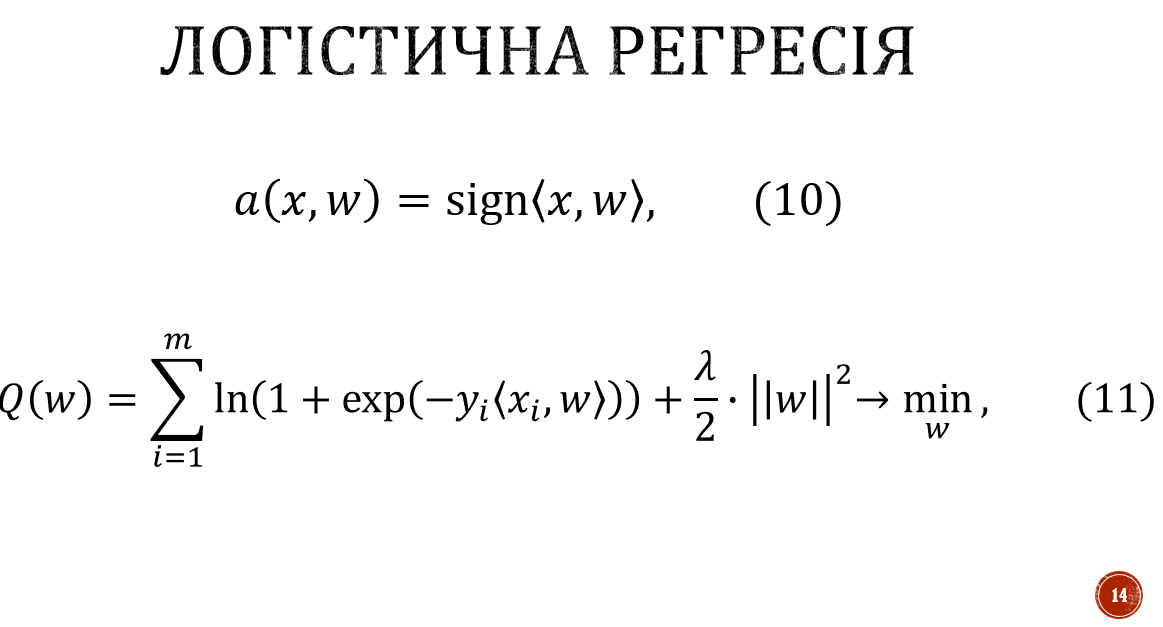
На першому графіку зображені траєкторії наближення до мінімуму відповідними методами. На двох наступних графіках зображена залежність значення цільової функції від номеру ітерації.

1. ****

У разі коли модифікація гессіаном застосовується і до методів важкої кульки ситуація змінюється на протилежну – методи важкої кульки стають ефективніші.

1. ****

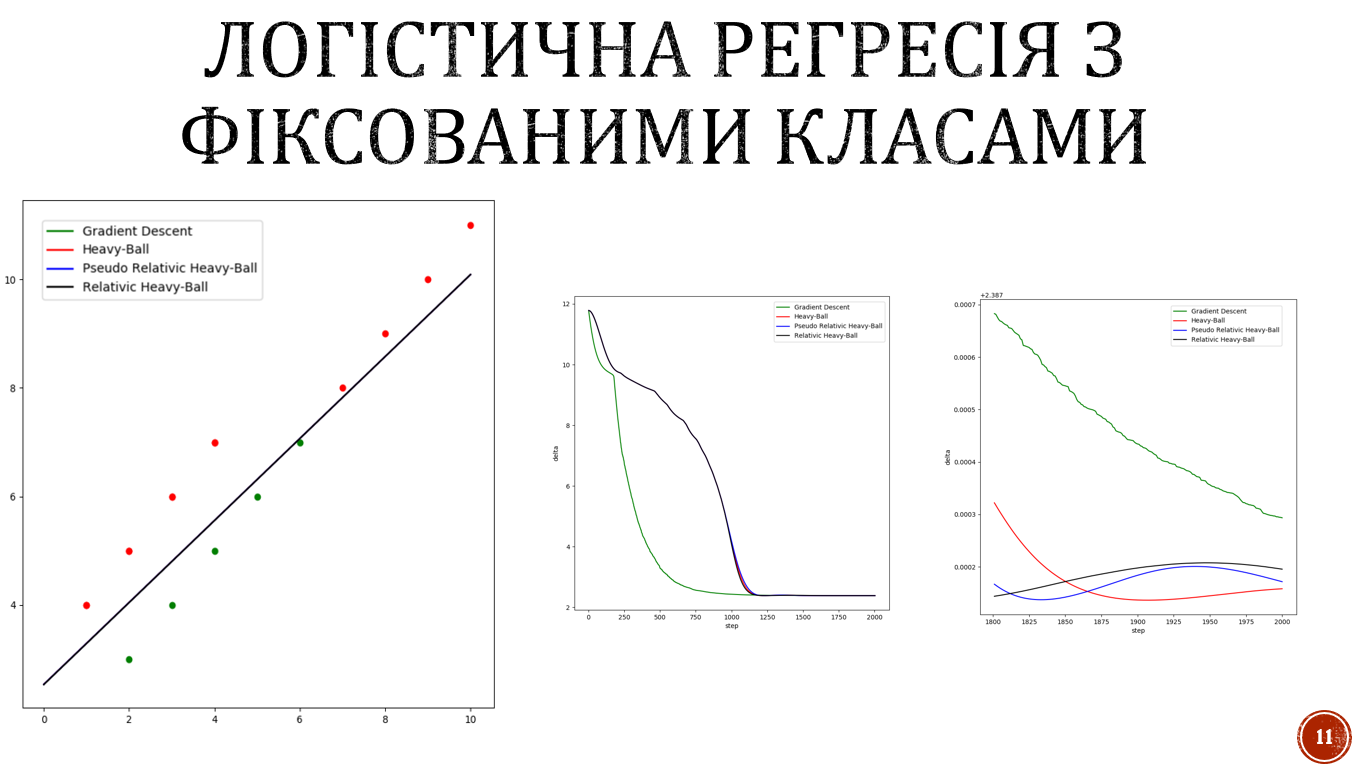
Природнім чином для релятивістських методів важливим параметром є швидкість світла. Якщо швидкість світла велика, то фактор Лоренца близький до одиниці і траєкторії методів важкої кульки та релятивістської важкої кульки співпадають між собою. У випадку коли швидкість світла обрана невеликою різниця між методами стає значною. Релятивістський варіант може на початку процесу значно швидше наближатись до мінімуму. Але для варіанту з традиційним фактором Лоренца звісно слід обирати швидкість світла такою щоб під коренем не було від’ємних величин.

1. ****

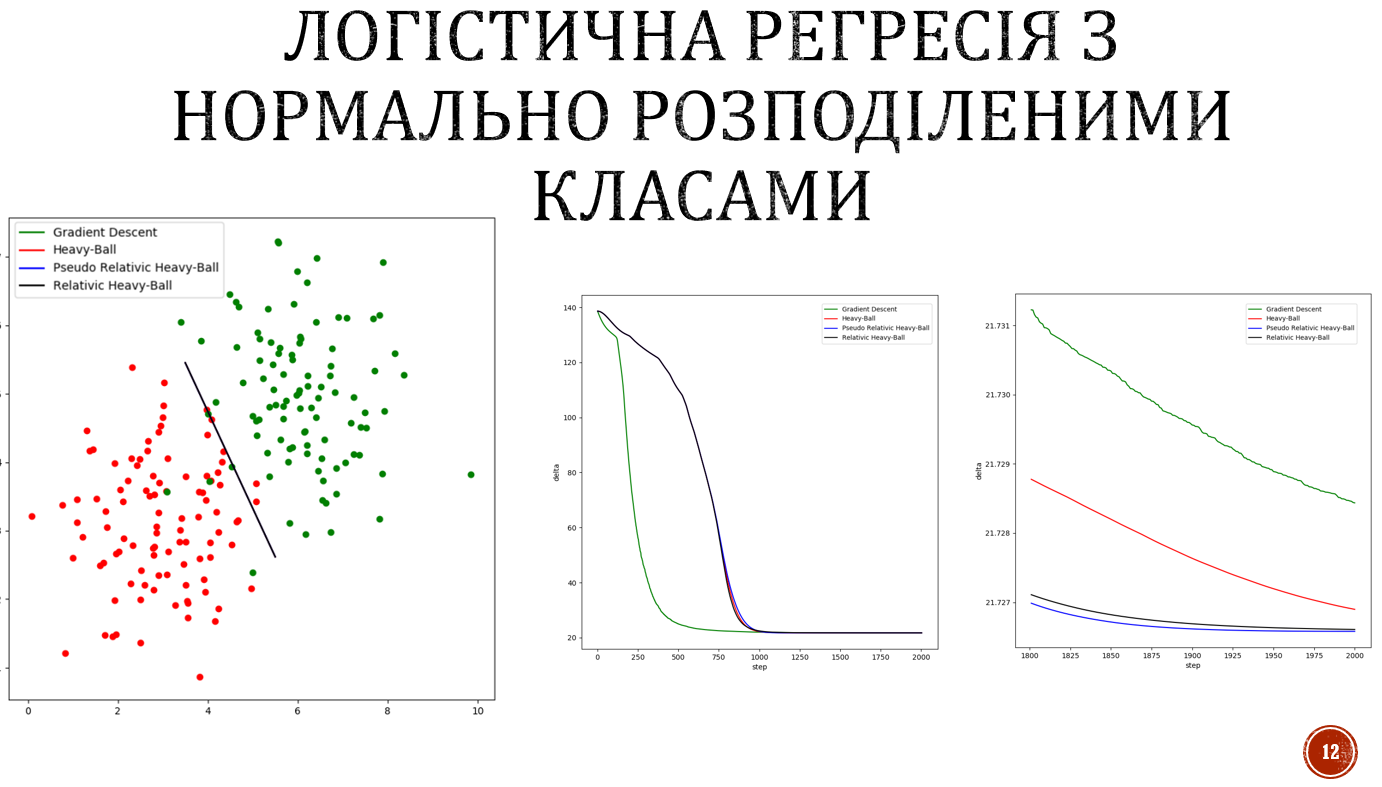
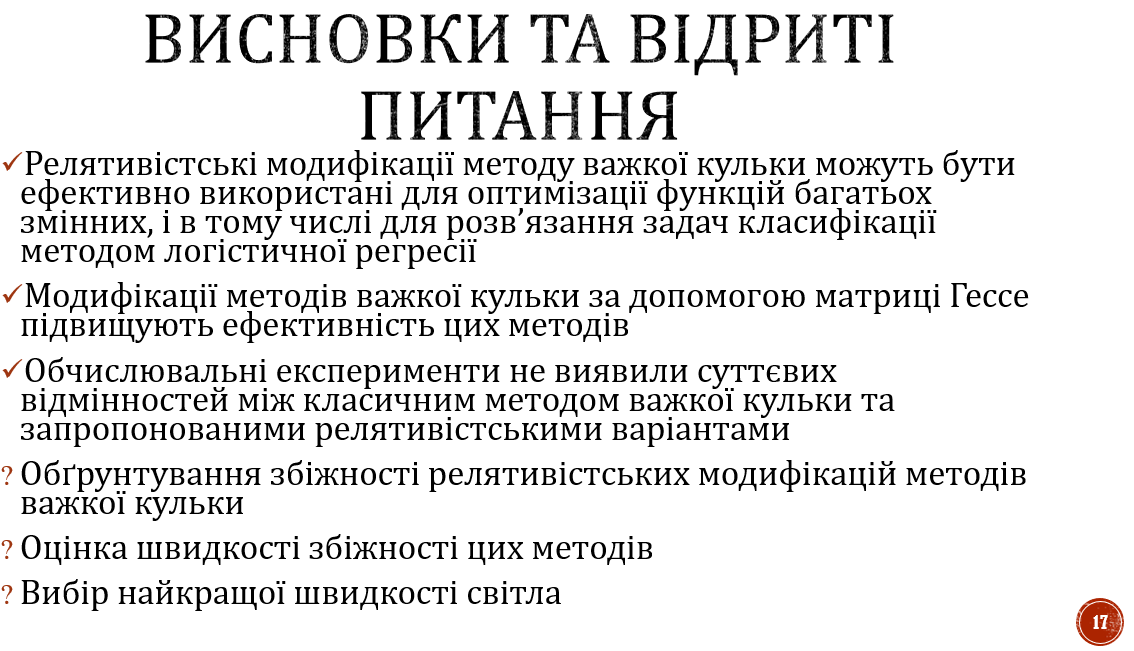
Цікавим і важливим з практичної точки зору є застосування методів оптимізації до задачі класифікації з використанням логістичної регресії.

У цьому випадку задача класифікації полягає в пошуку лінійної розділяючої функції, яка описується своїм вектором коефіцієнтів - формула . дає чи в залежності від знаку скалярного добутку вектору признаків на вектор розділяючої поверхні.

Коефіцієнти знаходяться з умови мінімальності емпіричного ризику, тобто середньої величини похибки алгоритму на навчальній вибірці. Ця умова приводить до необхідності мінімізації функції втрат спеціального вигляду. В цій формулі є додатковий член пропорційний нормі вектора з невідомих коефіцієнтів, який вводиться до цільової функції з метою обмеження величин цих коефіцієнтів. Такий засіб називається регуляризацією. В реальних задачах без застосування регуляризації при пошуку розв’язку коефіцієнти необмежено зростають, при цьому цільова функція може лише незначно зменшуватись. Регуляризація утримує розв’язок в певних межах, які залежать від вибору коефіцієнту регуляризації , хоча при цьому значення цільової функції стає дещо більшим від реальної нижньої межі. Вибір коефіцієнту це компроміс між глибиною мінімізації та утримання компонентів розв’язку в розумних межах. ( як правило обирається меньше 1)

1. ****

Методи важкої кульки практично дуже швидко знаходять коефіцієнти розділяючої лінійної функції. Як у випадку простих синтетичних прикладів так і у випадку вибірок з нормальними розподілами.

1. ****
2. ****

Значення з достатньою точністю релятивістські методи досягають швидко, а шукати надто велику точність не має сенсу, так як рішення зміщене регуляризацією.