

Алгебро-алгоритмічні засоби побудови штучних нейронних мереж та застосування їх до розв'язку початково-крайових задач

Юрій Жовнір

Кафедра інформаційних систем та мереж, Національний університет “Львівська політехніка”, УКРАЇНА,
м.Львів, вул.С.Бандери, 12, E-mail:zhovnir@gmail.com

Abstract – One of methods of computer design is investigated in this work, namely artificial neuron networks (ANN). An ordinary multinet method was taken that was improved by ANN, as the result in few times more rapid and exact method was received in compare to multinet. For an example was built artificial neuron network that had been used for the decision of boundary problems of mathematical physics, and winning was obtained in 5-10 times in compare to ordinary mathematical methods.

Key words: artificial neuron networks, multinet metho, boundary problems of mathematical physics

I. Вступ

Поведінка багатьох об'єктів реального світу описується настільки складними математичними моделями, що їх аналіз класичними чисельними методами недоступний навіть за умов швидкого зростання потужностей сучасної обчислюваної техніки. Це призводить до пошуку альтернативних шляхів розв'язування таких задач, одним з яких є застосування комп'ютерних моделей типу нейронних мереж.

Успіхи штучних нейромереж у моделюванні складних природних процесів наводять на думку їх застосування до математичних моделей. Візьмемо для прикладу звичайний мультисітковий метод і подивимося на нього зі сторони нейромереж, представимо, що вузол сітки – це нейрон, згладжування у кожному вузлі – суматор, точність ϵ – функція активації. Це наводить на думку, що ШНМ є також математичною моделлю, і якщо додаємо до чисельного методу потужність нейромережі, отримуємо новий досконаліший механізм для розв'язання математичних задач, що й використовується в даній роботі. Тут було взято звичайний мультисітковий метод, удосконалено його нейромережею і як наслідок отримано в декілька раз швидший та точніший метод в порівнянні з мультисітковим. Методу складається з трьох етапів.

На першому етапі для розв'язування крайових задач на ШНМ застосовуємо дискретизацію, яка складається з заміни області змінної неперервного аргумента на дискретну множину точок (сітку), і заміну диференціального рівняння – різницевим рівнянням.

Далі реалізуємо механізм мультисіткового ітераційного методу ШНМ. Він складається з трьох шарів, де кожен з них реалізує відповідний етап ітерації. Перший шар нейронів забезпечує обчислення нев'язки, значення якої використовується на другому шарі з метою визначення умов завершення алгоритму. Якщо умови завершення виконуються, то стан вихідного вектора ШНМ відповідає розв'язку рівняння. В протилежному випадку вектор нев'язки передається до третього етапу для формування нового наближення вектора похибки.

На третьому етапі реалізовується нейронна підмережа прямого поширення, результатом якої є апроксимоване значення похибки. Завдяки тому, що на цьому етапі обчислення проводяться на сітці в 2,

4 рази менший, отримуємо значний вигравш в швидкодії.

В даній роботі використано штучну нейрону мережу для розв'язання крайові задачі математичної фізики. Звичайний мультисітковий метод модифікується за допомогою нейромережі прямого поширення. Вона дозволяє при менших затратах ресурсів отримати апроксимоване значення похибки. В роботі було показано доцільність такої модифікації. Мультисітковий метод представимо у вигляді рекурентної нейромережі, яка на вхід отримує початкове значення та апроксимовану похибку, сумує ці значення та перевіряє на збіжність, в чому і полягає активаційна роль нейрона, відповідно чи передати на наступний крок, чи завершити процес

Метод розв'язання за допомогою ШНМ

Розглянемо систему різницевих рівнянь

$$L_h u_h(x) = f_h(x), x \in G$$

Яка є апроксимацією крайової задачі $Lu=f$ на сітці

$$G = \omega \cup \gamma, \text{де } \omega_h = \{x_i = ih, i = 1, \dots, N-1, h = 1/N\}$$

— множина внутрішніх вузлів;

$$\gamma_h = \{x_i = ih, i = 0, N, h = 1/N\} — множина$$

граничних вузлів.

Нехай $u_h(n)$ - апроксимація розв'язку uh , який відповідає ітераційному кроці n . Тоді $v_h(n)$ називають помилкою такої апроксимації за умови:

$$v_h(n) = u_h - u_h(n)$$

а нев'язку задають у вигляді різниці

$$d_h(n) = f_h - L_h u_h(n)$$

Використовуючи дві попередні формули запишемо рівняння нев'язки: $L_h v_h(n) = d_h(n)$

Розглянемо можливість апроксимації даної задачі на множині сіток $G^{(0)}, G^{(1)}, G^{(2)}, \dots, G^{(M)}$, починаючи з деякої базової сітки $G(0)$, яка характеризується найменшим кроком дискретизації h_0 . Величину кроку дискретизації для довільної сітки $G(i)$ можна записати: $h_i = k^{(i)} h_{i-1}$

Де $k(i)$ – коефіцієнт зростання кроку дискретизації. У більшості випадків $k(i) = k = 2$

Перехід від сітки $G(i)$ до $G(i+1)$ відбувається за допомогою лінійного оператора рестрикції $R_i^{i+1} : H_i \rightarrow H_{i+1}$, зворотній перехід задає лінійний оператор пролонгації: $P_{i+1}^i : H_{i+1} \rightarrow H_i$

Для параметрів $d(i)(n)$ та $v(i)(n)$ що відповідають значенням нев'язки та похибки у вузлах сітки $G(i)$, оператор R_i^{i+1} задає перехід: $d^{(i+1)}(n) = R_i^{i+1}d^{(i)}(n)$, а оператор $P_{i+1}^i : V^{(i)}(n) = P_{i+1}^iV^{(i+1)}(n)$

Позначимо через $S[u^{(0)}(n), L^{(0)}, f^{(0)}]$ ітераційну схему на сітці $G(0)$ а через $S[v^{(i)}(n), L^{(i)}, d^{(i)}(n)]$ — ітераційну схему на довільній грубій сітці $G(i) > 0$, на ітераційному кроці ρ . Тоді наступна послідовність операцій складатиме алгоритм реалізації згаданого методу.

1. Згладжування на сітці $G(0)$: $u^{(0)}(n) = S[u^{(0)}(n), L^{(0)}, f^{(0)}]$
2. Обчислення нев'язки на точній сітці: $d^{(0)}(n) = f^{(0)} - L^{(0)}u^{(0)}(n)$
3. Операція рестрикції на грубу сітку: $d^{(1)}(n) = R_0^1d^{(0)}(n)$
4. Згладжування похибки на сітці $G(1)$: $V^{(1)}(n) = S[v^{(1)}(n), L^{(1)}, d^{(1)}(n)]$
5. Пролонгація похибки: $V^{(0)}(n) = P_1^0V^{(1)}(n)$
6. Обчислення нової апроксимації: $u^{(0)}(n+1) = u^{(0)}(n) + V^{(0)}(n)$
7. Повторення пунктів 1-6 до досягнення збіжності.

Цей підхід став базовим при розв'язаннях крайових задач.

Архітектура нейронної мережі складається з кліткової двошарової субмережі та субмережі прямого поширення. Кожен із шарів кліткової субмережі виконує специфічні обчислення, що зумовлює їх різну структуру. Нехай нейронна мережа забезпечує розв'язування системи різницевих рівнянь $L_h u = f_h$. Тоді на першому шарі кліткової субмережі виконують обчислення мультисіткової корекції за формулою:

$$u_h(n) = W u_h(n) + \Delta u(n)$$

Він складається з операційного блоку Σ/Π та блоку активаційної функції ϕ . Призначення операційного блоку полягає в обчисленні корекції елемента u_j , вектора $u = (u_0, \dots, u_j, \dots, u_m)$ для поточного ітераційного кроku ρ . Активаційна функція ϕ для нейронів першого і другого шарів кліткової нейронної субмережі є кусково лінійною що характерно для обчислювальних нейронних мереж. Матриця вагових коефіцієнтів W задає коефіцієнти затухання.

Нейрон другого шару кліткової під мережі, забезпечує формування чергової апроксимації вектора розв'язку згаданої системи різницевих рівнянь шляхом реалізації ітераційної схеми :

$$u_h(n+1) = u_h(n) - \rho(L_h u_h(n) - \varphi)$$

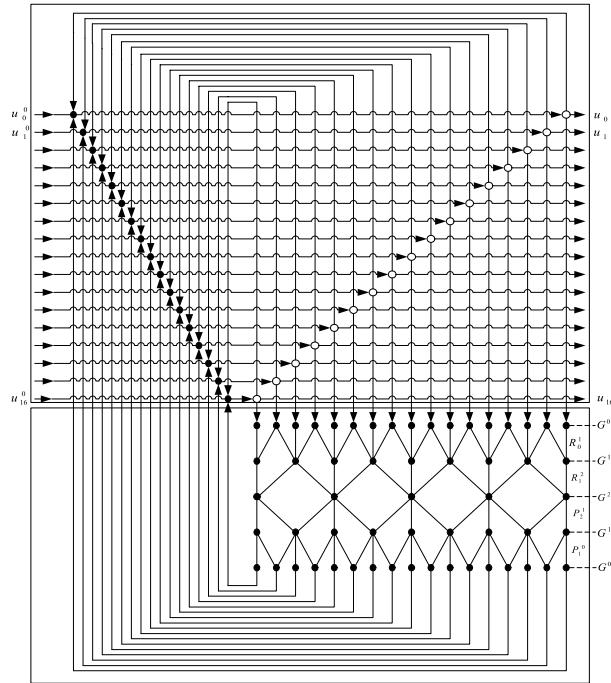


Рис. 1. Архітектура рекурентної ШНМ для розв'язування крайових задач

Умова завершення ітераційного процесу має вигляд: $\frac{|u_h(n+1) - u_h(n)|}{\max |u_h(n)|} < \varepsilon$

Справдження нерівності свідчить про збіжність ітераційної послідовності $\{u_h(n)\}$ до розв'язку. Для підвищення швидкості збіжності незалежно від величини кроків дискретизації застосовують мультисіткові методи. В ШНМ, архітектура якої показана на рис.1, для досягнення згаданих цілей використано нейронну під мережу прямого поширення.

Для реалізації процесу мультисіткового згладжування застосовують $2N+1$ шарів нейронів, оскільки обчислення на кожній із сіток, за винятком сітки $G(N)$, виконують два шари нейронів. Загальна кількість сіток, які застосовуються для реалізації мультисіткового згладжування, знаходитьться в діапазоні $1 < i < \log_2 t$, де t — кількість елементів вектора u_h , і залежить від вибору циклу мультисіткового методу та виду операторів пролонгації та рестрикції.

Для одновимірної нейронної субмережі, яка зображена на рис.1, використані одновимірні сітки $G(i)$ та $G(i+1)$ зі співвідношенням кроків дискретизації

Результат дії оператора пролонгації

$$P_{i+1}^i : V^{(i+1)} \rightarrow V^{(i)}, V^{(i)} \in R^{n_i}, V^{(i+1)} \in R^{m_{i+1}}$$

$$m_i = 2m_{i+1} + 1$$

визначають за формулою: $V_h^{(i)} = P_{i+1}^i V_h^{(i)}$

Результат дії оператора рестрикції

$$R_i^{i+1} : d^{(i)} \rightarrow d^{(i+1)}, d^{(i)} \in R^{m_i}, d^{(i+1)} \in R^{m_{i+1}}$$

$$m_i = 2m_{i+1} + 1$$

визначають за формулою: $d^{(i+1)} = R_i^{i+1} d^{(i)}$

На першому шарі згладжування виконують обчислення значень нев'язки $d_h^0 : d_h^0 = f - L_h u_h^0$

Передача одержаного значення нев'язки на наступний шар нейронів, ототожнюваний з більш грубою сіткою, відбувається з використанням операторів рестрикції, які виконують роль вагових коефіцієнтів для верхньої половини мультисіткової субмережі. На грубих сітках розв'язують рівняння нев'язки:

$$L_h^{(i)} V_h^{(i)} = d_h^{(i)}$$

Для цього можна застосувати ітераційну схему

$$V_h^{(i)}(n+1) = V_h^{(i)}(n) - \rho(L_h^{(i)} V_h^{(i)} - d_h^{(i)})$$

у випадку використання цифрової моделі нейрона або ж застосувати прямі методи обчислення при використанні неперервних моделей.

Одержані значення похибки передають до наступного шару, і застосовують оператори пролонгації, які визначаються як вагові коефіцієнти для нижньої половини мультисіткової субмережі. Результатом роботи мультисіткової субмережі є апроксимоване значення похибки, яке відповідає поточній макроітерації, визначуваною клітковою субмережею. Процес корекції поточного значення вектора розв'язку відбувається на першому шарі кліткової субмережі. Загальний алгоритм роботи нейронної мультисіткової субмережі може бути представлений у вигляді рекурсивної процедури.

Ітераційний процес на кожному з нейронних шарів складається з внутрішніх нейронних ітераційних субпроцесів і виконує функцію згладжування після операцій пролонгації та рестрикції. Навчання нейронної мережі полягає у модифікації матриці вагових коефіцієнтів, яка задає дію цих операторів. Основною метою навчання є досягнення максимальної швидкості збіжності в результаті ефективних операцій корекції.

III. АПРОБАЦІЯ

Розглянемо числову реалізацію, даного методу для тестового прикладу.

$$\frac{d^2u}{dx^2} = f(x), 0 \leq x \leq 1$$

$$u(0) = 0, u(1) = -7,148764$$

де $f(x) = -225000 \sin(150x)$

Перехід до різницевої крайової задачі виконуємо в два етапи:

Замінимо область неперервної зміни аргументу областю його дискретної зміни.

Замінимо диференціальний оператор деяким різницевим оператором і сформулюємо різницевий аналог для граничної умови та початкових даних.

Результатом цієї процедури буде перехід до алгебраїчної системи рівнянь, яка може бути розв'язана за допомогою асинхронного мультисіткового методу.

У межах реалізації першого етапу введемо множину вкладених сіток:

$\bar{G}^{(k)} = G^{(k)} \cup \Gamma^{(k)}$, де k – номер сітки, Γ – зафікована границя. Для даної задачі використаємо дві сітки.

На другому етапі використовуємо один із методів апроксимації диференціального оператора. Найпростіший підхід полягає в заміні похідних, які входять у L_h , на їх різницеві аналоги, які визначають L_h . Для даної задачі різницевий оператор має вигляд:

$$(L^{(k)} u^{(k)}) = (-2u_i^{(k)} + u_{i+1}^{(k)} + u_{i-1}^{(k)}) / h^{(k)^2}$$

Використовуючи різницевий оператор будемо проводити обчислення значень функції за ітераційною формулою:

$$u^{(k)}(n+1) = u^{(k)}(n) + \rho(L^{(k)} u^{(k)}(n) - F^{(k)}(n)),$$

де $\rho = h^{(k)^2} / 2$

Значення нев'язки визначимо за формулою

$$d_i^{(k)}(n) = F_i^{(k)}(n) - (L^{(k)} u^{(k)}(n))_i$$

та спроектуємо її на більш грубу сітку за допомогою оператора рестрикції:

$$u_j^{(i+1)} = (R_i^{i+1} u^{(i)})_j = \frac{1}{2} \left(\frac{u_{2j-1}^{(i)}}{2} + u_{2j}^{(i)} + \frac{u_{2j+1}^{(i)}}{2} \right)$$

Проміжне згладжування на грубих сітках виконуватимемо за ітераційною формулою:

$$v^{(k)}(n+1) = v^{(k)}(n) + \rho(L^{(k)} V^{(k)}(n) - F^{(k)}(n))$$

де $v(k)$ – поточне значення похибки на сітці $G(k)$. Операцію пролонгації виконуємо за формулою:

$$u_j^{(i)} = \begin{cases} (P_{i+1}^i u^{i+1})_{2j} = u_j^{(i+1)} \\ (P_{i+1}^i u^{i+1})_{2j+1} = \frac{1}{2} (u_j^{(i+1)} + u_{j+1}^{(i+1)}) \end{cases}$$

Корекцію наближеного значення функції на сітці $G(0)$ виконують за формулою:

$$(u_i^{(0)})' = u_i^{(0)} + V_i^{(0)}$$

Отримані розв'язки показані на рис. 2,

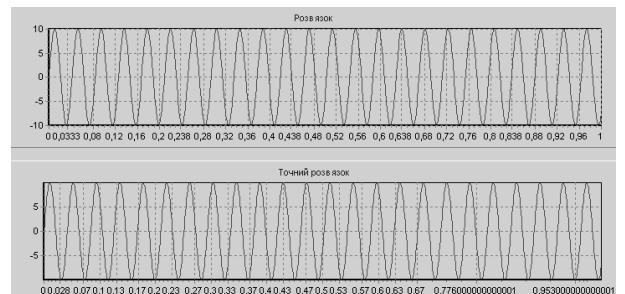


Рис. 2. Розв'язок отриманий ШНМ і точний розв'язок, при $N=451 \epsilon=0,0001$

IV. Аналіз методу

Даний метод є мультисітковим методом з нейромережею прямого поширення, яка повинна дати кращу швидкість збіжності, за рахунок того що на грубих сітках ми обчислюємо значення в $N/2, N/4 \dots 1$ точках. Попробуємо в цьому перевіритися, для цього використаємо наступну тестову задачу

$$\frac{d^2u}{dx^2} = f(x), 0 \leq x \leq 1$$

$$u(0) = 0, u(1) = -0,279415$$

де $f(x) = -36 \sin(6x)$

Візьмемо для даного експерименту фіксовану сітку $N=451$ і фіксоване $\epsilon = 0,0001$. Подивимося на скільки зросте швидкодія при використанні грубих сіток і при різних комбінаціях згладжування на цих сітках. Припустимо що одна ітерація мультисіткового методу займає 1 ум. Од. ресурсів, тоді не зменшуючи загальності одна ітерація на грубій сітці №1 займає $0,5*v1 + 0,5$, де $v1$ – кількість ітерацій згладжування на сітці №1 (так як кількість елементів в два рази менша, і операції пролонгації та рестрикції, менш ресурсоемкі ніж згладжування). Аналогічно на сітці №2 займає $0,25*v2 + 0,25$ де $v2$ – кількість ітерацій згладжування на сітці №2.

Виходячи з цих міркувань, попробуємо звести наші досліди в таблицю

№	$v1$	$v2$	Кількість ітерацій	Умовні ресурси
1	0	0	5000	5000
2	1	0	1736	3472
3	1	1	675	1687
4	2	2	401	1303
5	3	3	293	1172
6	3	0	844	2532
7	3	1	676	2366
8	5	5	195	1072
9	10	10	111	1026
10	12	12	95	1021

У випадку №1 використовується звичайний мультисітковий метод, орієнтуючись на це як за основу, ми бачимо що нейромережа прямого поширення, дас дуже хороший приріст в швидкодії у випадку №3 коли використовується лише одна ітерація згладжування на кожній грубій сітці в 3 рази, а у випадках №8-10 майже в 5 раз. З цього можна зробити такі висновки:

Використання нейромережі прямого поширення покращує мультисітковий метод у декілька раз, навіть при одній грубій сітці, не вимагаючи при цьому значних труднощів у реалізації;

Варто використовувати більше 2 шарів нейронів та кількість згладжувань більше 5;

На практиці для випадку №1 було використано близько 10с. для №10 менше 1 с.

Проаналізуємо як залежить розв'язок від точності ϵ . Нагадаємо, що

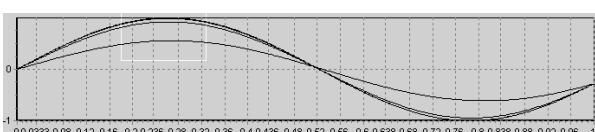
$$\frac{|u_h(n+1) - u_h(n)|}{\max |u_h(n)|} < \epsilon$$

Використаємо аналогічну тестову задачу, з фіксованою сіткою $N=451$, та фіксованими $v1 = 2$ та $v2 = 2$, максимальною кількістю ітерацій $N_{\max}=5000$.

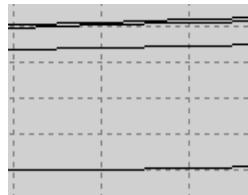
Виберемо для аналізу наступні значення ϵ :

$E = 0,01$, $E = 0,0001$, $E = 0,00001$, $E = 0,0000001$

Отримали наступні графіки розв'язку:



зближеному варіанті:



З графіків можна побачити, що розв'язок кардинально відрізняється в залежності від E , отже можна зробити висновок, є потрібно брати як умога менше: $0,00001$ і менше. Але ніякому разі більше від $0,00001$. І як показали дослідження оптимальним значенням для точності ϵ $0,000001$

Розглянемо рівняння

$$1,3 \frac{d^2u}{dx^2} + \frac{du}{dx} - 1,155u = 0,0 \leq x \leq 1$$

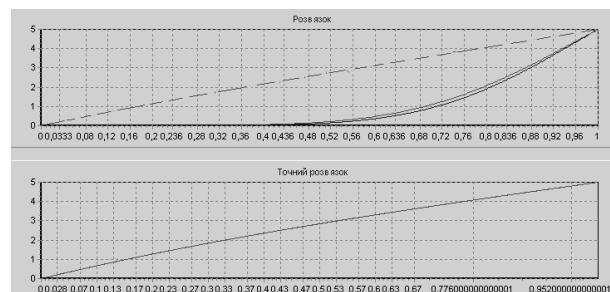
$$u(0) = 0, u(1) = 5$$

Для цього рівняння існує точний розв'язок

$$u = \frac{5}{e^{\frac{5}{13}} - e^{-\frac{15}{13}}} (e^{\frac{5}{13}x} - e^{-\frac{15}{13}x})$$

Це рівняння є більш складнішим для розв'язання чисельними методами, розв'язжемо його звичайним мультисітковим методом і нашим модифікованим нейромережею прямого поширення, з двома грубими сітками і коефіцієнтами $v1$ та $v2$ рівним 15. Зафіксуємо наступні значення $N=451$, $N_{\max} = 3000$, $\epsilon = 0,00001$.

Графіки роботи програми наведені на рисунку нижче:



Штрих-пунктиром на графіку позначений розв'язок отриманий ШНМ, суцільною лінією – звичайним мультисітковим при різних значеннях похибки $0,00001$ та $0,000001$, як видно покращення методу нейромережею прямого поширення дас дуже відчутний результат. Отримано досить точний результат, на відмінно від звичайного мультисіткового методу, і крім цього використавши приблизно в 5 разів менше ресурсів.

Висновки

В даній роботі розв'язано крайову задачу за допомогою штучних нейронних мереж. Наведені дослідження показують значний вигравш в швидкодії при використанні ШНМ над звичайним мультисітковим методом. Було продемонстровано, що рекурсивна нейромережа представляє собою ітераційні методи і крім цього удосконалює їх, що дає велику перевагу.

За допомогою тестових задач переконалися, що ШНМ дає виграш в 5-10 разів перед звичайним мультисітковим методом. Це є дуже суттєво. Йому під силу розв'язувати рівняння, для яких звичайний мультисітковий метод не може дати точного розв'язку. Разом з цим отримано результати, які показали, що метод дуже чутливий до точності. Це пов'язано з тим, що наступне апроксимоване значення залежить від попереднього. Відповідно зроблено висновок використовувати точність 0,00001, або оптимальніше 0,000001.

Було виявлено доцільність використання саме рекурентних ШНМ для такого типу задач, так як даний вид ШНМ дозволяє найпростішим способом застосувати ітераційні методи. Рекурентні ШНМ здатні до навчання без вчителя, а отже не потребують множини розв'язків для навчання, на відміну від ШНМ прямого поширення. В даній роботі показано, що доцільно застосовувати нейронну підмежеву прямого поширення

для підвищення швидкості збіжності, разом з цим отримано результати, згідно з якими доцільно використовувати більше 2 сіток та кількістю ітерацій згладжування на грубих сітках більше 5.

References

- [1] Трушевський, Н.Щербина «Розв'язання лінійних крайових задач мультисітковим ітераційним нейронним методом». Вісник Львів УН-ТУ Сер. Прикл. Мат. Та інформатики 2006 Вип. 11 С 82-98
- [2] М.А. Новотарський, Б.Б. Нестеренко Штучні нейронні мережі: обчислення // Праці Інституту математики НАН України. - Т51. -Київ: Ін-т математики НАН України. 2004. - 408 с.
- [3] Глушков В.М., Цейтлін Г.Е., Ющенко Е.Л. «Алгебра. Языки. Программирование» Київ 1978
- [4] Цейтлін Г.О. «Теорія клонів та її додатки», Київ 2003

Дослідження впливу металевого підкладу на потенціал міжелектронної взаємодії в плівці

Іван Задворняк

Кафедра прикладної математики, Національний університет “Львівська політехніка”, УКРАЇНА,
м.Львів, вул.С.Бандери, 12, E-mail: ivanZadv@yahoo.com

Problem of influence of metallic surface on electron-electron interaction in semiinfinity jellium is considered. Numerical and analytic results of calculation of effective potential are obtain.

Ключові слова – electron-electron interaction, effective potential.

I. Вступ

Досліджено вплив металевого підкладу на потенціал взаємодії між електронами в плівці. Плівка нанесена на поверхню підкладу, що моделюється напівобмеженим «желе». Обчислено ефективний і кулонівський потенціали міжелектронної взаємодії в плівці. За допомогою чисельних розрахунків отримано протабульовані значення потенціалу міжелектронної взаємодії по нормальніх до площини поділу координатах електронів z_1, z_2 та віддалі між електронами в площині поділу R_{\parallel} .

II. Моделювання потенціалу міжелектронної взаємодії

Розглядається система, яка містить N електронів, що знаходяться в плівці об'ємом $V = Sd$, де d – товщина плівки, S – площа поверхні плівки. Площу поверхні плівки вважаємо макроскопічно великою $(\frac{S}{\langle r \rangle^2} \gg 1)$, $\langle r \rangle$ – середня віддаль між

частинками, що визначається з умови $\frac{4\pi}{3} \langle r \rangle^3 = \frac{V}{N}$.

Для забезпечення умови електронейтральності вважаємо, що в об'ємі плівки рівномірно розподілено додатній заряд з густиноро $\rho = \rho(\bar{\eta})$, тобто виконується умова

$$\int d\bar{\eta} \rho(\bar{\eta}) = eN, e > 0$$

Гамільтоніан моделі, яку розглядаємо, запишемо в такому вигляді:

$$H = H_{nl} + H_{jell} + V_{nl-jell} = \\ = -\frac{\hbar^2}{2m} \sum_{i=1}^N \Delta_i + \sum_{i=1}^N V(\bar{R}_i) + \\ + \frac{1}{2} \sum_{i=1}^N \sum_{j=1, i \neq j}^N \frac{e^2}{|\bar{R}_i - \bar{R}_j|} + \frac{1}{2} \sum_{m=1}^N \sum_{j=1}^{N_{jel}} \frac{e^2}{|\bar{r}_j - \bar{R}_m|},$$

де $\sum_{i=1}^N V(\bar{R}_i)$ – енергія взаємодії електронів плівки з поверхнями поділу, N – кількість електронів у плівці, N_{jel} – кількість електронів у підкладі, e –