КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

Факультет комп’ютерних наук та кібернетики

Кафедра моделювання складних систем

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття ступеня бакалавра

за спеціальністю 113 Прикладна математика

на тему:

МАТЕМАТИЧНІ МОДЕЛІ. ЗБІЖНІСТЬ ПРОЦЕСІВ НЕЙРОДИНАМІКИ

Виконавець:

бакалавр четвертого курсу

Самілика Антона Анатолійовича ____________________

(підпис)

Науковий керівник:

професор, доктор фізико-математичних наук

Хусаінов Денис Ях'євич ____________________

(підпис)

Роботу заслухано на засіданні кафедри моделювання складних систем та рекомендовано до захисту, протокол № 18 від **10 червня 2022р.**

Завідувач кафедри МСС _____________ д.т.н., доцент Дмитро ЧЕРНІЙ

(підпис)

Київ – 2022

Анотація

В даній роботі розглянуто методи дослідження збіжності процесів нейродинаміки. Наведено основні підходи до вирішення проблем стійкості в процесах навчання нейронних мереж. Запропоновані варіанти застосування функцій Ляпунова для дослідження стійкості в певних моделях нейромереж.

Зміст

Вступ

Загальні поняття з теорії нейронних мереж

Дослідження штучних нейронних мереж пов’язано з тим, що спосіб обробки інформації людським мозком відрізняється від методів, які використовує процесор комп`ютера. Мозок являє собою досить складний, нелінійний, паралельний комп`ютер. Він має властивість організовувати власні структурні компоненти, які називаються нейронами, щоб вони могли виконувати конкретні задачі, набагато швидше ніж сучасні комп`ютери. Прикладом такої задачі може бути звичайний людський зір. Функція зору заключається у тому, щоб представляти навколишній світ в такому вигляді, який дає змогу взаємодіяти з цим навколишнім світом. Наприклад, розпізнавання обличчя звичайною людиною відбувається за 100-200 мілісекунд, в свою чергу комп`ютеру на вирішення даної задачі знадобиться набагато більше часу.

Таких чудових результатів людський мозок досягає через можливість будувати власні правила на основі досвіду, який накопичується з часом.

Поняття розвитку мозку пов`язано з поняттям ‘пластичності’ – можливість налаштовувати нервову систему у відповідність до навколишнього світу. Пластичність відіграє найважливішу роль в роботі нейронів у людському мозку. Аналогічно і в нейронних мережах робота проводиться зі штучними нейронами. В загальному нейронна мережа являє собою машину, яка моделює спосіб обробки певної задачі людським мозком. Для того щоб досягнути високої продуктивності, нейронні мережі використовують множину зв`язків між елементарними частинами, які називаються нейронами.

Нейронна мережа - це величезний розподілений паралельний процесор, що складається з елементарних одиниць обробки інформації, що накопичують експериментальні знання і надають їх для подальшої обробки. Нейронна мережа подібна до мозку з двох точок зору:

- Знання надходять в нейронну мережу з навколишнього середовища і використовуються в процесі навчання нейромережі;
- Для накопичення знань застосовуються зв`язки між нейронами, які називаються синаптичними вагами.

Процедура, яка використовується для процеса навчання, називається алгоритмом навчання. Алгоритм навчання забезпечує визначення (налаштування) синаптичних ваг, який визначить правильну поведінку нейронної мережі, задля вирішення конкретних задач.

Історія створення (об'єкти дослідження)

Філософія сформулювала найважливіші положення, що керують раціональною частиною мислення, але для їх формалізації необхідні були фундаментальні дослідження в іншій науці – математиці. Протягом кількох століть ці дослідження проводилися паралельно, взаємно збагачуючи обидві науки. Для штучного інтелекту найбільше вплинуло розвиток таких розділів математики як логіка, обчислення та ймовірність.

Досягнення в галузі філософії та математики сприяли створенню перших обчислювальних пристроїв.

Незважаючи на досить тривалий період досліджень природи мислення, практичних результатів було досягнуто мало. Це багато в чому пов'язано з тим, що методи наукових досліджень, що дозволили отримати значні результати в таких галузях, як астрономія, фізика та хімія, виявилися не ефективними щодо людини.

Значний прорив стався у 20 столітті, він пов'язаний із:

1. Досягненнями в галузі нейрофізіології, нейроанатомії та психології;
2. Інтеграцією досягнень різних наук у нову галузь науки - штучний інтелект.

У шістдесяті роки минулого століття група дослідників в областях нейробіології та нейроанатомії встановила, що мозок – це сотні мільярдів нейронів, з'єднаних один з одним.

Розуміння функціонування нейрона та його зв'язків дозволило дослідникам створити математичні моделі, які, у свою чергу, стали теоретичною основою створення штучних нейронних мереж.

Перші штучні нейронні мережі були реалізовані у вигляді електронних схем. Пізніше у зв'язку з розвитком обчислювальної техніки штучні нейронні мережі стали реалізовуватися як програми. Оскільки у штучному інтелекті систематизуються та автоматизуються інтелектуальні завдання будь-якої сфери інтелектуальної діяльності, штучний інтелект стає справді універсальною науковою областю.

Завдяки дослідженню таких вчених, як Кохонен, Гроссберг, Андерсон, сформувався теоретичний фундамент, на основі якого стало можливо конструювання потужних багатошарових мереж. Однак проблема полягала в їх навчанні.

1974 р. - П. Вербосом розроблено алгоритм зворотного розповсюдження помилки для навчання багатошарових перцептронів, перевідкритий в 1982 р. Д. Паркером та в 1986 року Девідом І. Румельхартом, Дж. Є. Хінтоном та Рональдом Дж. Вільямсом і незалежно і одночасно С.І. Барцевим та В.А. Охоніним. Цей систематичний метод для навчання багатошарових мереж долає обмеження, вказані Мінським.

Подальші дослідження показали, що цей метод не є універсальним, незважаючи на багато успішних практичних результатів. Проблема полягає в дуже довгому процесі навчання, а в деяких випадках мережа може взагалі не навчитися. Останнє можливе з двох причин: параліч мережі та потрапляння до локального мінімуму.

1975 р. - Фукусіма представляє когнітрон - мережу, що самоорганізується, призначену для інваріантного розпізнавання образів.

1980 р. - у спробах покращити когнітрон Фукусімою була розроблена потужна парадигма, названа неокогнітроном.

1982 р. - Дж. Хопфілд розробив нейронну мережу із зворотними зв'язками. Хоча мережа мала цілу низку недоліків і не могла бути використана на практиці, вчений заклав основи нейронних рекурентних мереж, після чого про штучні нейронні мережі стало можливим говорити як про асоціативну пам'ять.

1982 - Кохоненом представлені моделі мережі, яка навчається без вчителя на основі самоорганізації.

1987 р. - Роберт Хехт-Нільсон, вирішуючи тимчасові обмеження зворотної мережі розповсюдження помилки, розробив мережі зустрічного поширення. Час навчання у таких мережах в порівнянні зі зворотним розповсюдженням може бути зменшеним в сто разів.

Мета

Дослідження та вдосконалення великих систем нейронних мереж використовуючи знання з теорії динамічних систем. Виконання опису роботи моделі Хопфілда, теореми Коена-Гроссберга, функції Ляпунова для вивчення поняття стійкості в нейромережах великої розмірності. Дослідження використання звичайних диференціальних рівнянь та рівнянь з запізненням в теорії нейромереж. Висвітлення методів для отримання умов збіжностей для процесів навчання моделей нейродинаміки з післядією.

Можливості

На сьогодні розроблено досить багато моделей нейронних мереж, які мають свої особливості. У суспільстві багато проблем зводяться до питання управління слабоструктурованими, а часто і неструктурованими складними системами. Тому оцінити перспективи розвитку нейронних мереж можна, лише вирішивши проблеми процесів нейродинаміки, що дасть поштовх для розвитку надскладних нейромереж для моделювання склданих процесів.

Потенційними областями застосування штучних нейронних мереж є ті, де людський інтелект є малоефективним, а традиційні обчислення трудомісткі або фізично неадекватні (тобто не відображають або погано відображають реальні фізичні процеси та об'єкти). Актуальність застосування нейронних мереж багаторазово зростає, коли виникає необхідність вирішення погано формалізованих завдань. Тому дослідження нейромереж, а в особливості збіжності процесів нейродинаміки, дозволяє покращити вирішення ряду неформалізованих задач:

1) Фінанси –прогнозування курсу акцій за допомогою аналізу часових рядів є можливим завдяки побудові складної нейромережі , яка моделює процеси ринку;

2) Нейрокерування – нейромережі використовують для керування динамічними об'єктами, математичні моделі яких включають в себе велику кількість формалізованих параметрів та елементи динамічного програмування;
3) Економіка - нейромережа здатна моделювати проблеми економіко-статистичних процесів, зводячи задані моделі дослідження до реалістичних сценаріїв;
4) Дослідження інших складних процесів.

Розділ 1

Використання звичайних диференціальних рівнянь та рівнянь з запізненням в теорії нейромереж

Особливістю, яка виникає при дослідженні динамічних процесів у нейродинаміці, є облік часу післядії. Реальні процеси неспроможні відбуватися миттєво. Тому ігнорування цього чинника може призвести до помилкових висновків щодо динаміки навчання чи розпізнавання.

Слід зазначити, що при дослідженні динаміки потрібно враховувати час обробки та передачі сигналів. Тому найбільш адекватними моделями є моделі, що враховують запізнення. Апаратом дослідження таких систем може бути обраний метод порівняння та метод функціоналів Ляпунова-Красовського.

1.1 Місце теорії динамічних систем в нейронних мережах

Область знань, у якій нейронні мережі розглядаються як нелінійні динамічні системи та основна увага приділяється проблемам стійкості та збіжності називається нейродинамікою. Класична теорія стійкості має на увазі стійкість окремо взятої траєкторії (як правило положення рівноваги). Тому розглядаються моделі, що описуються динамічними системами (диференціальними або різницевими), що мають одне положення рівноваги.

Системи рівнянь, що описують динамічні процеси нейродинаміки, мають дуже велику розмірність. Тому пряме використання класичних методів дослідження стійкості тут неможливе. Класичні способи дослідження зіштовхуються з проблемами великих розмірностей. І для систем, які описують нейродинаміку, часто використовують поняття “системи зі слабкою нелінійністю”, тобто. системи з виділеною лінійною частиною та нелінійною, розташованою в деякому досить малому секторі. Вочевидь, що з “слабкої нелінійності” стійкість становища рівноваги нелінійної системи випливатиме з асимптотичної стійкості системи лінійного наближення. Звідси, зокрема, слідуватиме збіжність процесів навчання.

Ще одним підходом дослідження динамічних систем великої розмірності є метод порівняння. Система розбивається на підсистеми та будуються окремі функції Ляпунова. Висновок про стійкість вихідної системи виробляється виходячи з дослідження системи порівняння, отриманої з функцій Ляпунова окремих підсистем.

Одним з підходів подолання фактора "великої розмірності" є використання поняття "колективна динаміка", що з'явилося останнім часом. Воно дозволяє досліджувати групову поведінку систем великої розмірності. Однак у цьому напрямі робляться лише перші кроки і для коливальних систем.

1.2 Стійкість нульового розв’язку лінійних стаціонарних систем

Розглянемо однорідну систему диференціальних рівнянь

$$\dot{x} = Ax , \qquad (1.1)$$

де A – матриця зі сталими коефіцієнтами. Дослідимо на стійкість нульовий розв’язок $x(t) \equiv 0$.

Теорема 1.1 Для лінійної однорідної системи зі сталими коефіцієнтами можливі такі випадки:

1. Для того, щоб нульовий розв'язок $x(t) \equiv 0$ лінійної системи із сталими коефіцієнтами (1.1) був асимптотично стійким, необхідно і достатньо щоб усі корені характеристичного рівняння системи (1.1) мали від'ємну дійсну частину, тобто $\operatorname{Re} \lambda_i(\mathrm{A}) < 0$, $i = \overline{1, n}$.

2. Якщо хоча б один корінь характеристичного рівняння системи (1.1) має додатню дійсну частину, тобто існує λ_s, таке, що $\operatorname{Re} \lambda_s(\mathrm{A}) > 0$, то нульовий розв'язок $x(t) \equiv 0$ системи є нестійким.

3. Для того, щоб розв'язок $x(t) \equiv 0$ лінійної системи із сталими коефіцієнтами (1.1) був стійким по Ляпунову, необхідно і достатньо, щоб усі корені характеристичного рівняння системи (1.1) мали недодатню дійсну частину, тобто $\operatorname{Re} \lambda_i(\mathrm{A}) \le 0$, $i = \overline{1, n}$, причому числа з нульовою дійсною частиною мали прості елементарні дільники, тобто матриця Жордана зводилася до одного елемента.

Таким чином, для дослідження стійкості лінійної системи зі сталими коефіцієнтами треба розкрити характеристичне рівняння

$$\det(\mathrm{A} - \lambda E) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0,$$

переписати його у вигляді

$$\lambda^n + p_1 \lambda^{n-1} + \ldots + p_{n-1} \lambda + p_n = 0 \tag{1.2}$$

та дослідити його корені отриманого полінома.

Теорема 1.2*(Необхідна умова стійкості)* Якщо характеристичне рівняння (1.2) має корені з від'ємною дійсною частиною, то всі коефіцієнти характеристичного рівняння (1.2) додатні, тобто $p_i > 0$, $i = \overline{1,n}$.

1.3 Критерій Гурвіца. Критерій Михайлова

Як було показано в попередній частині, критерієм стійкості (асимптотичної стійкості) є розташування дійсних частин коренів характеристичного рівняння.

Нехай коефіцієнти характеристичного рівняння

$$\lambda^n + p_1\lambda^{n-1} + ... + p_{n-1}\lambda = 0 \tag{1.3}$$

записані у вигляді так званої матриці Гурвіца

$$\Gamma = \begin{bmatrix} p_1 & 1 & 0 & 0 & \cdots & 0 \\ p_3 & p_2 & p_1 & 1 & \cdots & 0 \\ p_5 & p_4 & p_3 & p_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{2n+1} & p_{2n} & p_{2n-1} & p_{2n-2} & \cdots & p_n \end{bmatrix},$$

де p_i, $i = \overline{1,n}$ – дійсні коефіцієнти, $p_i = 0$ при $i > n$. Має місце наступний критерій асимптотичної стійкості.

Теорема 1.3 *(критерій Гурвіца)* Для того, щоб характеристичне рівняння (1.3) мало корені з від'ємною дійсною частиною необхідно і достатньо, щоб усі головні діагональні мінори матриці Гурвіца були додатні, тобто

$$\Delta_1 = p_1 > 0,\ \Delta_2 = \begin{vmatrix} p_1 & 1 \\ p_3 & p_2 \end{vmatrix} > 0,\ \ldots\ \Delta_n > 0.$$

Наслідок 1.1 Нехай розглядається характеристичне рівняння системи другого порядку, тобто

$$\lambda^2 + p_1\lambda + p_2 = 0.$$

Для системи другого порядку необхідною і достатньою умовою асимптотичної стійкості є додатність коефіцієнтів $p_1 > 0$, $p_2 > 0$.

Іноді, особливо в технічних задачах, зручно користуватися так званими частотними критеріями. Перепишемо ліву частину характеристичного рівняння (1.3) у вигляді поліному

$$f(z) = z^n + p_1 z^{n-1} + \ldots + p_{n-1} z + p_n , \tag{1.4}$$

де p_i, $i = \overline{1,n}$ – дійсні коефіцієнти. Крива $\omega = f(i\omega)$, де $0 \le \omega < \infty$, $i = \sqrt{-1}$ – уявна одиниця, називається годографом Михайлова.

Лема 1.1. Нехай $f(z)$ – поліном степеня n, який не має чисто уявних коренів. Тоді кут повороту проти ходу годинникової стрілки ненульового вектора $f(i\omega)$ при $0 \le \omega < +\infty$ дорівнює

$$\Delta Arg f(i\omega) = \frac{\pi}{2}(n - 2m),$$

де m – число коренів поліному $f(z)$ з додатньою дійсною частиною враховуючи їх кратності. На базі цієї леми формулюється критерій стійкості Михайлова.

Теорема 1.4 *(Критерій Михайлова)* Для того щоб характеристичне рівняння мало корені з від'ємною дійсною частиною необхідно і достатньо, щоб кут повороту проти годинникової стрілки вектора $f(i\omega)$ при зміні $0 \le \omega < +\infty$ дорівнював

$$\Delta Arg f(i\omega) = \frac{\pi}{2} n .$$

1.4 Основні поняття другого методу Ляпунова. Геометрична та фізична інтерпретація

Розглянемо систему диференціальних рівнянь вигляду

$$\dot{x} = f(x,t), \; x = (x_1, x_2, ..., x_n), \tag{1.5}$$

де $f(x,t)$ та $\frac{\partial f}{\partial x_k}$ неперервні функції при $t_0 \le t < \infty$. Крім того $f(0,t) \equiv 0$, тобто система має нульовий розв'язок.

Основна ідея другого методу А.М.Ляпунова полягає в тому, що стійкість нульового розв'язку $x(t) \equiv 0$ системи (1.5) визначається з поведінки наперед визначеної функції Ляпунова.

Коливання маятника, чи рух кульки можна описати за допомогою систем диференціальних рівнянь. Ці системи мають два стани рівноваги – верхній та нижній.

Верхній стан рівноваги – нестійкий, нижній – стійкий. Якщо за характеристичну функцію брати повну енергію, то стійкість буде там, де енергія мінімальна, причому, при переході до стійкого стану рівноваги енергія зменшується. Таким чином дослідження стійкості фізичної системи можна проводити, використовуючи функцію повної енергії (енергетичної функції).

Розглянемо в деякій області $H \times R = \{t \ge t_0, \|x\| \le h\}$ функцію $n+1$ змінних $V(x_1, x_2, ..., x_n, t)$, або $V(x,t)$.

Означення 1.1 Функція $V(x)$ називається додатньо визначеною, якщо $V(x) > 0$ при $x \ne 0$ та $V(0) = 0$.

Означення 1.2 Функція $V(x,t)$ називається додатньо визначеною, якщо існує така додатньо визначена функція $W(x)$, що $V(x,t) \geq W(x)$ та $V(0,t) \equiv 0$.

Означення 1.3 Функція $V(x,t)$ допускає нескінченно малу нижню границю, якщо існує додатньо визначена функція $W_1(x)$ така, що $|V(x,t)| \leq W_1(x)$.

Означення 1.4 Повною похідною функції $V(x,t)$ у силу системи (1.5) називається функція

$$\frac{dV}{dt} = \frac{\partial V(x,t)}{\partial t} + (gradV(x,t), f(x,t)) =$$
$$= \frac{\partial V(x,t)}{\partial t} + \sum_{i=1}^{n} \frac{\partial V(x,t)}{\partial x_1} f_i(x_1, x_2, ..., x_n, t).$$

Наведемо основні теореми про стійкість та нестійкість руху. В основі другого методу Ляпунова лежать дві теореми про стійкість та теорема Четаєва про нестійкість.

Теорема 1.5 *(Перша Ляпунова про стійкість)* Нехай існує неперервно-диференційована функція $V(x,t)$, що задовольняє умовам:

1. $V(x,t)$ – додатньо визначена в деякому околі $x = 0$ (тобто $V(x,t) > 0$);
2. повна похідна функції $V(x,t)$ у силу системи не додатня (тобто $\frac{dv(x,t)}{dt} \leq 0$).

Тоді нульовий розв'язок $x(t) \equiv 0$ системи (1.5) стійкий по Ляпунову.

Теорема 1.6 *(Друга Ляпунова про асимптотичну стійкість).* Нехай існує неперервно-диференційована функція $V(x,t)$, що задовольняє умовам:

1. $V(x,t)$ – додатньо визначена (тобто $V(x,t) > 0$);
2. $V(x,t)$ – допускає нескінченно малу вищу границю (тобто $|V(x,t)| \le W_1(x)$);
3. повна похідна функції $V(x,t)$ у силу системи від'ємно визначена (тобто $\frac{dV(x,t)}{dt} < 0$).

Тоді нульовий розв'язок $x(t) \equiv 0$ системи (1.5) асимптотично стійкий.

Теорема 1.7 *(Четаєва, про нестійкість)* Нехай існує неперервно-диференційована функція $V(x,t)$, область додатності якої $\Pi(x,t) = \{(x,t) : V(x,t) > 0\}$ при кожному фіксованому $t \ge t_0$ має ненульовий відкритий перетин $D_t(x)$, що примикає до початку координат, а на границі області $\Pi(x,t)$ виконується рівність $V(x,t) = 0$.

Якщо виконуються умови:

1. функція $V(x,t)$ обмежена в $\Pi(x,t)$;
2. в області $\Pi(x,t)$ справедлива нерівність $\frac{dV(x,t)}{dt} > 0$;
3. існує функція $\beta(\alpha)$ така, що при $V(x,t) \ge a > 0$ буде $\frac{dV(x,t)}{dt} \ge \beta(\alpha) > 0$.

Тоді нульовий розв'язок $x(t) \equiv 0$ системи (1.5) нестійкий.

Метод функцій Ляпунова допускає просту геометричну інтерпретацію:

1. Існування додатньо визначеної функції Ляпунова означає існування

всюди щільної системи поверхонь рівня $V(x,t)=a$, які не розширюються та охоплюють початок координат.

2. Нескінченно мала вища границя функції $V(x,t)$ означає, що поверхні рівня $V(x,t)=\alpha$ – не стягуються при $t\to+\infty$ до початку координат.
3. Умова $\frac{dV(x,t)}{dt}<0$, $\left(\frac{dV(x,t)}{dt}\le 0\right)$ означає, що векторне поле системи спрямоване усередину областей, обмежених поверхнями рівня (або дотикається до них).

Перша терема Ляпунова, про стійкість представлена на Рис.1.1. Друга теорема Ляпунова, про асимптотичну стійкість представлена на Рис.1.2, теорема Четаєва на Рис.1.3

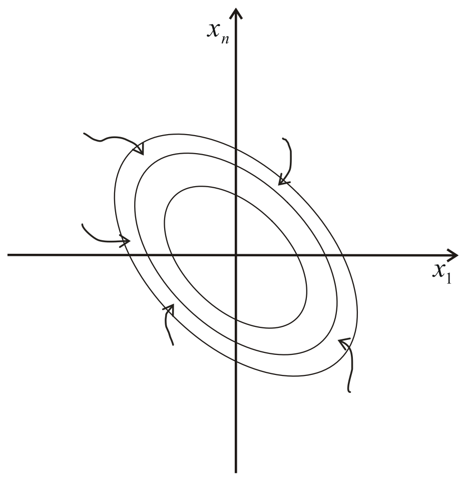

Рис.1.1 Геометрична інтерпретація 1-ї теореми Ляпунова

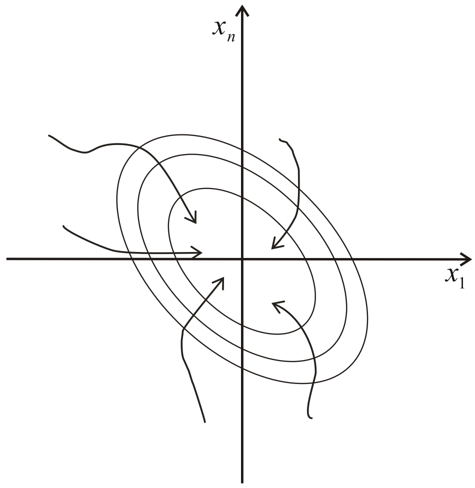

Рис.1.2 Геометрична інтерпретація 2-ї теореми Ляпунова

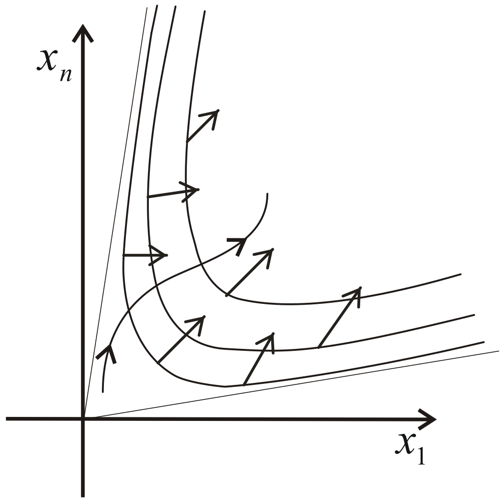

Рис.1.3 Геометрична інтерпретація теореми Четаєва

Розділ 2

Нейронні мережі

2.1 Модель нейрона (на прикладі фізичного явища)

У фізичних (електричних) термінах синаптичні ваги w_{ij}, є ємністю, а відповідні вихідні сигнали $x_i(t)$, $i=\overline{1,n}$ потенціали, де n - кількість входів. Ці сигнали подаються на з'єднуючі суматори. Загальний струм можна записати як

$$I = \sum_{i=1}^{n} w_{ij} x_j(t) + I_i \ ,$$

де перший доданок відображає збудження, що діють на синаптичні ваги, а другий доданок - це джерело струму, що представляє зовнішній зсув. Нехай, $v_i(t)$ індуковане локальне поле на вході нелінійної функції активації - $\varphi(\bullet)$. Тоді загальний струм, що виходить з вхідного вузла нелінійного елемента, можна виразити так

$$I = \frac{v_i(t)}{R_i} + C_i \frac{dv_i(t)}{dt} \ ,$$

де перший доданок викликаний опором (R_i), а другий ємністю (C_i). Застосувавши закон Кірхгофа, отримуємо

$$C_i \frac{dv_i(t)}{dt} = -\frac{v_i(t)}{R_i} + \sum_{j=1}^{n} \omega_{ij} x_j(t) + I_j \ .$$

Для заданого індукованого локального поля $v_i(t)$ можна визначити вихід нейрона i за допомогою відношення

$$x_j(t) = \varphi(v_j(t))$$

Функція активації $\varphi(\bullet)$, яка визначає відношення виходу $x_j(t)$ нейрона j до його ж індукованого локального поля $v_j(t)$ є неперервно - диференційованою функцією. Найчастіше як функцію активації використовують логістичну функцію

$$\varphi(v_j) = \frac{1}{1+\exp\{-v_j\}}, \ j = \overline{1,n}.$$

Тоді проведемо заміну

$$a_i = \frac{1}{R_i C_i}, \ w_i = \frac{\omega_{ij}}{C_i},$$

отримаємо систему

$$\frac{dv_i(t)}{dt} = -a_i v_i(t) + \sum_{j=1}^{n} w_{ij\,j}\varphi(v_j(t)) + K_j.$$

Ще одна модель нейродинаміки може бути описана системою диференціальних рівнянь

$$\frac{dx_i(t)}{dt} = -a_i x_i(t) + \varphi\left(\sum_{j=1}^{n} w_{ij} x_j(t)\right).$$

2.2 Модель Хопфілда

Мережа Хопфілда складається з багатьох нейронів, що формують систему з множиною зворотних зв'язків. Кількість зворотних зв'язків дорівнює кількості нейронів. Вихід кожного нейрона замикається через елемент одиничної затримки на решту нейронів мережі. І нейрон цієї мережі немає зворотних зв'язків із собою. У цьому випадку рівняння динаміки можна переписати як

$$C_i \frac{dv_i(t)}{dt} = -\frac{v_i(t)}{R_i} + \sum_{j=1}^{n} \omega_{ij} \varphi_j \left(v_j(t)\right) + I_j \ . \qquad (2.1)$$

Нехай мережу Хопфілда задано. Вона матиме динамічну рівновагу, якщо на її вхід подається образ, який відображається сам у себе. З погляду фізики повна енергія системи у цій точці буде мінімальною. Енергетична функція (по суті, функція Ляпунова) може бути задана у вигляді суми кінетичної та потенційної енергії

$$V(y) = \frac{1}{2} \sum_{i,j=1}^{n} w_{ij} y_i y_j + \sum_{k=1}^{n} \delta_k y_k \ . \qquad (2.2)$$

І, як випливає з теорем другого методу Ляпунова, стаціонарний стан мережі буде асимптотично стійким, якщо функція Ляпунова буде додатньо визначеною, а повна похідна вздовж ітерація буде від’ємно визначеною

$$C_i \frac{dy_i(t)}{dt} = -\frac{1}{R_i} y_i(t) + \sum_{j=1}^{n} v_{ij} \varphi_j \left(y_j(t)\right) + I_i \ , \ t \ge 0 \ , \ i = \overline{1,n} . \qquad (2.3)$$

Якщо враховується час обробки сигналу, то використовуються системи диференціально-різнецевих рівнянь із запізненням

$$C_i \frac{dy_i(t)}{dt} = -\frac{1}{R_i} y_i(t) + \sum_{j=1}^{n} v_{ij} \varphi_j \left(y_j(t-\tau)\right) + I_k \ , \ t \ge 0 \ , \ i = \overline{1,n} . \qquad (2.4)$$

Тут n - число нейронів в мережі, $y(t) = \left(y_1(t), y_2(t), ..., y_n(t)\right)^T$ - вектор стану нейронної мережі в момент часу $t > 0$; R_i, C_i, I_i, $i = \overline{1,n}$ - відповідно, опір, ємність і зовнішній струм. Заміною $a_i = \dfrac{1}{R_i C_i}$, $\omega_{ij} = \dfrac{v_{ij}}{C_i}$, $I_i^0 = \dfrac{I_i}{C_i}$, $i = \overline{1,n}$, $j = \overline{1,n}$ системи (2.1), (2.2) можливо звести до систем стандартного вигляду

$$\frac{dy_i(t)}{dt} = -a_i y_i(t) + \sum_{j=1}^{n} \omega_{ij} \varphi_j \left(y_j(t)\right) + I_i^0 \ , \ t \ge 0 \ , \ i = \overline{1,n} . \qquad (2.5)$$

А якщо враховувати час запізнення, то

$$\frac{dy_i(t)}{dt} = -a_i y_i(t) + \sum_{l=1}^{n} \omega_{ij}\varphi_j\left(y_j(t-\tau)\right) + I_i^0,\ t \ge 0,\ i = \overline{1,n} \quad (2.6)$$

2.3 Теорема Коена-Гроссберга

Задано певний клас нейронних мереж, який описується наступною системою нелінійних диференціальних рівнянь:

$$\frac{d}{dt}u_j = a_j(u_j)\left[b_j - \sum_{i=1}^{N} c_{ji}\,\varphi_i(u_i)\right], j = 1, \dots, N.$$

(2.7)

Відповідно, заданий клас нейронних мереж дозволяє визначити функції Ляпунова наступним чином:

$$E = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} c_{ij}\,\varphi_i(u_i)\varphi_j(u_j) - \sum_{j=1}^{N}\int_0^{u_j} b_j(\lambda)\varphi'_j(\lambda)d(\lambda),$$

(2.8)

де

$$\varphi'_j(\lambda) = \frac{d}{d\lambda}(\varphi_j(\lambda))$$

(2.9)

Для того щоб визначення (2.8) було коректним, необхідно щоб виконувалися наступні умови:

1. Умови симетрії . Синаптичні ваги повинні бути симетричними:

$$c_{ij} = c_{ji} \quad (2.10)$$

2. Умови невід’ємності:

$$a_j(u_j) \ge 0 \quad (2.11)$$

3. Умова монотонності. Нелінійна функція відображення входу на вихід повинна бути монотонною:

$$\varphi'_j(u_j) = \frac{d}{du_j}\varphi_j(u_j) \ge 0$$

(2.12)

Тепер можемо сформулювати теорему Коена-Гроссберга.

Теорема 2.1 (*теорема Коена-Гроссберга*) Якщо система нелінійних диференціальних рівнянь (2.7) задовольняє умовам симетрії (2.10), невід'ємності (2.11) та монотонності (2.12) , то функція Ляпунова системи (2.8) задовольняє умові

$$\frac{\mathrm{dE}}{\mathrm{dt}} \leq 0$$

Якщо функція Ляпунова має задану глобальну властивість, то відповідно першій теоремі Ляпунова, ця система є стікою глобально.

Розділ 3

Отримання умов збіжності для моделей нейродинаміки з післядією

3.1 Загальний підхід для нелінійних систем

Для систем диференціальних рівнянь із післядією (нелінійних із запізненням)

$$\dot{x}(t) = F\left(x(t), x(t-\tau), t\right),\ \tau > 0 \tag{3.1}$$

цей метод має такі особливості. Векторне поле системи (3.1) визначається не лише поточним положенням $x(t)$, але й попереднім $x(t-\tau)$. Тому, при дослідженні стійкості та отриманні оцінок збіжності другим методом Ляпунова було запропоновано величину повної похідної функції Ляпунова у момент $t > 0$ для кординати $x(t)$ оцінювати у припущенні, що запізнювальна координата $x(s),\ s < t$ знаходиться "всередині" поверхні рівня $V(x,t) = c$. Формально ця пропозиція була сформульована у вигляді залежності

$$V\left(x(s), s\right) < V\left(x(t), t\right), \text{ при } s < t .$$

Сенс цієї пропозиції наступний. Для довільного $\varepsilon > 0$ завжди знайдуться досить малі $\delta > 0$ і $\alpha > 0$, при яких розв'язок $x(s)$ системы (3.1), задовольняє при $0 \le s \le \tau$ початковій умові $|x(s)| < \delta$ буде знаходитися в області $V^{\alpha} = \{x \in R^n : V(x) < \alpha\}$, яка, в свою чергу, буде перебувати в $\varepsilon-$ околі нульового положення рівноваги $U_{\varepsilon} = \{x \in R^n : |x| < \varepsilon\}$. Якщо, від протилежного, в момент $t > 0$ розв'язок $x(t)$досягає межі цієї множини, тобто $|x(t)| = \varepsilon$, то він має перетнути кордон $\partial V^{\alpha} = \{x \in R^n : V(x) = \alpha\}$. І це неможливо, оскільки функція Ляпунова не зростає, що виходить з того що повна похідна від'ємно визначена. Геометрично умова від'ємної визначеності функції Ляпунова означає, що векторне поле системи на межі $\partial V^{\alpha} = \{x \in R^n : V(x) = \alpha\}$ спрямовано всередину.

Якщо функція Ляпунова має вигляд квадратичної форми $V(x) = x^T Hx$ (що зазвичай буває) із симетричною додатньо визначеною матрицею H, то умова Разуміхіна має вигляд

$$\lambda_{\min}(H)|x(s)|^2 \le x^T(s)Hx(s) = V(x(s)) < V(x(t)) = x^T(t)Hx(t) \le \lambda_{\max}(H)|x(t)|^2,$$
$$0 \le s \le t.$$

Тому при оцінці повної похідної функції Ляпунова через систему використовується нерівність

$$|x(s)| < \sqrt{\varphi(H)}|x(t)|,\ \varphi(H) = \lambda_{\max}(H)/\lambda_{\min}(H),\ 0 \le s \le t, \tag{3.2}$$

де $\lambda_{\max}(H)$, $\lambda_{\min}(H)$ максимальні та мінімальні власні числа симетричної, додатньо визначеної матриці H.

3.2 Лінійні системи з запізненням

Розглянемо систему однорідних диференціальних рівнянь зі сталими коефіцієнтами

$$\dot{x}(t) = Ax(t) + Bx(t-\tau),\ \tau > 0. \tag{3.3}$$

Дослідження стійкості проводилося методом квадратичних функцій Ляпунова $V(x) = x^T Hx$ із симетричною, додатньо визначеною матрицею H. Передбачалося, що "модельна система"

$$\dot{x}(t) = (A + B)x(t) \tag{3.4}$$

є асимптотично стійкою. Повна похідна функції Ляпунова $V(x)=x^T Hx$ в силу системи (3.3) записувалась у вигляді

$$\frac{d}{dt}V(x(t))=(\dot{x}(t))^T Hx(t)+x^T(t)H(\dot{x}(t))=$$
$$=(Ax(t)+Bx(t-\tau))^T Hx(t)+x^T(t)H(Ax(t)+Bx(t-\tau)).$$

Далі, додавався і віднімався другий член без запізнення. Виходило

$$\frac{d}{dt}V(x(t))=x^T(t)\left[(A+B)^T H+H(A+B)\right]x(t)+2x^T(t)HB[x(t-\tau)-x(t)]. \quad (3.5)$$

Якщо матриця $A+B$ асимптотично стійка, то матричне рівняння Ляпунова

$$(A+B)^T H+H(A+B)=-C \quad (3.6)$$

при будь-якій додатньо визначеній матриці $C=C_0$ мало розв'зком додатньо визначену матрицю $H=H_0$. Таким чином, залежність (3.5) має вигляд

$$\frac{d}{dt}V(x(t))=-x^T(t)C_0x(t)+2x^T(t)H_0B[x(t-\tau)-x(t)]. \quad (3.7)$$

3.3 Умови стійкості, рівномірні за запізненням

Використовуючи нерівність (3.2), отримуємо, що для повної похідної функції Ляпунова (3.7) має місце таке співвідношення

$$\frac{d}{dt}V(x(T))\le-\lambda_{\min}(C_0)|x(T)|^2+2|H_0B|\left[1+\sqrt{\varphi(H_0)}\right]|x(T)|^2=$$
$$=-\left\{\lambda_{\min}(C_0)-2|H_0B|\left[1+\sqrt{\varphi(H_0)}\right]\right\}|x(T)|.$$

І при виконанні умови

$$\lambda_{\min}(C_0)-2|H_0B|\left[1+\sqrt{\varphi(H_0)}\right]>0 \quad (3.8)$$

Повна похідна функції Ляпунова буде від'ємно визначеною. Це означає, що векторне поле системи із запізненням спрямоване всередину поверхні рівня $\partial V^{\alpha} = \{x : x^T H_0 x = \alpha\}$, і розв'язок $x(t)$ не покине області $V^{\alpha} = \{x : x^T H_0 x < \alpha\}$. Таким чином, нульовий розв'язок системи (3.3) буде стійким за Ляпуновим. Більше того, можна показати, що воно є експоненційно стійким. Дійсно, враховуючи нерівності квадратичних форм

$$\lambda_{\min}(H_0)|x(t)|^2 \le V(x(t)) = x(t)^T H x(t) \le \lambda_{\max}(H_0)|x(t)|^2, \qquad (3.9)$$

отримаємо, що

$$|x(T)|^2 \le \frac{1}{\lambda_{\min}(H_0)} V(x(t)).$$

І для повної похідної функції Ляпунова матиме місце нерівність

$$\frac{d}{dt} V(x(T)) \le \psi(H_0) V(x(T)), \quad \psi(H_0) = \frac{\lambda_{\min}(C_0) - 2|H_0 B| \left[1 + \sqrt{\varphi(H_0)}\right]}{\lambda_{\min}(H_0)}.$$

Вирішуючи записану диференціальну нерівність, отримуємо

$$V(x(T)) \le V(x(t_0)) \exp\{\psi(H_0)(t - t_0)\}. \qquad (3.10)$$

Знову використовуючи нерівність (3.9), остаточно отримуємо оцінку збіжності розв'язку системи диференціальних рівнянь (3.3).

$$|x(T)| \le \sqrt{\varphi(H_0)} \exp\left\{\frac{1}{2}\psi(H_0)(T - t_0)\right\}. \qquad (3.11)$$

3.4 Умови стійкості, що залежать від запізненя

Система диференціальних рівнянь (3.3) записувалася в інтегральному вигляді

$$x(t) = x(t-\tau) + \int_{t-\tau}^{t} [Ax(s) + Bx(s-\tau)]ds.$$

Звідси випливала нерівність

$$\left|x(t)-x(t-\tau)\right| \le \int_{t-\tau}^{t}\left[|A||x(s)|+|B||x(s-\tau)|\right]ds.$$

Нехай, як і в попередньому випадку, від протилежного, при $t_0-\tau \le t < T$ розв'язок $x(t)$ знаходиться всередині еліпса $V^{\alpha}=\{x : x^T Hx<\alpha\}$, а при $t=T$ досягає його межі $\partial V^{\alpha}=\{x : x^T Hx=\alpha\}$. Тоді

$$\left|x(T)-x(T-\tau)\right| \le (|A|+|B|)\sqrt{\varphi(H_0)}|x(T)|\tau. \qquad (3.12)$$

Знову повернемося до (3.7). Підставивши отримане співвідношення (3.12) (3.7), отримуємо

$$\frac{d}{dt}V(x(T)) \le -\left\{\lambda_{\min}(C_0)-2|H_0B|(|A|+|B|)\sqrt{\varphi(H_0)}\tau\right\}|x(T)|^2. \qquad (3.13)$$

І при

$$\tau<\tau_0,\ \tau_0=\frac{\lambda_{\min}(C_0)}{2|H_0B|(|A|+|B|)\sqrt{\varphi(H_0)}},\ \varphi(H_0)=\frac{\lambda_{\max}(H_0)}{\lambda_{\min}(H_0)}. \qquad (3.14)$$

повна похідна функції Ляпунова буде від'ємно визначеною. Таким чином, векторне поле системи (3.3) буде спрямоване всередину поверхні. $\partial V^{\alpha}=\{x : x^T Hx=\alpha\}$ і, як і в попередньому випадку, лінійна система із запізненням (3.3) буде стійкою.

Покажемо, що з виконання умови (3.14) розв'язок системи експоненційно сходиться до нульовго стану рівноваги. Дійсно, враховуючи нерівності квадратичних форм (3.9) та нерівності (3.14), отримуємо

$$\frac{d}{dt}V(x(T)) \le -\left\{\lambda_{\min}(C_0)-2|H_0B|(|A|+|B|)\sqrt{\varphi(H_0)}\tau\right\}\frac{1}{\lambda_{\max}(H)}V(x(T)).$$

Або маємо

$$\frac{d}{dt}V(x(T)) \le -\gamma(H_0)V(x(T)),\quad \gamma(H_0)=\frac{\lambda_{mim}(C_0)-2|H_0B|(|A|+|B|)\sqrt{\varphi(H_0)}\tau}{\lambda_{\min}(H_0)}, \quad (3.15)$$

Вирішуючи отриману диференціальну нерівність (3.11), отримуємо

$$V(x(T)) \le V(x(0)) \exp\{-\gamma(H_0)(t-t_0)\}.$$

І, знову використовуючи нерівності квадратичних форм, остаточно отримуємо

$$|x(t)| \le \sqrt{\varphi(H_0)} \exp\left\{\frac{1}{2}\gamma(H_0)(t-t_0)\right\}.$$

3.5 Системи зі слабкою нелінійністю

У роботі, як моделі безперервних нейронних мереж Хопфілда (НМХ) було запропоновано системи звичайних диференціальних рівнянь виду

$$C_i \frac{dy_i(t)}{dt} = -\frac{1}{R_i} y_i(t) + \sum_{j=1}^{n} \nu_{ij} \varphi_j(y_j(t)) + I_i, \; t \ge 0, \; i = \overline{1,n}. \tag{3.16}$$

Якщо враховується час обробки сигналу, то використовувалися системи диференціально-різнецевих рівнянь із запізненням

$$C_i \frac{dy_i(t)}{dt} = -\frac{1}{R_i} y_i(t) + \sum_{j=1}^{n} \nu_{ij} \varphi_j(y_j(t-\tau)) + I_k, \; t \ge 0, \; i = \overline{1,n}. \tag{3.17}$$

Тут n- число нейронів, $y(t) = (y_1(t), y_2(t), ..., y_n(t))^T$ - вектор стану мережі на момент часу $t > 0$; R_i, C_i, I_i, $i = \overline{1,n}$ - відповідно, опори, ємності та зовнішні струми. Заміною

$$a_i = \frac{1}{R_i C_i}, \; b_{ij} = \frac{\nu_{ij}}{C_i}, \; I_i^0 = \frac{I_i}{C_i}, \; i = \overline{1,n}, \; j = \overline{1,n}$$

системи (3.16), (3.17) можна привести до систем стандартного вигляду

$$\frac{dy_i(t)}{dt} = -a_i y_i(t) + \sum_{j=1}^{n} b_{ij} f_j(y_j(t)) + c_i, \; t \ge 0, \; i = \overline{1,n}. \tag{3.18}$$

А, якщо враховувати час обробки сигналу та побудови управління, то до систем із запізненням

$$\frac{dy_i(t)}{dt} = -a_i y_i(t) + \sum_{l=1}^{n} b_{ij} f_j\left(y_j(t-\tau)\right) + c_i\,,\ t \geq 0\,,\ i = \overline{1,n} \tag{3.19}$$

Розглянемо динамічні системи, що є моделями нейронних мереж, які описуються системами нелінійних диференціальних рівнянь із запізненням

$$\dot{y}_1(t) = -a_{11} y_1(t) + b_{11} f_{11}\left(y_1(t-\tau)\right) + b_{12} f_{12}\left(y_2(t-\tau)\right) + \ldots + b_{1n} f_{1n}\left(y_n(t-\tau)\right) + c_1$$

$$\dot{y}_2(t) = -a_{22} y_2(t) + b_{21} f_{21}\left(y_1(t-\tau)\right) + b_{22} f_{22}\left(y_2(t-\tau)\right) + \ldots + b_{2n} f_{2n}\left(y_n(t-\tau)\right) + c_2\,.$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\dot{y}_n(t) = -a_{nn} y_n(t) + b_{n1} f_{n1}\left(y_1(t-\tau)\right) + b_{n2} f_{n2}\left(y_2(t-\tau)\right) + \ldots + b_{nn} f_{nn}\left(y_n(t-\tau)\right) + c_n \tag{3.20}$$

Вважаємо, що $a_{ii} > 0$, функції $f_{ij}(y_j)$, $i, j = n$ є безперервно диференційованими, і система диференціальних рівнянь (3.20) має єдину точку спокою $M_0\left(y_1^0, y_2^0, \ldots, y_n\right)$, що є розв'язком системи рівнянь

$$-a_{11} y_1 + b_{11} f_{11}(y_1) + b_{12} f_{12}(y_2) + \ldots + b_{1n} f_{1n}(y_n) + c_1 = 0\,,$$

$$-a_{22} y_2 + b_{21} f_{21}(y_1) + b_{22} f_{22}(y_2) + \ldots + b_{2n} f_{2n}(y_n) + c_2 = 0\,.$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$-a_{nn} y_n + b_{n1} f_{n1}(y_1) + b_{n2} f_{n2}(y_2) + \ldots + b_{nn} f_{nn}(y_n) + c_n = 0$$

Зробимо заміну типу «паралельного перенесення» точки спокою $M_0\left(y_1^0, y_2^0, \ldots, y_n\right)$ в початок кординат

$$y_i(t) = x_i(t) + y_i^0\,,\ i = 1, 2, \ldots, n\,.$$

Тоді система (3.20) зведеться до

$$\dot{x}_1(t) = -a_{11} x_1(t) + b_{11} F_{11}\left(x_1(t-\tau)\right) + b_{12} F_{12}\left(x_2(t-\tau)\right) + \ldots + b_{1n} F_{1n}\left(x_n(t-\tau)\right),$$

$$\dot{x}_2(t) = -a_{22} x_2(t) + b_{21} F_{21}\left(x_1(t-\tau)\right) + b_{22} F_{22}\left(x_2(t-\tau)\right) + \ldots + b_{2n} F_{2n}\left(x_n(t-\tau)\right),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\dot{x}_n(t) = -a_{nn} x_n(t) + b_{n1} F_{n1}\left(x_1(t-\tau)\right) + b_{n2} F_{n2}\left(x_2(t-\tau)\right) + \ldots + b_{nn} F_{nn}\left(x_n(t-\tau)\right). \tag{3.21}$$

де

$$F_{ij}\left(x_j(t-\tau)\right) = f_{ij}\left(x_j(t-\tau) + y_j^0\right) - f_{ij}\left(y_j^0\right). \tag{3.22}$$

Так як

$$F_{ij}(0) = 0\,,\ i, j = 1, n\,,$$

то дослідження стійкості положення рівноваги $M_0\left(y_1^0, y_2^0, ..., y_n^0\right)$ і збіжності рішень до цієї точки зводиться до дослідження нульового положення рівноваги $O(0,0,...,0)$ системи (3.21).

Нехай функції $F_{ij}\left(x_j(t-\tau)\right),\ i, j=\overline{1,n}$ задовольняють так званим «умовам лінійного обмеження»

$$\left|F_{ij}\left(x_j(t-\tau)\right)\right| \le K_{ij}\left|x_j(t-\tau)\right|,\ i, j=\overline{1,n}. \tag{3.23}$$

У цьому випадку асимптотичну стійкість нульового положення рівноваги системи (3.21) та оцінку збіжності можна отримувати з використанням квадратичних функцій Ляпунова

$$V\left(x_1, x_2, ..., x_n\right)=\sum_{i=1}^{n} h_{ii} x_i^2,\ h_{ii}>0,\ i=\overline{1,n}. \tag{3.24}$$

Оскільки у диференціальному рівнянні є члени із запізненням, то для оцінки повної похідної системи будемо використовувати умову Розуміхіна.

Теорема 3.1 Нехай параметри $a_{ii}>0,\ h_{ii}>0,\ b_{ij},\ K_{ij}>0,\ i, j=\overline{1,n}$ системи (3.5.6) такі, що матриця $2C_1(h,a)-C_2(h,L)$, де

$$C_1(h,a)=\begin{bmatrix} h_{11}a_{11} & 0 & ... & 0 \\ 0 & h_{22}a_{22} & ... & 0 \\ . & . & ... & . \\ 0 & & ... & h_{nn}a_{nn} \end{bmatrix},$$

$$C_2(h,L)=\begin{bmatrix} 2h_{11}L_1 & h_{11}L_1+h_{22}L_2 & ... & h_{11}L_1+h_{nn}L_n \\ h_{11}L_1+h_{22}L_2 & 2h_{22}L_2 & ... & h_{22}L_2+h_{33}L_3 \\ . & . & ... & . \\ h_{11}L_{11}+h_{nn}L_n & h_{22}L_2+h_{33}L_3 & ... & 2h_{nn}L_{nn} \end{bmatrix} \tag{3.25}$$

$$L_i=\sqrt{\varphi(h)}\times\sum_{j=1}^{n}\left|b_{ij}\right|K_{ij},\ \varphi(h)=h_{\max}/h_{\min}\ h_{\min}=\min_{i=\overline{1,n}}\left\{h_{ii}\right\}, h_{\max}=\max_{i=\overline{1,n}}\left\{h_{ii}\right\}\ i=\overline{1,n}. \tag{3.26}$$

додатньо визначена. Тоді положення рівноваги $M_0\left(y_1^0, y_2^0, ..., y_n^0\right)$ системи (3.20) є стійким за Ляпунова.

Теорема 3.2 Нехай параметри $a_{ii} > 0$, $h_{ii} > 0$, b_{ij}, $K_{ij} > 0$, $i, j = 1, n$ системи (3.21) такі, що матриця

$$C_0(h, a, L, 0) = 2C_1(h.a) - \sqrt{\varphi(h)}C_2(h, L)$$

додатньо визначена. Тоді нульове положення рівноваги системи (3.21) буде асимптотично стійким і має місце наступна оцінка збіжності

$$\|x(t)\| \le \sqrt{\varphi(h)}\|x(0)\| \exp\{-\gamma_0 t/2\}, \ \|x(0)\| = \sqrt{\sum_{i=1}^{n} x_i^2(0)}. \qquad (3.27)$$

Тут $\gamma_0 > 0$ величина, за якої матриця

$$C_0(h, a, L, \gamma_0) = 2C_1(h, a) - \gamma_0 H - e^{\gamma_0 \tau/2}\varphi(h)C_2(h.L)$$

додатньо визначена.

Висновок

В наш час набуває популярності використання нейронних мереж для вирішення складних задач, які потребують величезних обчислювальних потужностей, через складність формальних моделей вивчення. Для вирішення таких задач використовують нейромережі великих розмірностей, які складно досліджувати, якщо спиратися тільки на загальні знання відносно теорії нейронних мереж. Процеси нейродинаміки описують нам ці складні процеси, але залишають відкритим питання дослідження їх збіжності.

В даній роботі описано методи дослідження збіжності процесів нейродинаміки. Використовуючи зннання з теорії стійкості, було наведено основні теореми та концепції щодо вирішення проблеми визначення умов збіжності , як глобальної так і асимптотичної стікості динамічних систем.

Також розглянуто поняття “системи зі слабкою нелінійністю”. За яким стало зрозуміло, що дослідження систем з виділеною лінійною частиною та нелінійністю, яка є обмеженою, зводиться до дослідження асипмтотичної стікості системи лінійного наближення (для “слабкої нелінійності”).

Розглянутим методом для дослідження умов збіжностей процесів нейродинаміки є метод функцій Ляпунова, який було застосовано до моделі нейронних мереж Хопфілда. Таким чином продемонстровані умови збіжності для систем з запізненням. Для цієї ж моделі було розглянуто відповідний клас нейронних мереж, для якого за допомогою теореми Коена-Гроссберга, були показані умови збіжності, використовуючи метод функцій Ляпунова.

Перелік джерел посилання

1. Simon Haykin, Neural Networks - A Comprehensive Foundation, Second Edition – 1998.
2. Куссуль Н.М., Шелестов А.Ю., Лавренюк А.М. Інтелектуальні обчислення. – Навчальний посібник, Київ, Наукова думка, 2006. – 186с.
3. Розенблат Ф. Принципы нейродинамики. – Москва, «Мир», 1965. – 480 с.
4. Борисюк Г. Н., Борисюк Р. М., Казанович Я. Б. и др. Модели динамики нейронной активности при обработке информации мозгом – итоги “десятилетия” // Успехи физических наук, 2002. – С.1189-1215.
5. А.І. Галушкіна, Нейронні мережі. Основи теорії; видавництво - Гаряча Лінія – Телеком, 2010.
6. Гітіс В.Б. Теорія і практика застосування нейронних мереж: посібник-Краматорськ, 2016.
7. Писаренко В.Г. Нова модель функціонування живої нейромережі, що враховує запізнілу взаємодію нейронів//Кібернетика та системний аналіз, №6, 2016. – С. 181-192.
8. Jack K. Hale, Theory of Functional Differential Equations, 1977 – 421 с.
9. Khusainov D.Ya., Diblik Y., Dfstinec Ja., Shatyrko A.V. Investigating Dynamics of One Weakly Nonlinear System with Delay Argument // Journal of Automation and Information Sciences, 2018. – pp.20-38.
10. Хусаінов Д.Я., Шатирко А.В., Бичков О.С., Шакотько Т.І. Стійкість та збіжність в моделях керуючих нейродинамічних систем. Актуальні проблеми теорії керуючих систем у комп’ютерних науках. Праці конференції, 21-24 грудня 2021 р., Слов’янськ. – С.121-126.
11. Шатирко А.В., Діблік Й., Хусаінов Д.Я., Баштинець Я. Збіжність процесів нейродинаміки в моделях Хопфілда// Штучний інтелект, 2017.
12. Барбашин Е.А. Введение в теорию устойчивости. – М., Наука, 1967. – 224 с.

13. Барбашин Е.А. Функции Ляпунова. – М., Наука, 1970. – 240 с.
14. Перестюк М.О., Чернікова О.С. Теорія стійкості. – ВПЦ "Київський університет", 2012. – 103 с.
15. Валєєв К.Г. Побудова функцій Ляпунова. – Київ, Наукова думка, 1981. – 412 с. (Ва 367980).