

дають (в Ом-м) 12 – 150 та 46,3. Виділені наступні різновиди ґрунтів: переважно слабо зволожені, незмінені, значення ρ 45 – 150, $\rho_{\text{ср}}$ 74,5; нормально зволожені, переважно незмінені, місцями слабо засолені до засолених, значення ρ 12 – 30 та $\rho_{\text{ср}}$ 22,2; локальне розповсюдження мають ґрунти перезволожені змінні і сильно забруднені (внаслідок розмиву міндобриг), значення ρ 0,4 – 2,2 та $\rho_{\text{ср}}$ 1,6.

В інтервалі глибин 1,5 – 4,0 м залягають неводонасичені лесовидні суглинки легкі. Згідно з узагальненими параметричними даними ρ та $\rho_{\text{ср}}$ суглинків складає відповідно 15 – 100 і 30,7, а коефіцієнт C , що визначає процентний вміст солей складає 0,27 – 0,036 %. Виділені три різновиди цих порід: менш зволожені, незасолені, значення ρ 50 – 120 та $\rho_{\text{ср}}$ 75, C до 0,12 %; частково змінні, незасолені і дуже слабо засолені, значення ρ 35 – 45, $\rho_{\text{ср}}$ 37,5, C 0,14 – 0,08 %; більш змінні і різною мірою засолені, C 0,6 – 0,12 %, ρ 12 – 36, $\rho_{\text{ср}}$ 22,2. Зустрічаються, але значно рідше, і різновиди сильно засолених (забруднених) неводонасичених лесовидних суглинків із значеннями ρ 8 – 12 та C 0,55 – 0,36.

Зміна складу і властивостей порід зони аерації та горизонту ґрунтових вод, викликані дією техногенних факторів призводять до значної зміни фізичних параметрів по-

рід. Внаслідок цього електричний опір та інші геофізичні параметри ґрунтів та порід, що складають ВЧР, залежать не стільки від первинного складу, скільки від їх стану – зволоженості (водопроникності), ступеню техногенних змін, засолення (забруднення), мінералізації вод, що створює фізичні передумови застосування геофізичних методів досліджень (в першу чергу електрометричних). Отримані результати свідчать про високу ефективність геофізичних методів при геоекологічних дослідженнях.

1. Вижва С.А. Геофізичний моніторинг небезпечних геологічних процесів. – Київ.: "Обрії", 2004. – 236 с. 2. Онищук І.І. Особливості застосування геофізичних методів при геоекологічних дослідженнях в Дарницькому районі м. Києва. // Вісник Київ. ун-ту. Геологія. – 2001. – Вип.19. – С.42-46. 3. Онищук І.І. Геоелектричне вивчення підтоплення центральної частини Національного комплексу "Експоцентр України". // Вісник Київ. ун-ту. Геологія. – 2002. – № 21. – С. 61-63. 4. Онищук І.І. Застосування екогеофізичних досліджень при вивченні підтоплених ґрунтів. Вісник Київ. ун-ту. Геологія. Вип. 20. Київ, 2001. с. 67 – 70. 5. Онищук І.І., Рева М.В. Принципи геофізичного картування стану земель. Вісник Київ. ун-ту. Геологія. – 2004. – №29 – 30. – С.36-40. 6. Вижва С.А., Онищук І.І., Безродний Д.А. Екогеофізичний моніторинг міських агломерацій // Вісник Київ. ун-ту. Геологія. – 2003. – №25. – С. 71-75. 7. Огильви А.А. Основы инженерной геофизики. М. Недра., 1990. – с.324-378.

УДК 550.837:622.241

В.М. Курганський, д-р геол.-мінералог. наук, проф.
І.В. Тішаєв, асист.

ДО ПИТАННЯ ЗНИЖЕННЯ РОЗМІРНОСТІ ДАНИХ В ГЕОЛОГО-ГЕОФІЗИЧНИХ ЗАДАЧАХ КЛАСИФІКАЦІЇ

В статті розглядаються питання зниження розмірності даних в контексті розв'язання широкого кола геолого-геофізичних задач класифікації. На прикладах роботи з модельними даними демонструються переваги та недоліки лінійного і нелінійного методів головних компонент.

In article briefly review the questions of data dimensionality reduction as applied to solving classification geological-geophysical problems. Linear and nonlinear principal-factor methods work showing on the model examples.

Актуальність питання. В практиці геолого-геофізичних досліджень дуже часто виникають різноманітні задачі класифікації. Зокрема, для такої галузі науково-практичної діяльності людини, як промислова геофізика, класифікаційні задачі навіть превалюють. Взагалі ж, якщо підходити до проблеми широко, то практично всяку задачу можна сформулювати як задачу класифікації, і не тільки в геологічній галузі. В геологічних дисциплінах розв'язок подібних задач зводився до узагальнення та візуалізації усього наявного фактичного матеріалу (у вигляді карт, розрізів, крос-плотів, гістограм), і фахівець-інтерпретатор на основі власного практичного досвіду визначав критерії та виробляв вирішальні правила для кожної конкретної ситуації. Такий комплексний аналіз усіх доступних матеріалів фактично означав опис геологічного середовища, що вивчалось, набором геолого-геофізичних (-геохімічних, -петрофізичних, -мінералогічних та інш.) ознак. Таким чином, говорячи формальною мовою, створювався простір ознак, кожна точка якого відповідала певному стану досліджуваного середовища. Розмірність простору може бути як завгодно великою, при чому інформативна значущість окремих ознак при цьому обов'язково буде різною. В цьому і полягають головні труднощі автоматизації процесу вирішення класифікаційних задач, а саме – у неможливості алгоритмізувати здатність людського мозку виявляти найбільш інформативні ознаки у контексті пошуку розв'язку конкретної задачі.

На фоні стрімкого розвитку засобів збору, реєстрації, обробки та зберігання величезних об'ємів даних в геологічній галузі дуже добре відчувається відсутність

відповідного методичного та програмно-алгоритмічного забезпечення, яке дозволило б проводити максимально ефективну обробку цих даних. Мається на увазі назріла необхідність створення систем обробки та аналізу різноманітної інформації, здатних моделювати (або імітувати) ту послідовність логічних дій, якою користується людина при розв'язанні складних задач розпізнавання образів (до яких відносяться і згадані вище задачі класифікації). Дана робота присвячена окремому важливому питанню, нерозривно пов'язаному з означеними проблемами інтеграції інформаційних технологій в геологічну сферу. Мова йде про методи зниження розмірності простору ознак. Вдалий (у сенсі збереження інформативної насиченості даних) перехід від початкового простору до простору з меншою розмірністю дає багато переваг при використанні різноманітних класифікаційних процедур [1, 2], які сьогодні є добре відомими та широко використовуються в геологічних науках, принаймні у найбільш формалізованих з них – геофізиці. З огляду на вищесказане, можна стверджувати про актуальність та своєчасність піднятого питання.

Стан розробки проблеми. Щільності розподілів даних, що зустрічаються на практиці, дуже рідко відповідають стандартним параметричним щільностям розподілів, які ґрунтовно досліджені в теоретичному плані, і для яких розроблені ефективні параметричні процедури класифікації. Тому багатьма дослідниками була запропонована велика кількість різних непараметричних методів, серед яких окремо можна виділити процедури,

призначені для перетворення простору ознак таким чином, щоб у перетвореному просторі можна було використовувати параметричні методи.

Добре відомий метод лінійного дискримінанта Фішера, головна ідея якого полягає у зменшенні розмірності вихідного простору з n вимірами шляхом проектування n -вимірних даних на пряму, тобто у представленні даних у одновимірному просторі. Обертаючи цю пряму у вихідному n -вимірному просторі, можна знайти таку орієнтацію, для якої спроектовані на пряму дані будуть добре розділені. Саме це є метою класичного дискримінантного аналізу [1].

Також класичними процедурами є аналіз головних компонент та факторний аналіз. Обидва названі методи зменшують розмірність шляхом формування лінійних комбінацій ознак. Метою аналізу головних компонент (відомого в літературі по теорії зв'язку як розклад Кархунена-Лоева) є ранжування ознак за величиною їх дисперсії. Простір меншої розмірності утворюється шляхом простого відкидання тих ознак, дисперсія яких є меншою за деякий встановлений поріг. Метою факторного аналізу є знаходження представлення у просторі меншої розмірності на основі врахування кореляції між ознаками [1, 3, 4]. В результаті дістають: оцінку кількості незалежних факторів, які керують розподілом даних; факторні навантаження, які характеризують вплив факторів; головні компоненти, які опосередковано відображають фактори через нормовані на власну дисперсію ознаки [5].

Інший відомий підхід до представлення даних у просторі меншої розмірності полягає у перетворенні ознак таким чином, щоб вони утворювали ортогональну систему. Це свідчить про їхню взаємну незалежність. Класичний процес ортогоналізації Грама-Шмідта полягає у наступному: по системі лінійно незалежних ненульових векторів (що відповідають окремим ознакам) $x_1, x_2, \dots, x_n$ рекурентно будується система ортогональних векторів $h_1, h_2, \dots, h_n$. В якості першого вектору h_1 вибирається вихідний вектор x_1 . Кожен наступний вектор робиться ортогональним усім попереднім, для чого від нього віднімаються його проекції на усі попередні вектори. При цьому, якщо деякий з векторів h_i виявляється рівним нулеві, то він відкидається. Можна показати, що за побудовою отримана система векторів виявляється ортогональною, тобто кожному вектору відповідає певна унікальна ознака [6].

Представлені методи зниження розмірності простору ознак, оперують лише лінійними перетвореннями, а отже можуть бути корисними для розв'язку вузького кола задач. Вочевидь, у випадку "складного" розподілу даних у вихідному просторі ознак лише лінійні перетворення останніх можуть не дати бажаного результату.

Постановка та розв'язок задачі. На противагу описаним вище методам розглянемо процедуру зниження розмірності простору ознак на основі нейромережових методик.

Штучні нейронні мережі (ШНМ) сьогодні розглядаються як високоефективні алгоритми розв'язання різноманітних задач, що важко піддаються формалізації. Елементарним елементом ШНМ є нейропроцесор, який функціонує за наведеною на рис.1 схемою. Тут в якості складових частин нейропроцесору виділяються: вектор вагів w , розмірність якого співпадає з розмірністю вхідного вектора x ; блок сумування NET, який, згідно виразу (1), являє собою не що інше, як скалярний добуток векторів x та w ; блок нелінійного перетворення OUT, в якому відбувається функціональне перетворення над скалярним аргументом $*$ (формула (2)), при чому функція $F(*)$, яка називається активаційною функцією нейрона, є нелінійною з обмеженою областю значень (як зверху, так і знизу) [7]. Приклади активаційних функцій наведені на рис.2. В результаті усіх перелічених дій на виході блоку нелінійного перетворення утворюється вихідний вектор y (в даному випадку матимемо однокомпонентний вектор, тобто скаляр).

Класичною архітектурою штучної нейронної мережі є модель багатшарового персептрону без зворотніх зв'язків. В такій моделі окремі нейропроцесорні елементи (які в літературі також позначаються як А-елементи) об'єднуються в шари; сукупність шарів y утворює архітектуру під назвою "багатшаровий персептрон". При цьому, передача інформації відбувається послідовно від шару до шару: сукупність виходів окремих нейронів попереднього шару являє собою вхідний вектор для усіх нейронів наступного шару. Приклад такої мережі представлений на рис.3.

$$NET = \sum_i w_i x_i \quad (1)$$

$$OUT = F(NET - \theta)$$

де θ – пороговий рівень даного нейрону.

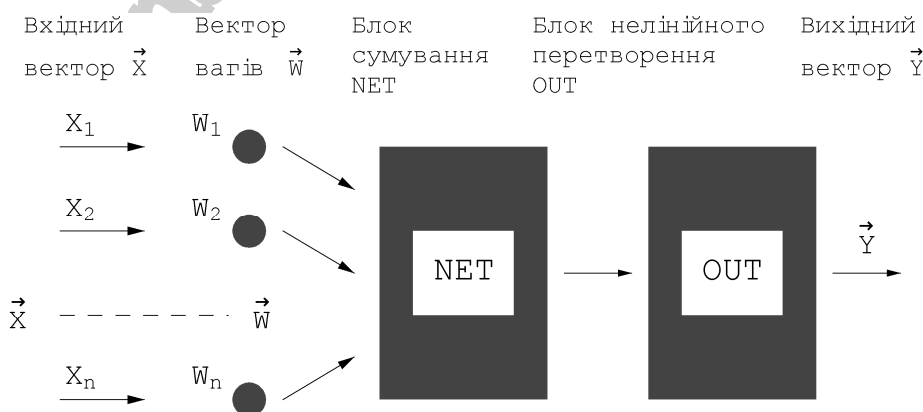


Рис.1. Функціональна схема математичної моделі елементарного нейропроцесору

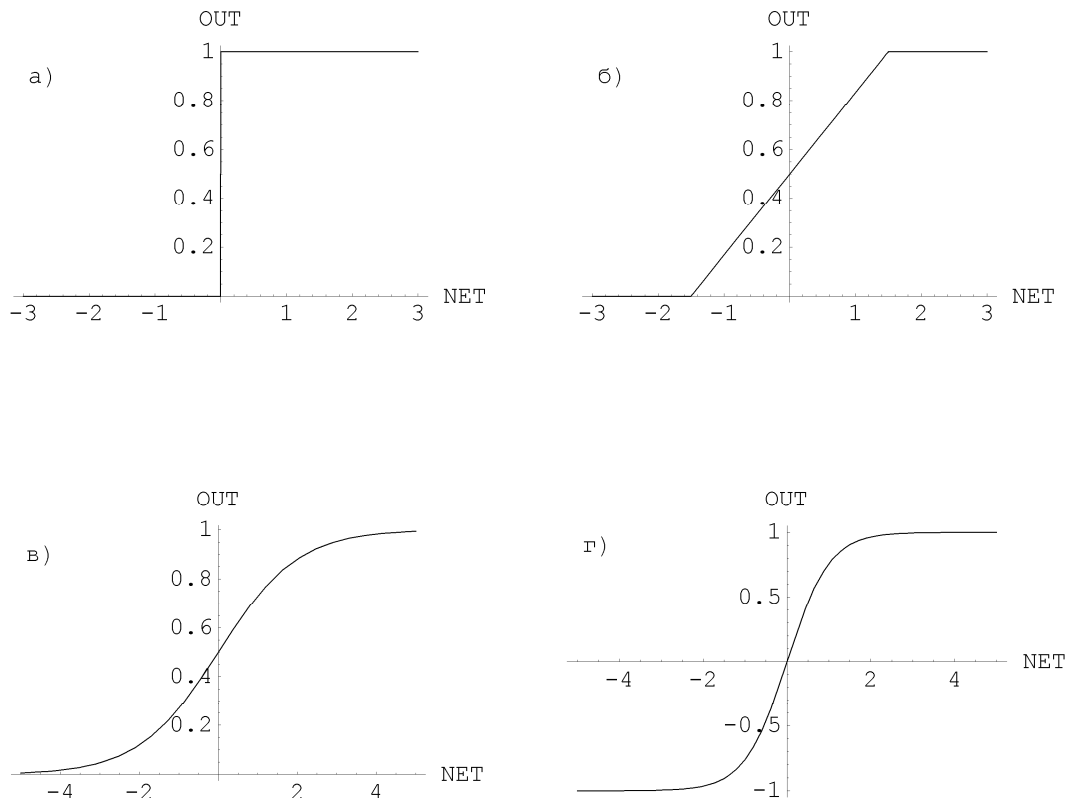


Рис.2. Приклади активаційних функцій: а) функція одиничного стрибку; б) похила сходінка; в) сигмоїда (логістична функція); г) гіперболічний тангенс.

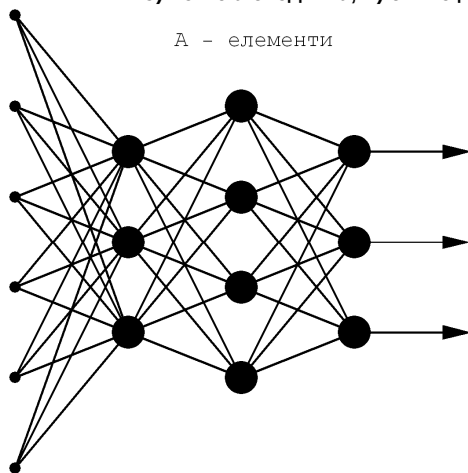


Рис.3. Приклад тришарової нейронної мережі. Тут: розмірність вхідного вектору дорівнює 6; елементарні нейропроцесорні елементи (А-елементи) зображені у вигляді великих чорних вузлів; вхідний шар нейромережі містить 3 нейрони, прихований – 4, вихідний шар містить 3 нейрони; розмірність вихідного вектору відповідає кількості нейронів у останньому (вихідному) шарі нейромережі й дорівнює 3.

В загальному випадку роботу наведеної на рис.3 нейромережі можна представити у наступному вигляді:

$$F(F(F(x \cdot W_1) \cdot W_2) \cdot W_3) = y \quad (3)$$

де: x – вхідний вектор; y – вихідний вектор; W_i – матриця вагових коефіцієнтів нейронів i – го шару мережі; $F(*)$ – активаційна функція, при чому

$$F(a) = \{F(a_1), F(a_2), \dots, F(a_n)\}.$$

Найбільш корисною і привабливою властивістю ШНМ є їхня здатність до навчання. Річ у тім, що фізична, або логічна, сутність вихідного вектору y ви-

значається лише зручністю представлення розв'язку задачі. Тоді, подаючи на вхід нейромережі вектори, для яких апріорно відома їхня належність до того чи іншого класу, можна вимагати появи на виході мережі цілком певного вихідного вектору, який і буде індикатором належності вхідного вектору до даного класу. Для забезпечення такого режиму роботи ШНМ необхідно настроїти вагові коефіцієнти та порогові рівні нейронів – навчити мережу. Тут найбільшої популярності набув метод навчання, відомий в літературі як метод зворотнього розповсюдження помилки (error back propagation method).

Згідно формальній постановці задачі [7] багатшаровий персептрон розраховує вихідний вектор y для довільного вхідного вектору x . Відповідно, умовою будь-якої задачі, яка може бути поставлена персептрону, повинна виступати множина векторів $X = \{x^1, \dots, x^S\}$ з n компонентами кожний. Розв'язком задачі буде множина векторів $Y = \{y^1, \dots, y^S\}$ з m компонентами кожний. Таким чином персептрон формує відображення $X \mapsto Y$ для $\forall x \in X$.

Багатшаровий персептрон навчається із вчителем. Це означає, що повинна бути задана множина пар векторів $\{x^s, d^s\}$, $s = \overline{1, S}$, де $\{x^s\} = \{x^1, \dots, x^S\}$ – формалізована умова задачі, $\{d^s\} = \{d^1, \dots, d^S\}$ – відомий розв'язок цієї задачі. Множина пар векторів $\{x^s, d^s\}$ складає навчальну множину. S – кількість елементів навчальної множини. Кількість пар у навчальній множині не регламентується, проте, якщо їх буде занадто мало або багато, мережа не навчиться і не буде здатною розв'язати поставлену задачу. Якщо вибрати один з векторів x^s і подати його на вхід нейромережі, то на виході буде отриманий деякий

вихідний вектор y^s . Тоді помилкою нейромережі можна вважати $E^s = \|d^s - y^s\|$.

Основна ідея зворотнього розповсюдження помилки полягає у можливості рекурсивного отримання значення помилки для нейронів прихованих шарів. Сенс навчання полягає у мінімізації функціонала квадратичної помилки мережі для даного набору навчальних прикладів:

$$E = \frac{1}{2} \sum_{s=1}^S \sum_{i=1}^m (d_i^s - y_i^s)^2. \quad (4)$$

Нейронна мережа може бути повністю охарактеризована вектором параметрів – набором усіх вагових коефіцієнтів та порогових рівнів:

$$P = \begin{pmatrix} W \\ \theta \end{pmatrix}. \quad (5)$$

Тоді, помилка нейромережі E буде функцією вектору параметрів:

$$E = E(P). \quad (6)$$

При навчанні нейромережі компоненти вектору параметрів корегуються в напрямку антиградієнта помилки E :

$$\Delta P = -\varepsilon \nabla E(P). \quad (7)$$

В теорії оптимізації доведено, що такий алгоритм забезпечує збіжність до одного з локальних мінімумів функції помилки, за умови правильного вибору $\varepsilon > 0$ на кожній ітерації. Цей метод оптимізації називається методом найшвидшого спуску. Алгоритм зворотнього розповсюдження – спеціальний спосіб швидкого (з малими обчислювальними витратами) розрахунку компонент вектору градієнта. Ідея методу полягає у тому, щоб представити E у вигляді складної функції та послідовно розраховувати частинні похідні за формулою для складної функції. Розрахунок виконується від останнього (вихідного) шару до першого (вхідного) за рекурентними формулами і не потребує безпосереднього обчислення градієнту.

Згідно наведених міркувань можна створити ШНМ такої архітектури, що передбачала б наявність прихованого шару, кількість нейронів у якому менша за розмірність вхідного вектору, а кількість нейронів у вихідному шарі задати рівною розмірності вхідного вектору. Якщо при цьому навчити нейромережу від-

творювати на виході безпосередньо вхідні вектори, то така ШНМ буде називатися *автоасоціативною*. Особливості архітектури цієї мережі (рис.4) дозволяють отримати на виході одного з прихованих шарів вектор, розмірність якого буде меншою за розмірність вхідного вектору.

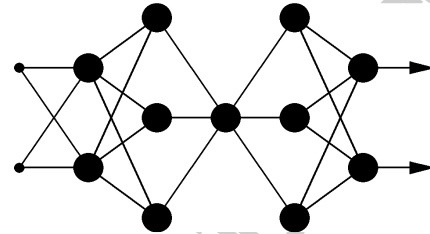


Рис.4. Автоасоціативна нейронна мережа (мережа з "вузьким горлом")

Отже, якщо при даній архітектурі нейромережі спроможтись отримати на виході безпосередньо вхідний вектор, то це означатиме, що вхідні дані були оптимальним чином стиснені у прихованому шарі (з одним нейроном) мережі, тобто при зниженні розмірності даних їхня інформативність практично не змінилася. На цьому принципі базується зниження розмірності простору ознак за допомогою ШНМ. Необхідно зазначити, що стискаючим прихованим шаром генеруються якісно нові ознаки, що майже ніколи не співпадають з початковими ознаками, але при умові якісного навчання мережі їхня інформативна насиченість зберігається.

Приклади

Розглянемо принцип роботи автоасоціативної нейронної мережі у порівнянні з методом головних компонент у контексті розв'язання задачі зниження розмірності простору ознак. На рис.5 наведені два модельні приклади: в одному випадку маємо п'ять лінійно роздільних класів (груп, кластерів), при чому центри класів лежать на одній прямій; в другому випадку загальний розподіл усієї сукупності даних є більш складним, хоча лінійна роздільність класів (чотири класи) збережена. Для обох випадків прийнятий нормальний розподіл точок даних в межах кожного окремого класу.

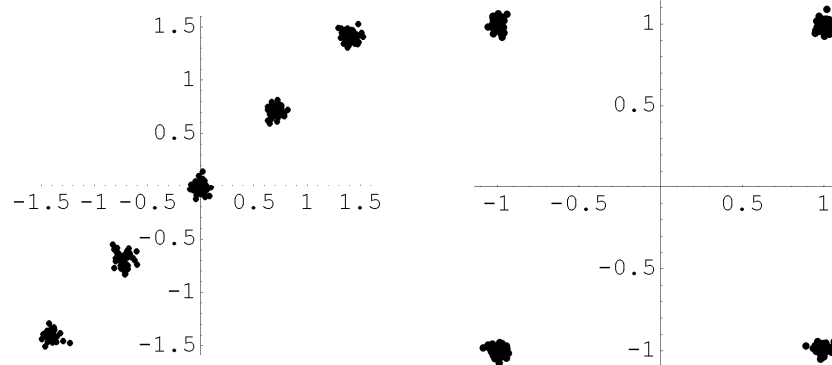


Рис.5. Два модельні приклади розташування окремих класів у просторі ознак. Класи є лінійно роздільними. Загальна сукупність даних в обох випадках центрована та нормована.

Як можна бачити, ідентифікація окремих класів є цілком очевидною, й побудова класифікатора (вирішальних правил) не буде викликати труднощів. Візьмемо за мету зменшення розмірності простору ознак. При цьому критерієм якості переходу до нового простору буде збереження кількості класів за результатами класифікації у новому просторі ознак.

Розглянемо спочатку дію метода головних компонент в обох випадках.

Аналізуючи коваріаційну матрицю даних першого прикладу, можна побачити, що обидві координати (ознаки) є добре корельовані між собою – недиагональні елементи коваріаційної матриці близькі до одиниці:

$$\begin{pmatrix} 1 & 0.997312 \\ 0.997312 & 1 \end{pmatrix}$$

Справді, в даному випадку центри усіх кластерів лежать на одній прямій, а отже, можна казати навіть про функціональну залежність ознак (рис.6, а). Перейдемо тепер до нової системи координат, напрямки вісей якої

співпадають з почергово визначеними напрямками найбільших дисперсій усієї сукупності даних. Коваріаційна матриця даних в такій координатній системі матиме вигляд:

$$\begin{pmatrix} 1.99731 & 3.27048 \times 10^{-16} \\ 3.27048 \times 10^{-16} & 0.00268765 \end{pmatrix}$$

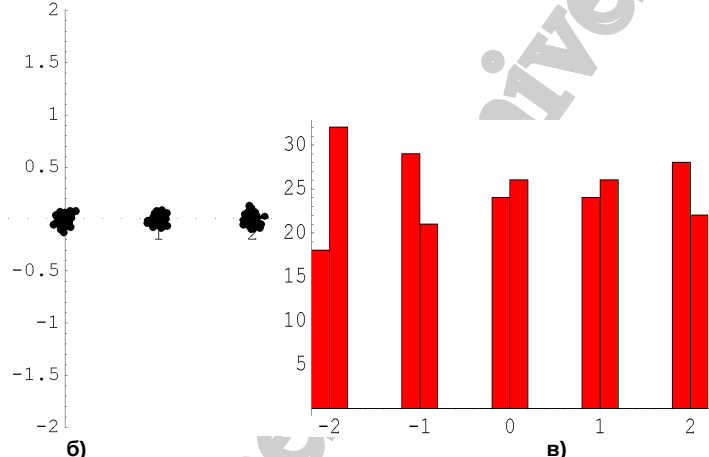
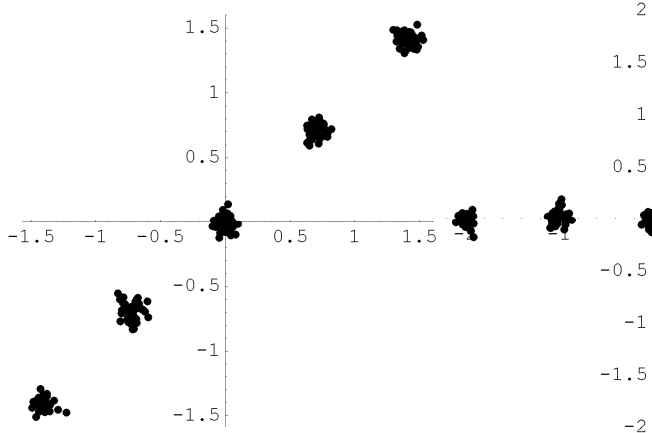


Рис.6. а) модельний розподіл двовимірних даних; б) перетворення даних до головних компонент; в) гістограма перетворених даних.

Як і слід було очікувати, напрямком найбільшої дисперсії даних співпав з прямою, на якій лежать центри кластерів. Це є перша головна компонента, уздовж якої спрямовується перша координатна вісь (ознака) даних в новій системі координат (рис.6, б). Друга координатна вісь спрямована у напрямку максимальної дисперсії перпендикулярно до першої головної компоненти. Близькі до нуля значення недиагональних компонент коваріаційної матриці кажуть про незалежність ознак, а співвідношення між першою і другою головними компонентами (майже 10^3) свідчить про неінформативність останньої. Отже, можна просто відкинути другу ознаку і представити дані у одновимірному просторі практично без втрат інформації. Дійсно, побудувавши гістограму одновимірних даних, побачимо чітку відповідність характерних максимумів окремим класам (рис.6, в).

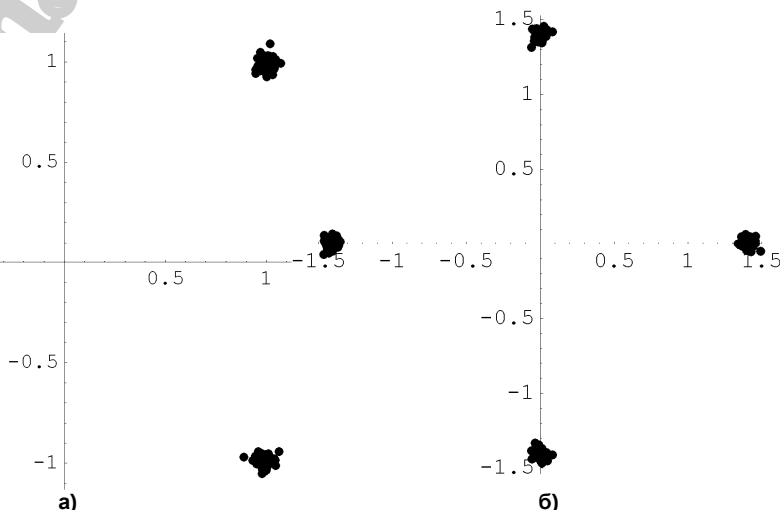
В другому випадку (рис.7, а) загальна структура даних є більш складною у тому сенсі, що за оцінкою функції розподілу неможливо виокремити окремі кластери. Коваріаційна матриця вказує на незалежність ознак, якими характеризуються дані:

$$\begin{pmatrix} 1 & 0.00385628 \\ 0.00385628 & 1 \end{pmatrix}$$

В цьому випадку можна очікувати, що дисперсії, узяті за любыми напрямками будуть незначно відрізнятися. Перевіримо це припущення за допомогою метода головних компонент. Коваріаційна матриця даних в новій координатній системі матиме вигляд:

$$\begin{pmatrix} 1.00386 & -3.52872 \times 10^{-17} \\ -3.52872 \times 10^{-17} & 0.996134 \end{pmatrix}$$

Як можна бачити, діагональні компоненти практично не змінилися, що вказує на їхню однакову інформативність. Отже, в даному випадку не можливо знизити розмірність даних, і метод головних компонент є зовсім неідеальним, хоча усі класи й продовжують залишатися лінійно роздільними (рис.7, б). Аналіз гістограм трансформованих даних (рис.7, в та рис.7, г) також не дозволяє зробити правильного висновку стосовно кількості класів, що складають представлену вибірку даних.



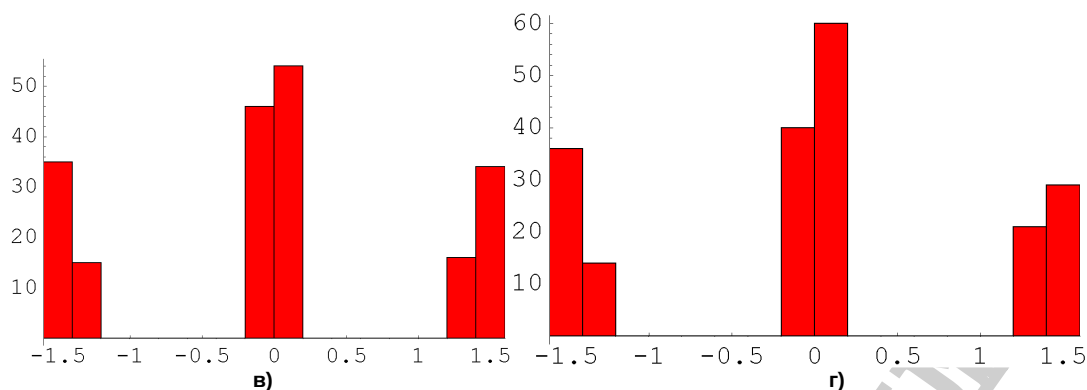


Рис.7. а) модельний розподіл двовимірних даних; б) перетворення даних до головних компонент; в)-г) гістограми перетворених даних уздовж вісей X та Y відповідно.

Спробуємо тепер побудувати та навчити згідно наведених вище міркувань автоасоціативну штучну нейронну мережу, архітектура якої наведена на рис.4. В якості функції активації використаємо гіперболічний тангенс (рис.2, г). Навчальну множину складемо з наборів векторів, що являють собою незначно ускладнені випадковими завданнями (з рівномірним розподілом похибки) вектори математичних сподівань кожного з кластерів на вході та "чисті" вектори математичних сподівань кластерів на виході. Будемо проводити навчання ШНМ за допомогою методу зворотнього розповсюдження похибки. Необхідно зазначити, що для автоасоціативних мереж проблема застрягання помилки навчання E в локальному мінімумі повер-

хні можливих розв'язків є надзвичайно актуальною, і навіть для таких простих прикладів знадобилося витратити досить багато часу для отримання прийнятних варіантів параметрів нейромереж.

Оскільки перший приклад є досить тривіальним й дуже просто може бути оброблений звичайним лінійним методом головних компонент, то розглянемо роботу ШНМ тільки на другому (рис.7, а) прикладі. На рис.8 наведені схематичне зображення принципу стиснення даних за допомогою мережі з "вузьким горлом" та гістограма розподілу одновимірних даних, отриманих на виході прихованого стискуючого шару нейромережі.

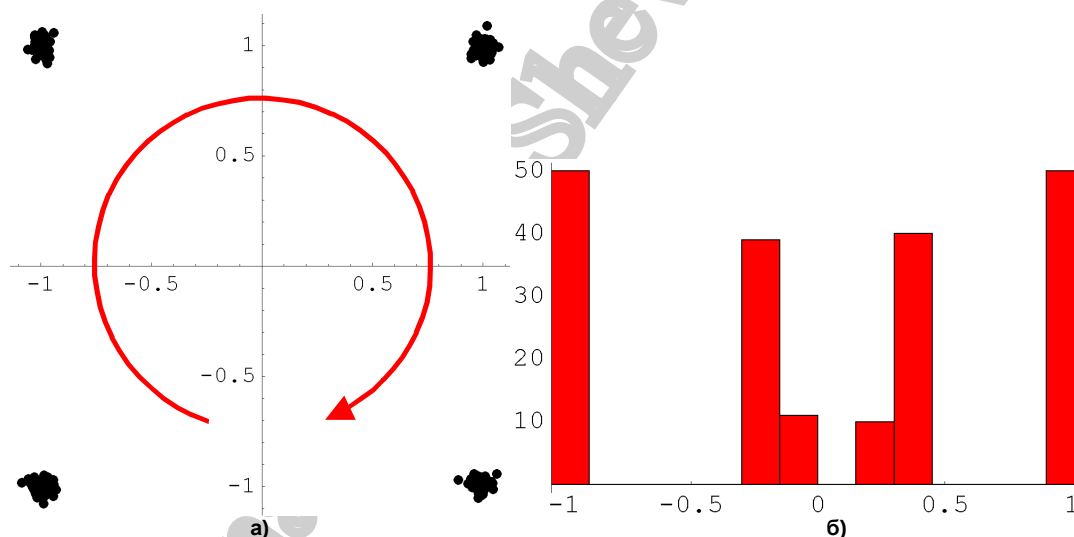


Рис.8. а) автоасоціативна нейромережа дозволяє визначати нелінійні напрямки, в яких дисперсія даних є найбільшою; б) гістограма одновимірних даних, спроектованих на певну (нелінійну!) траєкторію, на якій зберіглася лінійна роздільність окремих класів.

На вході ШНМ – вектори з тестового прикладу. На виході – тіж самі вектори, відтворені мережею. Це означає, що в процесі передачі інформації від нейрона до нейрона вона трансформувалася, але не втрачалася. А оскільки особливості архітектури автоасоціативної нейромережі передбачають наявність прихованого шару з кількістю нейронів, меншою за розмірність вхідних даних, то виходи цього шару можна інтерпретувати як оптимальним чином стиснуті вхідні дані. Таким чином відбувається зниження розмірності даних без втрати інформації. При цьому напрямки найбільших дисперсій даних, які виділяє нейронна мережа, можуть бути представлені довільними кривими. Фактично, нейромережа із стискуючим прихо-

ваним шаром реалізує нелінійний метод головних компонент [8].

Висновки. Головна перевага використання ШНМ полягає у можливості значного зниження розмірності даних без суттєвих втрат інформації, що дає змогу, наприклад, візуалізувати багатовимірні дані, а також проводити їх обробку із значно меншими витратами машинних ресурсів. В той же час, неможливість задання в якості навчальної множини довільної вибірки векторів є принциповою. Звідси – втрати інформації при стисненні даних, а отже, і при їх відновленні.

Важливим негативним моментом в практичній реалізації шнм є технічні труднощі проведення процесу навчання нейромережі. Процес навчання являє собою по суті задачу багатопараметричної оптимізації, і на сьогодні не існує чітких критеріїв та алгоритмів, які б гарантували

успішність навчання мережі за скінчену кількість кроків. Тут велику роль відіграє досвід дослідника.

Якщо зуміти провести навчання автоасоціативної мережі на достатньо репрезентативній вибірці двовимірних даних (яка описує усю множину можливих значень даних), то, вочевидь, можна побудувати універсальний кодер-декодер із гарним коефіцієнтом стиснення, рекурсивно подаючи виходи стиснутого шару нейромережі знову їй на вхід. На жаль, практична реалізація цієї оригінальної авторської ідеї поки ще не здійснена.

1. Р.Дуда, П.Харт. Распознавание образов и анализ сцен. Москва: Мир, 1976. – 512 с. 2. Б.Дюрран, П.Оделл. Кластерный анализ. – Москва: Статистика, 1977. – 128 с. 3. Никитин А.А. Теоретические основы обработки геофизической информации. – Москва: Недра, 1986. – 343 с. 4. Вычислительная математика и техника в разведочной геофизике / Справочник геофизика под ред. В.И.Дмитриева. – Москва: Недра, 1990. – 499 с. 5. М.Н.Жуков. Статистический анализ геологических данных (довідкові відомості, задачі, приклади розв'язання): Навч. посібник. – Київ: ІСДО, 1995. – 552 с. 6. Терехов С.А. Лекции по теории и приложениям искусственных нейронных сетей. http://alife.narod.ru/lectures/neural/Neu_index.htm 7. Заенцев И.В. Нейронные сети: основные модели. – Воронеж: ВГУ, 1999. <http://www.scintific.narod.ru/NNBooks/nncourse.zip> 8. А.А.Ежов, С.А.Шумский. Нейрокомпьютеринг и его применения в экономике и бизнесе. Серия "Учебники экономико-аналитического института МИФИ" под ред. проф. В.В. Харитонова. Москва: МИФИ, 1998, – 224 с. <http://neurolec.chat.ru/>

УДК 550.831

ДОСЛІДЖЕННЯ ТЕКТОНІЧНИХ ЗОН ОБЕРНЕНИМИ ЛІНІЙНО-НЕЛІНІЙНИМИ МЕТОДАМИ ГРАВІМЕТРІЇ ДЛЯ ВИБОРУ МІСЦЬ БУДІВНИЦТВА ЕКОЛОГІЧНО НЕБЕЗПЕЧНИХ ОБ'ЄКТІВ

Розроблено ряд методів розв'язку обернених задач гравіметрії, комплексовання яких є необхідною умовою для ефективного вивчення тектонічних зон в масивах гірських порід кристалічного фундаменту.

Methods of the decision of return problems gravity which complex use is a necessary condition for effective studying tectonic zones in files of rocks of the crystal base are developed.

Стан проблеми. Будівництво електростанцій, аеродромів, дамб, відстійників потребує детального вивчення характеру монолітності масивів гірських порід у кристалічному фундаменті, якщо потужність наносів невелика. У таких випадках розташування екологічно небезпечних об'єктів над широкими тектонічними зонами фундаменту являє загрозу як для навколишнього середовища, так і для самого об'єкта. Якщо тектонічні зони не заповнені дайками магнітних порід, їхнє вивчення практично можливе тільки дорогими свердловинами і тільки в приповерхній частині – методами електророзвідки.

Однак у даний час уже відомі ефективні і стійкі методи розв'язку обернених лінійних (а також і нелінійних) задач гравіметрії (ОЛЗГ, ОНЗГ), що дають надійні фізично і геологічно змістовні представлення про структуру геологічних об'єктів взагалі і тектонічних зон зокрема [1,2].

Метою цієї роботи є розробка оптимізуємих методів розв'язку ОЛЗГ і ОНЗГ для більш точного вилучення впливу великих масивів гірських порід зі спостереженого поля і забезпечення упевненого трасування тектонічних зон по залишкам поля сили тяжіння.

Поставлена мета досягається тим, що застосовується методика інтерпретації карти гравітаційного поля з почерговим використанням розв'язків оберненої задачі декількома методами [3].

Методи досліджень. Спочатку ОЛЗГ розв'язана методом ітерацій для одношарової моделі з апроксимацією нижнього півпростору на 20х20 прямокутних паралелепіпедів (блоків) невеликої висоти і на 4 обмежених горизонтальних уступи великої потужності, довжини і ширини з лінійно змінною з глибиною та невідомою аномальною щільністю x_i , що також підлягає визначенню одночасно з аномальною щільністю всіх паралелепіпедів. Проекції 400 паралелепіпедів цілком покривають площу карти гравітаційного поля g_j таким чином, що над кожним з них є хоча б один пункт з вимірюваним в ньому полем сили тяжіння, і під кожним таким пунктом є хоча б один паралелепіпед. Це умова коректності оберненої лінійної задачі гравіметрії. Чим більше пунктів поля перебуває за межами проекцій блоків на площину карти поля, і чим більше блоків розташовано так, що їх проекції не попадають на карту гравітаційного поля, тим більше некоректною є модель ОЛЗГ і тим більше не-

стійким і фізично незмістовним буде її розв'язок. Система уступів використовується для автоматичного визначення і вилучення практично всього регіонального фону. Кількість уступів може бути будь-якою, але над кожним з них повинні бути пункти поля сили тяжіння. Застосовуються, в основному, безумовні алгоритми з кількістю ітераційних параметрів τ_i , рівною кількості блоків і невідомих компонентів $x_{k,i}$ змінної щільності x_i уступів у моделі задачі. Метод і алгоритм розв'язку ОЛЗГ для цієї частини задачі має вигляд.

П.О. Міненко, канд. геол.-мінералог. наук
Р.В. Міненко, студ.

Система уступів використовується для автоматичного визначення і вилучення практично всього регіонального фону. Кількість уступів може бути будь-якою, але над кожним з них повинні бути пункти поля сили тяжіння. Застосовуються, в основному, безумовні алгоритми з кількістю ітераційних параметрів τ_i , рівною кількості блоків і невідомих компонентів $x_{k,i}$ змінної щільності x_i уступів у моделі задачі. Метод і алгоритм розв'язку ОЛЗГ для цієї частини задачі має вигляд.

$$x_{i,n+1} = x_{i,n} - \tau_{n+1} * B_{i,n}; \quad x_{i,n+1} = x_{i,n} - \tau_{i,n+1} * B_{i,n};$$

$$\lambda_i = \sum_j a_{ij}; \quad \lambda_j = \sum_i a_{ij}; \quad (1)$$

$$r_{j,n} = (a_{ij}, x_{i,n}) - g_j; \quad r_{j,n+1} = r_{j,n} - (a_{i,j}, B_{i,n} * \tau_{i,n+1});$$

$$B_{i,n} = (a_{i,j} / \lambda_i / \lambda_j, r_{j,n});$$

$$F_4 = \sum_i (B_{i,n+1})^2 \xrightarrow{\tau_i \neq const} \min;$$

Ітераційні параметри знаходять шляхом розв'язку системи 408 рівнянь на кожному кроці ітерацій. В більшості випадків для розв'язку ОЛЗГ методом (1) досить 3 ітерацій за 3 хвилини.

Далі залишкове поле інтерпретується шляхом розв'язку ОЛЗГ для моделі з аномальною щільністю, що змінюється з глибиною z_i усередині кожного паралелепіпеда за законом полінома четвертого ступеня. Можлива також добавка до полінома лінійних членів для горизонтальних напрямків. Таким чином, для кожного паралелепіпеда маємо від 5 до 7 компонентів параметра змінної щільності. При тих же 400 блоках і 4 уступах маємо 2808 невідомих для карти поля сили тяжіння g_j , вимірюваного в 65х53=3445 пунктах.

Застосовуються, в основному, безумовні алгоритми з одним ітераційним параметром τ_c для всіх невідомих компонентів $x_{k,i}$ перемінної щільності x_i блоків і уступів.

$$x_{i,n+1} = x_{i,n} - \tau_{n+1} * B_{i,n}; \quad \lambda_i = \sum_{j,k} a_{k,i,j}; \quad \lambda_j = \sum_i a_{k,i,j};$$

$$x_{i,n} = \sum_{k=0}^{k=m} x_{k,i,n} z_i^k; \quad (2)$$