УДК 004.8
DOI: 10.31673/2412-9070.2021.063540

В. В. ЖЕБКА, доктор техн. наук, професор;
В. О. КОРЕЦЬКА, канд. пед. наук, доцент,
Державний університет телекомунікацій, Київ

УДОСКОНАЛЕННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ НА ОСНОВІ МЕТОДУ МЕТРИЧНОГО ПРОКСИМАЛЬНОГО ГРАДІЄНТА З УВЕДЕННЯМ ДІАГОНАЛЬНОГО КРОКУ

Розглянуто окремі аспекти використання методів оптимізації в телекомунікаційних мережах. Особливо актуальне використання оптимізаційних методів засобами машинного навчання з метою уникнення різних позаштатних ситуацій в мережах. Доцільно застосування методів машинного навчання під час отримання інформації про якість сигналу, трафіку тощо. Водночас можна робити прогнози різного роду несправностей, маршрутизації, контролю безпеки.

Визначено, що в процесі моделювання неоднорідних мереж ефективне використання моделі марковського випадкового поля. Такий підхід дає можливість моделювати експоненціальний розподіл вузлів у гетерогенних мережах.

Наведено модифікацію алгоритму проксимального градієнта — метод змінного метричного проксимального градієнта. Забезпечення швидкої збіжності досягається засобами діагонального розміру кроку, який є ефективнішим, ніж скалярний.

Розкрито адаптивне правило вибору метрики, тобто діагонального кроку, що ґрунтується на методі Барзілая-Борвейна (ББ). У поданому алгоритмі поєднуються два підходи: стандартний метод проксимального градієнта та проксимальний метод Ньютона. Реалізовано встановлення чітких правил вибору діагонального розміру кроку для алгоритмів опуклої оптимізації.

Ключові слова: опукла оптимізація; методи оптимізації; машинне навчання; діагональний розмір кроку; метод Барзілая-Борвейна; проксимальний градієнт.

Вступ

Методи оптимізації традиційно використовуються в інформаційних технологіях, зокрема телекомунікаційних мережах, що підтверджується позитивними результатами в широкому діапазоні різних даних. Застосування оптимізаційного підходу потребує попереднього кроку — моделювання та спрощення реальності. Такий підхід передбачає проведення значної роботи і в переважній більшості може призвести до прийняття неефективних рішень. Тому особливо актуальним є використання оптимізаційних методів засобами машинного навчання, що в майбутньому на основі аналізу та навчання на великих масивах даних забезпечить прогнозування різних позаштатних ситуацій у мережах.

Актуальним є використання методів машинного навчання в разі потреби отримання висновків із результатів різного моніторингу, а саме: якість трафіку чи сигналу. Крім того, використовуючи зазначені методи в застосунках у мережах, можна робити прогнози несправностей, контролю безпеки, вторгнень, маршрутизації тощо. Проте постає проблема з різнорідними даними. Наприклад, вихід із ладу компонентів гетерогенної мережі аналізують на основі сукупності різних параметрів мережі та чинників, що на неї впливають. А отже, для моделювання таких неоднорідних мереж використовують модель марковського випадкового поля.

Попарно-експоненціальне марковське випадкове поле належить до класу багатовимірних експоненціальних розподілів. Цей метод полягає в представленні спільного розподілу, що дасть змогу компактно подати різнорідні змінні, забезпечуючи зі свого боку отримання більш швидкого методу вивчення структури, для розподілу вузлів з невідомими параметрами. Тобто такий підхід дає можливість моделювати експоненціальний розподіл вузлів у гетерогенних мережах. На увагу заслуговує модифікація алгоритму проксимального градієнта — метод змінного метричного проксимального градієнта. Забезпечення швидкої збіжності досягається засобами діагональним розміром кроку, який є більш ефективним, ніж скалярний.

У статті проаналізовано адаптивне правило вибору метрики, тобто діагонального кроку, що ґрунтується на методі Барзілая-Борвейна (ББ). У розглядуваному алгоритмі поєднуються два підходи: стандартний метод проксимального градієнта та проксимальний метод Ньютона.

Постановка проблеми. Підвищення ефективності та функціональної стійкості системи керування гетерогенною мережею передбачає вдосконалення моделей та методів машинного навчання. Використання методів машинного навчання дає можливість уникнути дестабілізуючих чинників у мережах. Для моделювання таких мереж використовують графічні моделі, оцінювання яких є досить складним. Тому актуальною є удосконалена модель попарно-експоненціального марковського випадкового поля.

Мета і задачі дослідження. Метою дослідження є встановлення чітких правил вибору діагонального розміру кроку для алгоритмів опуклої оптимізації. Для реалізації мети запропоновано використовувати адаптивне правило вибору метрики, тобто діагонального кроку, що має назву крок Барзілая-Борвейна (ББ). У поданому алгоритмі поєднуються два підходи: стандартний метод проксимального градієнта та проксимальний метод Ньютона.

Аналіз останніх досліджень і публікацій. Проєктування, реалізацію та керування мережею доречно здійснювати з використанням методів оптимізації та машинного навчання. Застосування методів оптимізації в телекомунікаційних мережах забезпечує ефективні результати. Актуально підкріплення оптимізаційних методів засобами машинного навчання, що дає змогу аналізувати та навчатися на великих масивах даних, а також прогнозувати можливі позаштатні ситуації в мережі.

Галузь телекомунікацій уже підготовлена до впровадження машинного навчання. Оператори мереж працюють із великим обсягом даних: дані про клієнтів, дані про продуктивність в інтернеті, дані про інтернет-трафік та дані соціальних мереж тощо. Водночас для планування мережі та аналізу щодо пошуку закономірності в даних оператори використовують різні застосунки. Це зумовлює появу багатьох програм машинного навчання в телекомунікаційній сфері.

Машинне навчання сьогодні можна трактувати як відхід від шаблонів у проєктуванні майбутніх мереж і систем. Методи машинного навчання використовуються в застосунках у мережі для прогнозування несправностей, виявлення вторгнень, контролю безпеки, маршрутизації, реконфігурації пропускної здатності з огляду на трафік і багато іншого.

Основна частина

Опукла оптимізація як математична теорія вивчалася дуже давно і знайшла своє застосування під час синтезу керування, оброблення сигналів, у разі роботи з великими даними та машинним навчанням.

Для дослідження опуклої оптимізації за основу було взято таку модель:

$$\min_{x \in R^n} F(x) = f(x) + q(x), \tag{1}$$

де $x \in R^n$ — змінна розв'язку; функція $f : R^n \to R$ опукла та диференційовна, а функція $q : R^n \to R \cup \{\infty\}$ — опукла та може бути недиференційовна. Функція q може бути використана для кодування обмежень на змінну x.

Структурована модель (1) опуклої оптимізації застосовна в багатьох задачах машинного навчання: регресії, класифікації, побудові матриць тощо. Для розв'язування таких задач оптимізації дуже часто використовуються алгоритми проксимального градієнта, які уможливлюють покращення процесу обчислення, надаючи практичні правила для вибору кроку, теоретичні гарантії за помірних умов, підвищення точності результату та ін.

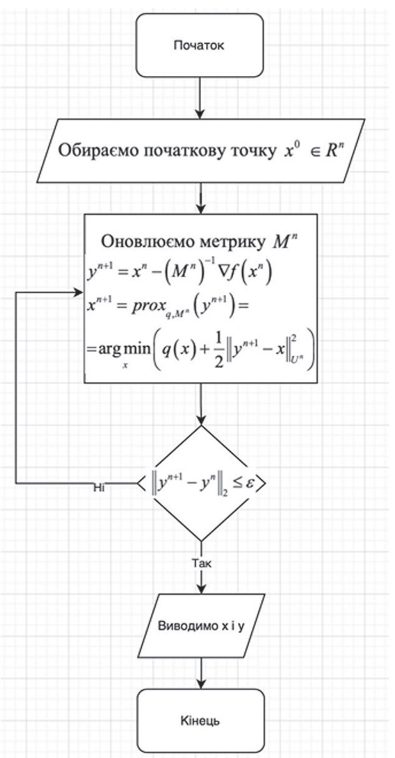

Рис. 1. Алгоритм роботи метричного проксимального градієнта

Більшість модифікацій алгоритму проксимального градієнта підпорядковуються одній формі, яка має назву метод змінного метричного проксимального градієнта (ЗМПГ).

Змінна x^k на k^{th}-й ітерації, де $M^k \in S^n_{++}$, має метрику $\|z\|_M = \sqrt{z^T M z}$, де M і $prox_{q,M}$ маштабоване проксимальне відбиття q відносно метрики M.

Алгоритм роботи метричного проксимального градієнта такий (рис. 1):

1. Вибрати початкову точку $x^0 \in R^n$.

2. Оновити метрику M^n:

$$y^{n+1} = x^n - (M^n)^{-1}\nabla f(x^n)$$
$$x^{n+1} = prox_{q,M}(y^{n+1})$$
$$\arg\min_x \left(q(x) + \frac{1}{2}\left\|y^{n+1} - x\right\|^2_{M^n}\right).$$

3. Якщо $\left\|y^{n+1} - y^n\right\|_2 \le \varepsilon$, то алгоритм зупиняється. В іншому разі необхідно повторити крок 2.

Від вибору метрики M^n буде залежати збіжність алгоритму. Вибір як метрика скаляру $M_n = (\beta^n)^{-1}$, де $\beta^n \in R$, дає негативні результати, а от вибір як метрика $M^n \approx \nabla^2 f(x^n)$ дозволяє використати метод Ньютона, який хоча і потребує значних витрат на покрокове обчислення, проте є швидкозбіжним. Що ж стосується методів проксимального градієнта, то вони мають низькі витрати на покрові обчислення, проте для забезпечення збіжності можуть вимагати значної кількості ітерацій.

Під час вибору кроку алгоритму, що забезпечує швидку збіжність, встановлено, що діагональний розмір кроку є більш ефективним, ніж скалярний. Отже, слід визначити чіткі правила вибору діагонального розміру кроку для алгоритмів опуклої оптимізації.

У статті запропоновано використовувати адаптивне правило вибору метрики, тобто діагонального кроку, що має назву крок Барзілая-Борвейна (ББ). У розглядуваному алгоритмі поєднуються два підходи: стандартний метод проксимального градієнта і проксимальний метод Ньютона. Метричний проксимальний градієнт разом із діагональним кроком ББ забезпечує низькі витрати на обчислення кроків, набагато краще гессіанське наближення на кожній ітерації, а отже, швидку збіжність алгоритму (порівняно з методом проксимального градієнта зі скалярним кроком). Проведені емпіричні дослідження довели, що запропонований метод з діагональною метрикою забезпечує кращу збіжність проти методу проксимального алгоритму зі скалярним кроком.

До найбільш популярних методів визначення діагонального кроку слід віднести: спектральний скалярний розмір кроку, змінну нескалярну метрику, діагональну метрику.

Зазвичай метод спектрального кроку застосовується для градієнтних методів типу ББ. Адитивне правило для вибору спектральної метрики використовує саме метод спектрального кроку. З метою зменшення визначених обмежень запропоновано нову адаптивну діагональну метричну стратегію вибору з гарантіями конвергенції за допомогою рядкового пошуку.

Крок проксимального градієнта можна розглядати як мінімізацію функції F, де диференційована частина f апроксимується у свою форму другого порядку для x^n, відносно $M^n \in C^h_{++}$:

$$prox_{q,M^n}\left(x^n-\left(M^n\right)^{-1}\nabla f\left(x^n\right)\right)=\arg\min_x q(x)+f\left(x^n\right)+\nabla f\left(x^n\right)^T\left(x-x^n\right)+\frac{1}{2}\left\|x-x^n\right\|^2_{M^n}.$$

Це наближення свідчить про те, що $M^n \approx \nabla^2 f(x^n)$ є оптимальним вибором після проксимального методу Ньютона. Проте використання гессіана зазвичай зумовлює високу вартість ітерацій. Альтернативним варіантом є апроксимація гессіана за допомогою січної стану:

$$M^n c^n \approx y^n, \tag{2}$$

для кроку $c^n = x^n - x^{n-1}$ і зміни градієнта $y^n = \nabla f(x^n) - \nabla f(x^{n-1})$.

Метод Барзілая–Борвейна — підхід, що оцінює скалярне наближення гессіана, встановлюючи $M^n = (\alpha^n)^{-1}$, яке задовольняє формулу (2). Метод є досить популярним.

Найбільш поширеними кроками в методі ББ є такі:

$$\begin{aligned}\beta^n_{ББ1} &:= \left\|c^n\right\|^2_2 \Big/ \left(c^n y^n\right);\\ \beta^n_{ББ2} &:= \left(c^n y^n\right) \Big/ \left\|y^n\right\|^2_2,\end{aligned} \tag{3}$$

де $\beta^n_{ББ2} \le \beta^n_{ББ1}$ завжди виконується через нерівність Коші–Шварца.

Диференційовна функція $f: R^n \to R$ є L-гладкою, якщо $\left\|\nabla f(x)-\nabla f(y)\right\|_2 \le L\left\|x-y\right\|_2$ виконується для всіх $x, y \in R^n$. І f є m-сильно опуклою, якщо $\left\langle \nabla f(x)-\nabla f(y), x-y\right\rangle \ge k\left\|x-y\right\|^2_2$ виконується для всіх $x, y \in R^n$.

Л е м а 1: Нехай диференційовна функція $f: R^n \to R$ є L-гладкою та m-сильно опуклою, тоді

$$\frac{1}{L} \le \beta^n_{ББ2} \le \beta^n_{ББ1} \le \frac{1}{k}.$$

Доведення зазначеної леми випливає з означення L-гладкості та m-сильно опуклості. Для багатьох вироджених сценаріїв із великими L або малими m оцінювання тривіальне і необхідним є гарантування чисельної стабільності оновлень у виразі (3). З цією метою до початкового етапу ББ прийнято кілька модифікацій та захисних заходів. Одним із таких числових заходів для c^n і y^n пропонується використання гібридного вибору між цими двома кроками:

$$\beta^n_{ББ} := \beta_{ББ}\left(c^n y^n\right) = \begin{cases}\beta^n_{ББ2}, \text{ якщо} \beta^n_{ББ1} < \partial\beta^n_{ББ2}\\ \beta^n_{ББ1} - \dfrac{1}{\delta}\beta^n_{ББ2}, \text{ в інших випадках,}\end{cases} \tag{4}$$

де гіперпараметр $\partial \in R$ зазвичай береться таким, що дорівнює 2. Якщо $\beta^n_{ББ}$ у виразі (4) є від'ємним, то вибирається попередній крок $\beta^n_{ББ} = \beta^{n-1}_{ББ}$.

Такі модифікації та заходи призначено для усунення нестабільності в початковому ББ з кроком (4) для погано зумовлених f. Проте навіть під час таких модифікацій скалярний ББ все ще схильний до невідповідностей. Слід зазначити, що $(\beta^n_{ББ1})^{-1}J$ і $(\beta^n_{ББ2})^{-1}J$ можна розглядати як наближення гессіана в евклідовому просторі. Водночас за погано зумовлених умов ці скалярні наближення можуть бути далекими від істинної неевклідової геометрії Гессе. В іншому ж разі, наприклад після апроксимальних наближень, зокрема проєкції напрямлення кроку c^n і зміни градієнта y^n, вони можуть бути близькими до ортогональності. Це зумовлює вироджені сценарії з $\beta_{ББ1} \to \infty$ або $\beta_{ББ2} \to 0$. Для таких випадків скалярні оцінки можуть значно відрізнятися від умови січної (2) і геометрії Гессе.

Далі в роботі розглядатиметься запропонований діагональний крок ББ.

Для відображення гессіанської геометрії f вводимо діагональну метрику M^n, яка на кожній ітерації n обчислюється в такий спосіб:

$$\min_{m\in R^n}\left\|Mc^n - y^n\right\|_2^2 + \eta\left\|M - M^{n-1}\right\|_F^2 \tag{5}$$

$$(\beta^n_{ББ1})^{-1}J \le M \le (\beta^n_{ББ2})^{-1}J$$

$$M = \mathrm{diag}(m),$$

де гіперпараметр $\eta > 0$ керує компромісом між задоволенням січної стану (2) та узгодженістю з попередньою метрикою M^{n-1}. Якщо гессіан змінюється швидко, то необхідно вибрати значення μ достатньо малим. У такому разі цей параметр відіграватиме роль числового захисту. Якщо ж гессіан під час ітерацій не сильно змінюється, то потрібно вибрати велике значення μ. Врешті решт діагональні елементи обмежені, тобто гарантовані кроком ББ (3).

Однією з переваг запропонованого формулювання (5) є те, що воно має просту замкнену форму розв'язання. Для $M = \mathrm{diag}(m_n)$, де $m^n = [m_1^n,..., m_h^n] \in R^n$ розв'язання задачі (5) задається так:

$$m_i^n = \begin{cases} \dfrac{1}{\beta^n_{ББ1}}, & \dfrac{c_i^n y_i^n + \eta m_i^{n-1}}{\left(c_i^n\right)^2 + \eta} < \dfrac{1}{\beta^n_{ББ1}} \\ \dfrac{1}{\beta^n_{ББ2}}, & \dfrac{c_i^n y_i^n + \eta m_i^{n-1}}{\left(c_i^k\right)^2 + \eta} > \dfrac{1}{\beta^n_{ББ2}} \\ \dfrac{c_i^n y_i^n + \eta m_i^{n-1}}{\left(c_i^n\right)^2 + \eta}, & \text{в інших випадках,} \end{cases} \tag{6}$$

де c_i^n та y_i^n — i-й елемент відповідно c^n і y^n.

Діагональ метричного вибору (6) швидше за все буде краще задовольняти умову січної (2) порівняно зі скалярним кроком ББ зі збереженням меншої вартості однієї ітерації порівняно з проксимальними методами типу Ньютона (таблиця). Наприклад, у вироджених випадках, коли $(c^n, y^n) \approx 0$ (за умов $\beta^n_{ББ1} \approx \propto$, $\beta^n_{ББ2} \approx 0$) нев'язка січної зі скаляром ББ може бути досить велика. Проте нев'язка $\left\| M^n c^n - y^n \right\|$ у разі діагональної метрики може бути набагато меншою при достатньо малих μ.

Вартість обчислення метрики M^n і кроку вперед $(x^n - (M^n)^{-1}\nabla f(x^n))$

	ПГ (ББ)	ЗМПГ (ДББ)	Проксимальний метод Ньютона
Метричний	$O(n)$	$O(n)$	$O(n^2)$
Кращий	$O(n)$	$O(n)$	$O(n^3)$

Причому m^n на кожній ітерації є скінченним доти, доки $0 \le m^{n-1} < \propto$ і $\eta > 0$. Це забезпечує стабільність алгоритму ЗМПГ із діагональним кроком ББ порівняно з гібридним скаляром ББ в (4).

Крім того, хоча обидва гібридні скаляри ББ (4) і діагональний ББ залежать від попередньої метрики, гібридний ББ використовує обмежену інформацію, де попереднє значення розміру кроку просто копіюється для негативного початкового кроку. Що ж до діагоналі ББ, то вона краще використовує цю додаткову інформацію через визначений користувачем параметр μ, додаючи взаємодію між кращими гессіанським наближенням та числовою стабільністю. Установлення великого значення η аналогічне копіюванню розміру попереднього кроку, що є характерним для гібридного скаляра ББ (рис. 2).

Рис. 2. Графічне подання методу ББ

Діагональна метрика M^n еквівалентна масштабуванню координат на кожній ітерації подальшим градієнтом та проксимальним кроком. Отже, алгоритм ЗМПГ із діагональною метрикою можна розглядати як виконання послідовності масштабування координат, де масштаб на кожній ітерації n залежить від локальної кривизни, тобто гессіанської. Це робить алгоритми діагональної метрики менш чутливими до величезних коливань у масштабі координат. Подібно до багатьох методів ББ алгоритм із діагональною метрикою в (6) все ще може не гарантувати збіжності без пошуку рядків для штрафних не квадратичних задач. Отже, слід використовувати (6) як початкову метрику та додатково виконувати пошуки рядків.

Висновки

Запропонована діагональна метрика забезпечує кращу оцінку погано зумовленого локального гессіана порівняно зі стандартним скалярним кроком Барзілая-Борвейна ББ, що призводить до швидшої збіжності алгоритму. У поєднанні з немонотонним лінійним пошуком загальний алгоритм є гарантовано збіжним. Нарешті, для кількох програм машинного навчання зі штучними та реальними наборами даних емпіричні результати демонструють кращу поведінку збіжності для запропонованої методології.

Упровадження запропонованої методики метричного проксимального градієнта в керування гетерогенною телекомунікаційною мережею забезпечить ефективне децентралізоване керування ресурсами гетерогенної мережі та зменшить кількість службової інформації в мережі. Це дасть змогу уникнути перевантаження мережі в разі виникнення позаштатних ситуацій. Проте все ще залишається відкритим питання навантаження на обладнання мережі в процесі керування нею за умов великої кількості користувачів.

Список використаної літератури

1. ***Protection*** *of telecommunication network from natural hazards of global warming / P. Anakhov, V. Zhebka, G. Grynkevych, A. Makarenko // Eastern-European Journal of Enterprise Technologies. 2020. 3(10 (105)). P. 26–37.*

2. ***Systematization*** *of measures on lightning protection of the objects of telecommunications network / P. Anakhov, A. Makarenko, V. Zhebka [et al.] // International Journal of Advanced Trends in Computer Science and Engineering. 2020. № 9(5). P. 7870–7877.*

3. ***Berkman L., Kriuchkova L., Zhebka V., Strelnikova S.*** *Universal Method of Multidimensional Signal Formation for Any Multiplicity of Modulation in 5G Mobile Network Lecture Notes in Electrical Engineeringthis link is disabled, 2022, 831. P. 305–321.*

4. ***Measuring*** *the effectiveness of a radio-identification system / O. S. Bichkov, V. S. Nakonechnyi, N. V. Lukova-Chuiko [et al.] // Journal of Communications. 2020. №15 (9). P. 669–675.*

5. ***Canale A., Lunardon N.*** *Churn prediction in telecommunications industry. A study based on bagging classifiers telecom // Carlo Alberto Notebooks, 2014. Vol. 350. P. 1–11. URL: https://www.carloalberto.org/assets/working-papers/no.350.pdf.*

6. ***Hughes J., Murali H., Petruta C Caragea.*** *Autologistic models for binary data on a lattice // Environmetrics 2011. 22.7. P. 857–871.*

7. ***Khan A. A., Sanjay J., Sepehri M. M.*** *Applying data mining to customer churn prediction in an Internet service provider // Int. J. Comput. Appl., 2010. Vol. 9. No. 7. P. 8–14. URL: http:// www.ijcaonline.org/volume9/number7/pxc3871889.pdf.*

8. ***The Diagnostics*** *Methods for Modern Communication Tools in the Armed Forces of Ukraine Based on Neural Network Approach / O. Klymovych, V. Hrabchak, O. Lavrut [et al.] // MOMLET 2020 (Modern Machine Learning Technologies Workshop). P. 198–208. [Електронний ресурс]. URL: http://ceur-ws.org/Vol-2631.*

9. ***Wainwright M. J., Jordan M. I.*** *Graphical Models, Exponential Families, and Variational Inference. Found. and Tr. in Mach. Learn., 1(1–2):1–305, 2008.*

10. ***Zucchini W., MacDonald I. L., Langrock R.*** *Hidden Markov Models for Time Series: An Introduction Using R. Chapman and Hall, 2016.*

11. ***Жебка В. В.*** *Моделювання марковського випадкового поля з метою його подальшої оптимізації та застосування // Зв'язок. 2020. №5. С. 35–40.*

12. ***Жебка В. В.*** *Управління фінансовими ризиками за допомогою машинного навчання // Зв'язок. 2018. №6. С. 32–35.*

13. ***Оптимизация*** *работы алгоритма градиентного бустинга с помощью перекрестной проверки / В. В. Жебка, В. И. Виноградов, А. П. Бондарчук, М. Н. Степанов // Актуальні проблеми економіки. Київ: 2019. №12 (222). С. 189–197.*

14. **Жебка В. В., Негоденко Е. В., Аронов А. А.** *Алгоритм максимально эффективного использования памяти для попарного марковского случайного поля // Актуальні проблеми економіки. 2020. №1 (223). С. 180–191.*

15. **Лаврут О. О., Лаврут Т. В.** *Модель та метод управління трафіком в мережах зв'язку критичного призначення. Prospects and priorities of research in science and technology: Collective monograph. Vol. 2. Riga, Latvia: Baltija Publishing, 2020. P. 36–60.*

В. В. Жебка, В. А. Корецкая

СОВЕРШЕНСТВОВАНИЕ ИНФОРМАЦИОННОЙ ТЕХНОЛОГИИ НА ОСНОВЕ МЕТОДА МЕТРИЧЕСКОГО ПРОКСИМАЛЬНОГО ГРАДИЕНТА С ВВЕДЕНИЕМ ДИАГОНАЛЬНОГО ШАГА

Рассмотрены отдельные аспекты использования методов оптимизации в телекоммуникационных сетях. Особенно актуально использование оптимизационных методов средствами машинного обучения во избежание различных внештатных ситуаций в сетях. Целесообразно использование методов машинного обучения при получении информации о качестве сигнала, трафика и т.п. Вместе с тем можно делать прогнозы разного рода неисправностей, маршрутизации, контроля безопасности.

Определено, что при моделировании неоднородных сетей эффективно использование модели марковского случайного поля. Данный подход позволяет моделировать экспоненциальное распределение узлов в гетерогенных сетях.

Приведена модификация алгоритма проксимального градиента — метод переменного метрического проксимального градиента. Обеспечение быстрой сходимости достигается средствами диагонального размера шага, который более эффективен, чем скалярный.

Раскрыто адаптивное правило выбора метрики, т. е. диагонального шага, основанного на методе Барзилла-Борвейна (ББ). В представленном алгоритме сочетаются два подхода: стандартный метод проксимального градиента и проксимальный метод Ньютона. Реализовано установление четких правил выбора диагонального размера шага для алгоритмов выпуклой оптимизации.

Ключевые слова: выпуклая оптимизация; методы оптимизации; машинное обучение; диагональный размер шага; метод Барзилла-Борвейна; проксимальный градиент.

V. V. Zhebka, V. O. Koretska

IMPROVEMENT OF INFORMATION TECHNOLOGY ON THE BASIS OF METHOD OF METRIC PROXIMAL GRADIENT WITH INTRODUCTION OF DIAGONAL STEP

The article presents some aspects of the use of optimization methods in telecommunications networks. The use of optimization methods by means of machine learning is especially important in order to avoid various emergency situations in networks. It is advisable to use machine learning methods to obtain information about signal quality, traffic, etc. At the same time it is possible to make forecasts of various malfunctions, routing, safety control.

It is determined that the model of Markov random field is effective in modeling inhomogeneous networks. This approach allows modeling the exponential distribution of nodes in heterogeneous networks.

A modification of the proximal gradient algorithm is presented — a method of variable metric proximal gradient. Ensuring fast convergence is achieved by means of diagonal step size, which is more efficient than scalar.

The article reveals an adaptive rule for choosing a metric, a diagonal step based on the Barzilai-Borvain (BB) method. The presented algorithm combines two approaches: the standard proximal gradient method and the proximal Newton method. The establishment of clear rules for choosing the diagonal step size for convex optimization algorithms has been implemented.

The proposed diagonal metric provides a better estimate of the ill-conditioned local Hessian compared to the standard Barzilai-Borwein BB scalar step, leading to faster convergence of the algorithm. Combined with non-monotonic linear search, the general algorithm is guaranteed to converge. Finally, for several machine learning programs with artificial and real datasets, empirical results demonstrate better convergence behavior for the proposed methodology.

Implementation of the proposed metric proximal gradient technique in the management of a heterogeneous telecommunication network will ensure effective decentralized management of heterogeneous network resources and reduce the amount of service information in the network. This will avoid overloading the network in the event of emergency situations. However, the question of the load on the network equipment during its management in the presence of a large number of users still remains open.

Keywords: convex optimization; optimization methods; machine learning; diagonal step size; Barzilai-Borvain method; proximal gradient.