ВИКОРИСТАННЯ ПОЗИТИВНО ВИЗНАЧЕНИХ МАТРИЦЬ В ЗАДАЧАХ МАТЕМАТИЧНОГО ПРОГРАМУВАННЯ

А.О. АНТОНЮК,
Університет державної фіскальної служби України, Ірпінь Київської області, Україна
tolik__@ukr.net
Н.Г. АНТОНЮК,
Національний університет «Києво-Могилянська академія», Київ, Україна
tolik__@ukr.net
Т.В. БЄЛИХ,
Інститут кібернетики НАН України, Київ, Україна
krainaz@ukr.net

Анотація. *Приводяться особливості задачі мінімізації матричних функцій як задачі негладкої оптимізації. Розглядаються приклади задач мінімізації функцій на множині позитивно визначених матриць. Пропонується процедура зведення задачі мінімізації на множині матриць до задачі мінімізації в просторі більшої розмірності.*

Ключові слова: *мінімізація матричних функцій, позитивно визначена матриця, негладка оптимізація, нев'язка, ортогональна матриця.*

В [1,2] розглядаються задачі мінімізації матричних функцій. Під задачею мінімізації матричної функції (або матричної оптимізації) розуміється така задача, яку зручно формулювати в термінах матричного обчислення. Багато важливих характеристики матриць (власні числа симетричних квадратних матриць, їх суми) представляють собою негладкі функції від елементів матриць. Тому матричні задачі оптимізації часто виявляються негладкими, що ускладнює їх розв'язки класичними методами. Одна з перших робіт, в якій вивчалися властивості субградієнтів негладких функцій, що виникають у зв'язку з обмеженнями позитивної визначеності, коли обрані параметри входять лише в діагональні елементи матриць, була робота Флетчера [3]. В роботі [4] описані властивості деяких матричних функцій і наведені алгоритми розв'язку матричних задач математичного програмування. Один з найбільш практично ефективних алгоритмів розв'язання такого роду задач – r-алгоритм [5].

Якщо розглянути задачу

$$\min\{f(x) \mid f_i(x) \le 0,\ i = 1,...,k,\ x \in M\},$$

де $x \in R^n$, f, f_i – опуклі неперервні функції, M – опукла множина, то неважко побачити, що до неї зводяться такі матричні екстремальні задачі на графах: знаходження оцінки максимально зваженої внутрішньо стійкої множини вершин графа $G(V,E)$ і побудова оцінки максимального розрізу графа. Наприклад, в першій задачі цільова функція $f(x)= \lambda_{max}(A(x))$, $\varphi(x) = -\lambda_{min}(B(x))$, где λ_{max}, λ_{min} – максимальні і мінімальні власні значення матриці, $A(x) = W + U(x)$, $B(x) = 2I + U(x)$, W – вагова матриця, $U(x)$ – матриця, в якій елементи, що відповідають E, змінні, інші – нульові (вершини графа з'єднані послідовно), $\varphi(x) = \max f_i(x)$. Тут для перевірки виконання умови позитивної визначеності матриці B використовується критерій $\lambda_{min} \ge 0$. Так як $\lambda_{min}(B)$ – вгнута функція від елементів матриці, то $\varphi(x)$ – опукла функція від параметра x, причому цільова функція задачі також буде опуклою. Тому для її розв'язку можна застосовувати вищеописаний алгоритм.

Вищевказані задачі на графах відносяться до задач мінімізації матричних параметричних функцій з обмеженнями на позитивну визначеність. Більшість таких задач, як було зазначено раніше, є негладкими. В роботі [6] побудований алгоритм розв'язку загальної матричної задачі оптимізації, що використовує штрафні функції. Якщо всі функції задачі неперервно диференціюються, то алгоритм індукує послідовність, що сходиться до стаціонарної точки вихідної задачі. Тобто знаходиться точка, в якій виконуються умови Каруша – Куна – Таккера [1].

Серед інших оптимізаційних матричних задач відзначимо також задачі мінімізації деякої функції $f(A)$, де матриця A порядку n, визначається видом обмежень на елементи цієї матриці – вони повинні бути такими, щоб матриця A була позитивно визначеною.

Але обмеження такого виду неможливо представити в традиційному для оптимізаційних задач вигляді типу, наприклад $g(x) \le 0$.

Розглянемо приклади математичних моделей, що використовують саме позитивно визначені матриці.

При розробці моделей масопереносу в процесах адсорбції багатокомпонентних сумішей речовин із розчинів найбільш розповсюдженими вважаються два підходи. В першому з них моделі представляються у вигляді системи звичайних диференціальних рівнянь, яка має вигляд [7]

$$\dot{x} = B(\Phi(x) - x),$$

де $x \in E^n, \Phi(x)$ – відома вектор-функція рівноважних концентрацій, x – вектор концентрацій речовин, B – матриця (nхn). Аналіз закономірностей кінетики адсорбції та розробка методів розрахунку технології розділення сумішей речовин вимагають знання чисельних величин елементів саме матриці коефіцієнтів B. Задача ідентифікації матриці B може бути зведена до оберненої задачі, тобто до задачі мінімізації функції нев'язки

$$F(B) = \sum_{i=1}^{N} \left\| x_{i\,\mathrm{exp}} - x(t_i, B) \right\|^2$$

де x_{iexp} – експериментально отримані значення концентрацій, $x(t_i,B)$ – значення концентрацій, які отримуються як розв'язок системи рівнянь в моменти часу t_i при певних значеннях матриці B.

Проте елементи матриці B не можна обирати довільно. Це пов'язано з тим, що реальний процес повинен бути стійким в околі точки x_*, для якої $\Phi(x_*) = x_*$. Тобто в процесі мінімізації нев'язки матрицю B слід вибирати таким чином, щоб матриця (похідна правої частини системи) $-B(\Phi'(x_*) - I)$ мала власні числа з позитивними дійсними частинами. Зрозуміло, що якщо вона буде позитивно визначеною, то така вимога буде автоматично виконуватися.

Нехай $\{A\}$ – множина позитивно визначених матриць. Тоді, вважаючи, що існує $(\Phi'(x_*) - I)^{-1}$, для будь-якої матриці $A \in \{A\}$ покладемо

$$B(A) = -A(\Phi'(x_*) - I)^{-1},$$

і бачимо, що матриця $-B(A)(\Phi'(x_*) - I) = A$ буде завжди позитивно визначеною за способом побудови. Отже, задача мінімізіції нев'язки тепер звелася до задачі мінімізації функції $F(B(A))$ на множині позитивно визначених матриць.

Інший підхід до моделювання процесів масопереносу пов'язаний із заданням густин дифузійних потоків компонентів суміші у вигляді узагальненого закону Фіка. В цьому випадку рівняння кінетики адсорбції будуть рівняннями з частковими похідними [8]. Тут надамо розв’язок такої системи рівнянь

$$\bar{x} = (I - \frac{6}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{m^2}\exp(-\frac{m^2\pi^2 t}{r_0^2}D))x_0 ,$$

де $\bar{x}$ – середня по об’єму сферичної частинки радіусу r_0 величина адсорбції, x_0 – початкова концентрація речовин. В подальшому формується функція нев’язки, аналогічная попередній задачі. Але, як бачимо, вона також містить матрицю D, яка також повинна бути позитивно визначеною як умова збіжності наведеного матричного ряду.

Далі розвиваються ідеї із [7,8] побудови такої процедури, яка дозволяє установити взаємнооднозначну залежність між позитивною визначеністю матриці A (тобто змінною цільової функції) і деяким довільним «довгим» вектором x розмірності n^2. «Довільність» вектора x означатиме, що в задачі оптимізації на змінну не буде обмежень.

Введемо деякі поняття. Квадратна дійсна матриця називається позитивно (негативно) визначеною [9], якщо $(Ax,x) > 0$ ($(Ax,x) < 0$) для будь-яких $x \neq 0$. Відомо також [9], що будь-яка дійсна матриця A може бути єдиним чином представлена у вигляді суми симетричної A_c та кососиметричної $A_к$ матриць, тобто $A = A_c + A_к$, причому $A_c = (A + A^*)/2$ і $A_к = (A - A^*)/2$. Тут A^* – транспонована до A матриця. Крім того, матриця A буде позитивно визначеною тоді і тільки тоді, коли її симетрична складова A_c також буде позитивно визначеною.

Далі, як відомо [9], будь-яку дійсну симетричну матрицю A_c завжди можна привести до діагонального вигляду Λ деяким ортогональним перетворенням U, тобто $U^*A_cU = \Lambda$. Таким чином, задаючи матриці Λ з позитивними елементами і змінюючи ортогональне перетворення U якімось чином, можна завжди отримати позитивно визначені симетричні матриці $A_c = U\Lambda U^*$.

Нехай деякий $x \in E^N$, де E^N – евклідів простір і $N = n(n\text{-}1)/2$. Якщо Λ – діагональна матриця, то позначимо через λ вектор з відповідними їй компонентами, причому $\lambda \in E^n$. В [7] запропоновано спосіб побудови ортогонального перетворення U, за допомогою якого кожному N-вимірному вектору $x \in E^N$ ставиться у відповідність матриця $U(x)$. Тоді $A_c(x,\lambda)= U(x)\Lambda U^*(x)$. Тобто, вдається побудувати матрицю A_c за допомогою довільних векторів x і λ. Далі, розміщуючи елементи N-вимірного вектора z на місцях під головною діагоналлю матриці $A_к$ і ці ж елементи зі знаком мінус над головною діагоналлю, отримаємо кососиметричні матриці $A_к(z)$ з нульовою діагоналлю. Тобто для матриці A отримано наступне представлення $A = U(x)\Lambda U^*(x) + A_к(z)$, причому на змінні (x,λ,z) не накладаються ніякі обмеження.

В результаті функція $f(A)$, яку необхідно мінімізувати, набуває вигляду

$$f(A) = f(U(x)\Lambda U^*(x) + A_к(z)),$$

а процес мінімізації буде проходити в просторі змінних (x,λ,z) загальної розмірності n^2.

Таким чином, запропоновано процедуру, яка дозволяє звести задачу мінімізації функції $f(A)$ на множині позитивно визначених матриць $\{A\}$ до задачі мінімізації більшої роозмірності.

Література

1. О матричных задачах оптимизации [Електронний ресурс] / Э.И. Ненахов // Теорія оптимальних рішень. – 2010. – № 9. – С. 79-85. – Режим доступу: http://nbuv.gov.ua/UJRN/Tor_2010_9_12.
2. Ненахов Э.И. Методы решения негладких выпуклых задач математического программирования и их приложения: Дис... д-ра физ.-мат. наук: 01.05.01 / НАН Украины; Институт кибернетики им. В.М.Глушкова. - К., 2000. - 318 л. - Бібліогр.: л. 299-318.
3. Fletcher R. Semidefinite matrix constrains in optimization // SIAM J. Control Optim. – 1985. – 23. – P. 493–513.
4. Шор Н.З. Задачи минимизации матричных функций и недифференцируемая оптимизация // Обозрение прикладной и промышленной математики. – 1995. – 2, вып. 1. – С. 113–138.

5. Шор Н.З. Методы минимизации недифференцируемых функций и их приложение. – Киев: Наук. думка, 1979. – 199 с.
6. Kanzow C., Nagel C., Kato H., Fukushima M. Successive linearization methods for nonlinear semidefinite programs // Computational Optim. And Appl. – 2005. – 31. – P. 251–273.
7. Михалевич В.С., Редковский Н.Н., Антонюк А.А. Некоторые методы минимизации на множестве неотрицательно определенных матриц // Кибернетика. - 1986. - № 6. – С. 84-97.
8. Антонюк А.А., Марутовский Р.М., Редковский Н.Н. Численное решение обратной задачи нестационарной массопроводности многокомпонентных смесей // Инженерно-физический журнал. – 1987. – Т. 53, №1. – С.113-117.
9. Гантмахер Ф.Р. Теория матриц. – М.: Наука, 1967. – 576 с.