Шаговая регрессия и рекуррентный метод наименьших квадратов:

постановки задач, использование в машинном обучении, алгоритм Гревилля и его обращение как возможный инструмент реализации

В.И. Ерохин

[1] Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург, Россия
vka@mil.ru

Семинар по оптимизации, машинному обучению
и искусственному интеллекту «O&ML»

9 марта 2023 г.

Шаговая регрессия, рекуррентный МНК - неформальная схема

$$\frac{1}{\tilde{m}-\tilde{n}}\left\| \underbrace{\left.\begin{array}{c} \overset{1}{\boxed{b}}_{m} \\ \updownarrow \\ \overset{1}{\boxed{\beta}}_{1} \end{array}\right._{\tilde{m}}}_{\tilde{b}} - \underbrace{\begin{array}{ccc} \overset{n}{\boxed{A}}_{m} & \leftrightarrow & \overset{1}{\boxed{c}}_{m} \\ \updownarrow & \ddots & \vdots \\ \overset{n}{\boxed{d}}_{1} & \cdots & . \end{array}}_{\tilde{A}} \cdot \underbrace{\left.\begin{array}{c} \overset{1}{\boxed{x}}_{n} \\ \updownarrow \\ \overset{1}{\boxed{\chi}}_{1} \end{array}\right._{\tilde{n}}}_{\tilde{x}} \right\|_2^2 \to \min_{x,\chi,\dots}$$

$$\Updownarrow \tilde{A}\tilde{x} \cong \tilde{b}$$

$$\frac{1}{\tilde{m}-\tilde{n}}\left\| \tilde{b} - \tilde{A}\,\underbrace{\color{red}{\tilde{A}^{+}}\color{black}{\tilde{b}}}_{\hat{x}} \right\|_2^2 \to \min_{\tilde{m},\tilde{n},A,c,d,b,\beta}$$

Псевдообратная матрица

Пусть $A \in \mathbb{R}^{m \times n}$ – некоторая матрица, $\operatorname{rank} A = r \leqslant \min\{m, n\}$. *Псевдообратной* к матрице A будем называть матрицу $A^+ \in \mathbb{R}^{n \times m}$, для которой выполняются

уравнения Пенроуза:

$$AA^+A = A, \tag{1}$$

$$A^+AA^+ = A^+, \tag{2}$$

$$(AA^+)^\top = AA^+, \tag{3}$$

$$(A^+A)^\top = A^+A. \tag{4}$$

Утверждение 1. Для *любой* матрицы A псевдообратная матрица A^+ *существует* и является *единственной*.

Алгоритм Гревилля

Необходимые обозначения

1. Рассмотрим следующие блочные представления:

$$A_k = [A_{k-1} \,\vdots\, a_k],$$

$$B_k = (A_k)^+ = \begin{bmatrix} B_{k-1} \\ --- \\ b_k \end{bmatrix},$$

где A_{k-1} и A_k – матрицы размеров $m \times (k-1)$ и $m \times k$, составленные соответственно из первых $k-1$ и k столбцов матрицы A, a_k – k-й столбец матрицы A, $B_k \in \mathbb{R}^{k \times m}$ – матрица, псевдообратная к матрице A_k, B_{k-1} – подматрица, составленная из первых $k-1$ строк матрицы B_k, b_k – последняя строка матрицы B_k.

2. Пусть $c_k \in \mathbb{R}^m$ и $d_k \in \mathbb{R}^k$ – вспомогательные векторы.

Алгоритм Гревилля

1.

$$A_1 = a_1, A_1^+ = B_1 = b_1 = \begin{cases} a_1^\top / a_1^\top a_1 \text{ -- если } a_1 \neq 0, \\ 0 \in \mathbb{R}^{1\times m} \text{ иначе.} \end{cases} \quad (5)$$

2. $k = 2, 3, \ldots, n$

$$d_k = A_{k-1}^+ a_k, \quad (6)$$

$$c_k = a_k - A_{k-1} d_k, \quad (7)$$

$$\text{Если } c_k \neq 0 \text{ то } b_k = c_k^\top / c_k^\top c_k, \quad (8)$$

$$\text{Иначе } b_k = \frac{d_k^\top A_{k-1}^+}{1 + d_k^\top d_k}, \quad (9)$$

$$B_{k-1} = A_{k-1}^+ - d_k b_k, \quad (10)$$

$$A_k^+ = B_k = \begin{bmatrix} B_{k-1} \\ --- \\ b_k \end{bmatrix}. \quad (11)$$

«Обращение» алгоритма Гревилля

Постановка задачи

Известны

$$A = [\tilde{A} \mid a], \; A^+ = \begin{bmatrix} B \\ -- \\ b \end{bmatrix}.$$

Найти $\tilde{A}^+$.

Теорема (Об «обращении» алгоритма Гревилля)

$$\tilde{A}^+ = \begin{cases} B\left(I - \frac{b^\top b}{bb^\top}\right) & \text{если } ba = 1, \\ B\left(I + \frac{ab}{1-ba}\right) & \text{в противном случае .} \end{cases}$$

Замечание

Если $ba = 1$, то $\operatorname{rank} \tilde{A} = \operatorname{rank} A - 1$, в противном случае $\operatorname{rank} \tilde{A} = \operatorname{rank} A$.

Обоснование обращения алгоритма Гревилля

Лемма 1 (О необходимых и достаточных условиях выполнения условия $ba = 1$)

$ba = 1 \Leftrightarrow c \neq 0$, где, в силу (6), (7),

$$c = (I - \tilde{A}\tilde{A}^{+})a. \tag{12}$$

Доказательство.
1. Пусть $c \neq 0$. Тогда $c^\top c = a^\top(I - \tilde{A}\tilde{A}^{+})^\top(I - \tilde{A}\tilde{A}^{+})a =$ (в силу (3)) $=$
$= a^\top(I - \tilde{A}\tilde{A}^{+})a$. В то же время, в силу (8),

$$b = \frac{a^\top(I - \tilde{A}\tilde{A}^{+})}{a^\top(I - \tilde{A}\tilde{A}^{+})a},$$

откуда и получаем $ba = 1$.
2. Пусть $c = 0$. Тогда, в силу (9),

$$b = \frac{d^\top \tilde{A}^{+}}{1 + d^\top d},$$

откуда в силу (6) получаем

$$ba = \frac{d^\top d}{1 + d^\top d} < 1.$$

Обоснование обращения алгоритма Гревилля

Лемма 2 (О выполнении условия $\tilde{A}^+ b^\top = 0$)

$c \neq 0 \Rightarrow \tilde{A}^+ b^\top = 0$.

Доказательство. Пусть $c \neq 0$. Тогда, в силу (8),

$$\tilde{A}^+ b^\top = \frac{\tilde{A}^+ c}{c^\top c}.$$

Но в силу (12) и (2), $\tilde{A}^+ c = \tilde{A}^+ (I - \tilde{A}\tilde{A}^+) a = 0 \Rightarrow \tilde{A}^+ b^\top$.

Обоснование обращения алгоритма Гревилля

Доказательство теоремы

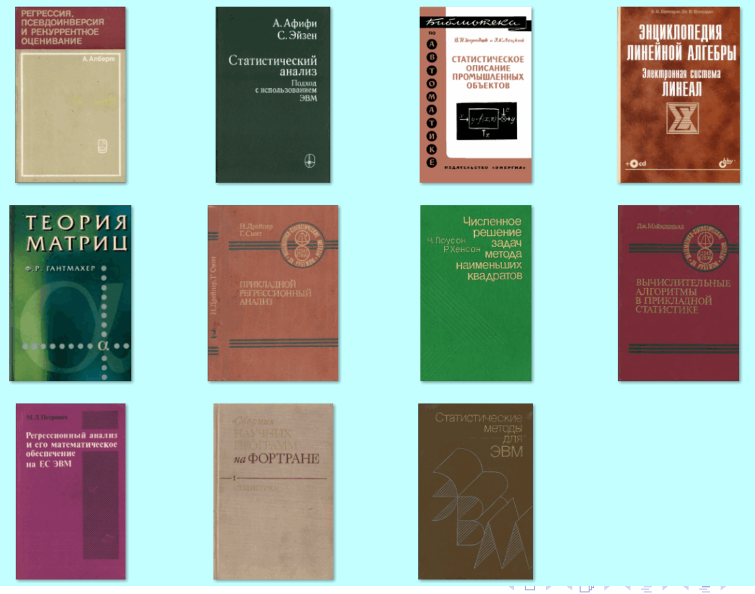
В силу (6),(10)

$$B = \tilde{A}^{+} - db = \tilde{A}^{+}(I - ab).$$

1. Пусть выполняется условие $ba \neq 1$. Тогда матрица $(I - ab)$ невырождена и, в силу тождества Шермана-Моррисона, $\tilde{A}^{+} = B(I - ab)^{-1} = B\left(I + \frac{ab}{1-ba}\right)$.
2. Пусть выполняется условие $ba = 1$. Умножив матрицу B справа на вектор $b^{\top}$, в силу лемм 1 и 2 получаем: $Bb^{\top} = \tilde{A}^{+}b^{\top} - dbb^{\top} = -dbb^{\top}$, откуда

$$d = -\frac{Bb^{\top}}{bb^{\top}}, \; \tilde{A}^{+} = B + db = B\left(I - \frac{b^{\top}b}{bb^{\top}}\right).$$

Литература (книги)



Литература (книги)

- *Гантмахер Ф.Р.* Теория матриц. – 5-е изд., – М.: ФИЗМАТЛИТ, 2010. – 560 с. (Псевдообратная матрица, с. 30 – 37, Алгоритм Гревилля, с. 37 – 38)
- *Бородюк В.П., Лецкий Э.К.* Статистическое описание промышленных объектов. – М.: Энергия, 1971. – 112 с. (Рекуррентный МНК, с. 61 – 86)
- Сборник научных программ на Фортране. Вып. 1. Статистика. – М.: Статистика, 1974. – 316 с. (Шаговая регрессия, с. 54 – 64)
- *Петрович М.Л.* Регрессионный анализ и его математическое обеспечение на ЕС ЭВМ: Практическое руководство. – М.: Финансы и статистика, 1982. – 199 с. (Рекуррентный МНК, с. 113 – 122)
- *Афифи А., Эйзен С.* Статистический анализ: Подход с использованием ЭВМ. – М.: Мир, 1982. – 488 с. (Шаговая регрессия, с. 194 – 208)
- *Алберт А.* Регрессия, псевдоинверсия и рекуррентное оценивание. – М.: Наука, 1977. – 224 с. Рекуррентный МНК, с. 151 – 160, Шаговая регрессия, с. 177 – 188, Фильтр Калмана, с. 203 – 207)

Литература (книги)

- *Лоусон Ч., Хенсон Р.* Численное решение задач метода наименьших квадратов. – М.: Наука, 1986. – 232 с. (..., Модификация QR-разложения матрицы при добавлении или удалении столбцов, с. 134 – 137, Модификация QR-разложения при добавлении или удалении строки, с. 160 – 169)
- Статистические методы для ЭВМ/Под ред. К. Энслейна, Э. Рэлстона, Г.С. Уилфа. – М.: Наука, 1986. – 646 с. (Шаговая регрессия, с. 77 – 122)
- *Дрейпер Н., Смит Г.* Прикладной регрессионный анализ: В 2-х кн. Кн. 2 - М.: Финансы и статистика, 1987. – 351 с. (Шаговая регрессия, с. 22 – 29)
- *Мэйндонал Дж.* Вычислительные алгоритмы в прикладной статистике. - М.: Финансы и статистика, 1988. - 350 с. (Шаговая регрессия, с. 51 – 64)
- *Воеводин В.В., Воеводин Вл. В.* Энциклопедия линейной алгебры. Электронная система ЛИНЕАЛ. – СПб.: БХВ-Петербург, 2006. – 544 с. (Псевдорешение и псевдообратная матрица, с. 209 – 214, Свойства псевдообратной матрицы, с. 214 – 224)

Литература (статьи)

- *Ерохин В.И., Кашмет В.В., Лисицын Н.В.* Модификация алгоритма Гревилля для удаления столбцов (строк) при построении псевдообратной матрицы // Журн. вычисл. матем. и матем. физ., 1989, Т.29, № 11, С. 1758.
- *Ерохин В.И., Лисицын Н.В., Кашмет В.В.* Модификация алгоритма Гревилля для удаления столбцов (строк) при построении псевдообратной матрицы // Рук. деп. в ВИНИТИ, 1989, № 4302–В89, 10 с.
- *Ерохин В.И.* Обобщение тождества Шермана-Моррисона на случай одноранговой модификации псевдообратной матрицы полного ранга// Журн. вычисл. матем. и матем. физ., 1999, Т.39, № 8, С. 1280–1282.
- *Икрамов Х.Д., М. Матин фар.* Пересчет нормальных псевдорешений при одноранговых модификациях матрицы // Журн. вычисл. матем. и матем. физ., 2003, Т. 43, № 4, С. 493–505.
- *Greville T.N.E.* Some applications of the pseudoinverse of marix // SIAM Rev., 1960, Vol. 2, No. 1, pp. 15–22.
- *Meyer C.D., JR.* Generalized inversion of modified matrices // SIAM J. Appl. Math., 1973, Vol. 24, No. 3, pp. 315–323.

Спасибо за внимание!
Вопросы?

Шаговая регрессия и рекуррентный метод наименьших квадратов:

постановки задач, использование в машинном обучении, алгоритм Гревилля и его обращение как возможный инструмент реализации

В.И. Ерохин

[1] Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург, Россия
vka@mil.ru

Семинар по оптимизации, машинному обучению
и искусственному интеллекту «O&ML»

9 марта 2023 г.