ДОКЛАДЫ АКАДЕМИИ НАУК, 2008, том 422, № 5, с. 601–604

МАТЕМАТИКА

УДК 519.853.3/.32

ФЕЙЕРОВСКИЕ ПРОЦЕССЫ С МАЛЫМИ ВОЗМУЩЕНИЯМИ

© 2008 г. Е. А. Нурминский

Представлено академиком С.К. Коровиным 12.11.2007 г.

Поступило 03.06.2008 г.

Для решения систем выпуклых неравенств широко применяется аппарат фейеровских процессов, сводку результатов по современному состоянию их теории и приложений можно найти в [1]. В общем виде эти процессы имеют вид итеративной схемы $x^{s+1} = F(x^s)$, $s = 0, 1, 2, \ldots$, где $F(x)$ – фейеровский оператор. Отличительной особенностью фейеровских операторов и процессов, построенных по этой схеме, является свойство притяжения и, в конце концов, сходимости к определенным множествам, в связи с чем они и используются как теоретические модели вычислительных алгоритмов. Особенно привлекательными эти процессы выглядят с точки зрения декомпозиции и использования параллельных вычислений, поэтому они широко применяются для решения задач большой размерности при обработке данных в компьютерной томографии, планировании радиационной терапии, распознавании образов, обработке изображений и других ситуациях, связанных с обработкой больших объемов данных. В данном сообщении рассматривается поведение фейеровских процессов с исчезающим возмущением, порожденным малым сдвигом в аргументе фейеровского оператора. Показано, что в случае локально сильной фейеровости $F(x)$ наличие исчезающей возмущающей добавки не препятствует сходимости к притягивающему множеству. Вместе с тем такой добавкой можно воспользоваться для придания процессу дополнительных свойств, которые обеспечивают сходимость к выделенным подмножествам притягивающего множества. Такая схема позволяет, в частности, предложить новый принцип декомпозиции экстремальных задач, не требующий наличия специфической структуры ограничений.

Институт автоматики и процессов управления Дальневосточного отделения Российской Академии наук, Владивосток

СХОДИМОСТЬ ФЕЙЕРОВСКИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ С ВОЗМУЩЕНИЯМИ

Основные события данной работы будут происходить в конечномерном евклидовом пространстве E со скалярным произведением xy и нормой $\|x\| = \sqrt{xx}$. Отличие от умножения вектора x на скалярный множитель α обычно ясно из контекста. Стандартный N-мерный симплекс будем обозначать $\Delta_N = \left\{ w_i \geq 0,\ i = 1, 2, \ldots, N;\ \sum_{i=1}^{N} w_i = 1 \right\}$. Окрестностью нуля будем считать произвольное открытое множество, содержащее $0 \in E$.

Согласно принятому определению, фейеровский оператор определяется относительно некоторого заданного множества V следующим образом.

О п р е д е л е н и е 1. Оператор $F: E \to E$ называется ф е й е р о в с к и м (относительно заданного множества V), если $F(v) = v$ для $v \in V$ и

$$\|F(x) - v\| \leq \|x - v\| \qquad (1)$$

для всех $v \in V$.

Множество V обычно подразумевается из контекста, и далее будем предполагать его замкнутость и ограниченность. Дополнительно к данному определению будет также предполагаться непрерывность $F(x)$ на некотором открытом расширении V.

С помощью фейеровского оператора F можно построить, начиная с какой-либо исходной точки x^0, итеративный фейеровский процесс $x^{s+1} = F(x^s)$, $s = 0, 1, \ldots$, который служит моделью вычислительного алгоритма для определения точки или точек множества V (задача допустимости). Свойство (1), а точнее его различные усиления, гарантируют сходимость в том или ином смысле элементов последовательности $\{x^s\}$ к множеству V. В обзорной работе [2] приводится далеко не полный список из более чем 100 работ, посвященных этому направлению. Как правило, для гарантий достаточно сильной сходимости фейеровского процесса требуются более сильные свойства, чем приведенные в определении 1, такие, например, как псевдофейеровость, квазисжимаемость и пр. Имея в виду дальнейшие приложения, мы предположим следующее свойство F, которое также усиливает (1).

Определение 2. Фейеровский оператор F будем считать локально сильно фейеровским в точке $\bar{x} \notin V$, если существует окрестность нуля U, такая, что

$$\|F(x) - v\| \leq \alpha \|x - v\| \tag{2}$$

для всех $v \in V$, $x \in \bar{x} + U$ и некоторого $\alpha < 1$.

Оператор $F(x)$ будем называть локально сильно фейеровским, если он локально сильно фейеровский в любой точке $x \notin V$. Коэффициент притяжения α зависит при этом, конечно, от выбранной точки. Справедлива следующая

Теорема 1. *Пусть V замкнуто и ограничено, F локально сильно фейеровский, последовательность $\{x^s\}$, полученная с помощью рекуррентных соотношений*

$$x^{s+1} = F(x^s + z^s), \quad s = 0, 1, 2, \ldots, \tag{3}$$

ограничена, $z^s \to 0$ при $s \to \infty$.

Тогда для произвольного x^0 все предельные точки $\{x^s\}$ принадлежат множеству V.

Полученные результаты можно обобщить на случай нестационарной фейеровской последовательности вида

$$x^{s+1} = F_s(x^s + z^s), \quad s = 0, 1, 2, \ldots, \tag{4}$$

где F_s выбирается из некоторого конечного семейства фейеровских операторов, что и позволяет использовать аппарат фейеровских процессов в декомпозиционных алгоритмах.

Теорема 2. *Пусть $\mathcal{F} = \{P_1, P_2, \ldots, P_m\}$ – семейство операторов P_i, таких, что для любой $x \notin V$ существует P_i, локально сильно фейеровский в x, $z^s \to 0$ при $s \to \infty$ и $F_s = P_{i_s}$, где P_{i_s} – локально сильно фейеровский в x^s оператор.*

Тогда если последовательность $\{x^s\}$, построенная по правилу (4), *ограничена, то все ее предельные точки принадлежат V.*

Приведенные теоремы показывают, что при достаточно щадящих условиях исчезающая помеха не препятствует основному результату теории фейеровских процессов – сходимости к некоторому множеству неподвижных точек. Далее приведены результаты об использовании малых дополнительных воздействий z^s с целью получения сходимости процессов (3), (4) к определенным подмножествам V.

Введем понятие локализованного аттрактанта как векторного поля, такого, что внутри V оно, содержательно говоря, направлено в сторону некоторого подмножества V.

Определение 3. Точечно-множественное отображение Φ: $V \to 2^E$ называется локализованным аттрактантом $Z \subset V$ в точке x, если из того, что $x \in V \setminus Z$, следует $g(z - x) \geq 0$ для всех $g \in \Phi(x)$ и $z \in Z$.

На самом деле для обоснования желаемой сходимости необходимо несколько более сильное свойство.

Определение 4. Аттрактант Φ называется сильным локализованным в точке x', если из того, что $x' \in V \setminus Z$, следует существование окрестности нуля U, такой, что

$$g(z - x) \geq \delta > 0$$

для всех $z \in Z$, $x \in x' + U$, $g \in \Phi(x)$ и некоторого $\delta > 0$.

Имея в виду некоторое фиксированное Z, будем называть Φ просто сильным локализованным аттрактантом, если приведенное свойство имеет место в каждой точке $V \setminus Z$. Константа δ будет, очевидно, зависеть от выбранной точки. Справедлива следующая

Теорема 3. *Пусть F – локально сильно фейеровский оператор, Φ – ограниченный и полунепрерывный сверху локализованный сильный аттрактант $Z \subset V$ и последовательность $\{x^s\}$ построена по правилу*

$$x^{s+1} = F(x^s + \lambda_s \Phi(x^s)), \tag{5}$$

где начальное состояние x_0 произвольно, $\lambda_s \to +0$, $\sum \lambda_s = \infty$.

Тогда если $\{x^s\}$ ограничена, то ее любая предельная точка принадлежит Z.

Так же, как и для теоремы 1, справедливо и ее обобщение на нестационарные фейеровские операторы.

Теорема 4. *Пусть F_s – локально сильно фейеровский в точке x^s оператор, выбираемый из конечного семейства $\mathcal{F} = \{P_1, P_2, \ldots, P_m\}$ непрерывных операторов P_i, таких, что $P_i(v) = v$, $i = 1, 2, \ldots, m$, для всех $v \in V$ и для любой $x \notin V$ существует P_i, локально сильно фейеровский в x, $z^s \to 0$ при $s \to \infty$, Φ – ограниченный полунепрерывный сверху сильный локализованный аттрактант $Z \subset V$ и последовательность $\{x^s\}$ построена по правилу*

$$x^{s+1} = F_s(x^s + \lambda_s \Phi(x^s)), \tag{6}$$

где начальное состояние x^0 произвольно, $\lambda_s \to +0$, $\sum \lambda_s = \infty$.

Тогда если $\{x^s\}$ ограничена, то ее любая предельная точка принадлежит Z.

Утверждения теорем 1–4 нетривиальны и не следуют из существующей теории фейеровских процессов, поскольку, вообще говоря, ни один из

(3), (4), (5), (6) даже просто фейеровским процессом не является. Для доказательства использовались общие условия сходимости итеративных процессов [3].

В приведенных теоремах использовалось на первый взгляд весьма ограничительное и трудно проверяемое глобальное условие ограниченности последовательности $\{x^s\}$. Однако алгоритмические схемы (теоремы 1–4) легко модифицировать с помощью какого-либо ретракта $R: E \to \tilde{V}$, возвращающего процесс $\{x^s\}$ в некоторое ограниченное множество $\tilde{V}$, такое, что $V + U \subset \tilde{V}$. Например,

$$x^{s+1} = \tilde{F}(x^s + z^s) =$$

$$= \begin{cases} F(x^s + z^s), & x^s \in \tilde{W} \supset \tilde{V} + U, \\ R(x^s) = y^s \in \tilde{V} \text{ в противном случае}, \end{cases} \quad (7)$$

где U – окрестность нуля, а y^s выбирается в $\tilde{V}$ произвольно.

МЕТОД ПРОЕКЦИЙ ГРАДИЕНТА С ДЕКОМПОЗИЦИЕЙ СИСТЕМЫ ОГРАНИЧЕНИЙ

Практическая польза от полученных результатов заключается, с одной стороны, в возможности использовать широкий спектр фейеровских операторов для решения проблемы допустимости, т.е. в гарантии того, что предельные точки построенной последовательности $\{x^s\}$ принадлежат допустимому множеству V даже при наличии возмущающего воздействия z^s, $s = 0, 1, 2, \ldots$ С другой стороны, полученный в работе результат позволяет использовать и различные способы улучшения допустимой точки, приближающие ее к множеству Z выделенных в V точек. В качестве таковых можно рассматривать, например, точки, являющиеся решением экстремальной или иной задачи на допустимом множестве V, и здесь также существует много способов определения, скажем, релаксационных направлений, приближающих текущую допустимую точку к множеству решений. Суть теорем 3, 4 заключается в том, что при соблюдении их условий можно достаточно свободно комбинировать схемы получения допустимых точек и алгоритмы достижения выделенного подмножества.

В связи с этим в качестве объекта приложений рассмотрим выпуклую задачу математического программирования

$$f_\star = \min_{x \in V} f(x) = f(x^\star), \quad x^\star \in Z \subset V, \quad (8)$$

с выпуклой конечной целевой функцией f и выпуклым допустимым множеством V. Одним из общих способов решения задачи (8) является комбинация градиентных шагов и проекции на множество V

$$x^{s+1} = \Pi(x^s - \lambda_s g^s), \quad g^s \in \partial f(x^s), \quad (9)$$

где Π – проекция на множество V, но последняя операция для множества общего вида весьма трудоемка и фактически редко применяется за исключением простейших V. Однако практически всегда допустимое множество V представлено как пересечение семейства выпуклых множеств V_i, $i = 1, 2, \ldots, N$: $V = \bigcap_{i=1}^{N} V_i$. Для решения по крайней мере задачи допустимости предложено множество фейеровских алгоритмов, использующих лишь отдельные операции проектирования на элементы V_i в представлении $V = \bigcap_{i=1}^{N} V_i$. В отличие от V элементы V_i могут быть достаточно простыми множествами, например полупространствами, линейными многообразиями, брусами, сферами и пр., операции проектирования на которые легко реализуемы с вычислительной точки зрения.

Для применения теоремы 2 к операторам проектирования достаточно показать их локально сильную фейеровость.

Теорема 5. *Пусть V – замкнутое ограниченное множество, представимое в виде пересечения конечного или бесконечного семейства выпуклых множеств: $V = \bigcap_{\tau \in T} V_\tau$, обозначим через $\Pi_\tau(x)$ проекцию точки x на V_τ.*

Тогда если $x \notin V_{\tau'}$ для некоторого $\tau' \in T$, то оператор $F = \Pi_{\tau'}$ является локально сильно фейеровским в точке x.

В силу теоремы 2 отсюда немедленно получаем, что при конечном множестве $T = \{1, 2, \ldots, N\}$ оператор F_s, построенный по принципу выбора в точке x^s оператора $F_s = \Pi_{i_s}$ с $x^s \notin V_{i_s}$ гарантирует, сходимость метода простого итерирования $x^{s+1} = F_s(x^s + z^s)$ решения задачи допустимости при весьма слабых условиях на возмущающее воздействие z^s. Заметим, что способ выбора множества V_{i_s} не имеет никакого значения, поэтому с позиций теории сходимости практически все построчные или сингулярные (row-action [4]) методы – циклической проекции, проектирования на самое удаленное множество, перемежающиеся (intermittent), максимальной невязки и пр., покрываются одним утверждением теоремы 2.

Однако теорема 4 дает возможность добиться и дополнительного эффекта за счет аттрактанта $\Phi(x^s) = -\partial f(x_s)$ – субдифференциального отображения целевой функции задачи (8). В силу неравенства $0 \le f(x) - f_\star \le g(x - z)$ для $g \in \partial f(x)$, $x \in V \cap Z$ отображение $\Phi(\cdot)$ является полунепрерывным сверху ограниченным локализованным сильным аттрактантом для Z.

Полагая $z^s = \lambda_s g^s$, $g^s \in \Phi(x^s) = -\partial f(x^s)$, получим, в силу теоремы 4, сходимость различных вариантов метода перемежающихся проекций градиента на элементы декомпозиции допустимого множества V:

$$x^{s+1} = \Pi_s(x^s - \lambda_s g^s), \quad s = 0, 1, 2, \ldots, \tag{10}$$

где Π_s – оператор проекции на множество V_{i_s}, такое, что $x^s \notin V_{i_s}$, $\lambda_s \to +0$, $\sum_{s=0}^{\infty} \lambda_s = \infty$. Такой подход позволяет произвести декомпозицию (10) метода проекций градиента (9), заменив сложно реализуемую проекцию на множество V проектированием на элементы V_i.

Используя теорему 5, можно показать и локально сильную фейеровость оператора

$$F(x) = \sum_{i=1}^{N} w_i \Pi_i(x), \tag{11}$$

где Π_i – проекция на множества V_i, $w = (w_1, w_2, \ldots, w_N) \in \Delta_N$.

Т е о р е м а 6. *Пусть для оператора F, заданного соотношением* (11), *имеет место* $\sum_{i:\, x \notin V_i} w_i \geq \gamma > 0$.

Тогда $F(x)$ в точке x является локально сильно фейеровским.

Полученный результат позволяет обосновать и использование параллельной версии декомпозиционного метода проекций градиента с фейеровским оператором вида (11)

$$x^{s+1} = \sum_{i=1}^{N} w_i^s \Pi_i(x^s - \lambda_s g^s),$$

где операции проектирования могут выполняться параллельно. Условия теоремы 6 на веса w_i^s выполняются, например, в том случае, когда они все равномерно отделены от нуля $w_i^s \geq \epsilon > 0$. Для шаговых множителей λ_s достаточно выполнения традиционных для негладких градиентных схем условий $\lambda_s \to +0$, $\sum_{s=0}^{\infty} \lambda_s = \infty$, хотя это, как показывает вычислительная практика, приводит к довольно медленной сходимости и подобные методы нуждаются в дальнейшем усовершенствовании.

Работа выполнена в рамках проекта ДВО РАН 06–III–А–01–459.

СПИСОК ЛИТЕРАТУРЫ

1. *Васин В.В., Еремин И.И.* Операторы и итерационные процессы фейеровского типа. Теория и приложения. Екатеринбург: УрО РАН, 2005. 210 с.
2. *Bauschke H.H., Borwein J.M.* // SIAM Revs. 1996. V. 38. T 3. P. 367–426.
3. *Нурминский Е.А.* Численные методы выпуклой оптимизации М.: Наука, 1991. 168 с.
4. *Censor Y.* // SIAM Revs. 1988. V. 23. P. 444–466.