УДК 519.8

ТОЧНЫЕ МЕТОДЫ МИНИМИЗАЦИИ КВАДРАТИЧНЫХ ФУНКЦИЙ. ЧАСТЬ 1

Ю.М. ДАНИЛИН, И.А. ШУБЕНКОВА

Рассматривается единый подход к построению различных по характеру точных методов минимизации квадратичных функций. Описываются возможности применения этих методов к решению задач нелинейной оптимизации и систем линейных уравнений.

Уже длительное время для построения быстросходящихся алгоритмов безусловной нелинейной оптимизации на каждой итерации решается квадратичная задача минимизации, в которой вместо матрицы вторых производных используется некоторая ее аппроксимация, строящаяся с применением только первых производных исходной функции или, даже, только ее значений [1–7]. Алгоритмы такого типа (методы сопряженных направлений, переменной метрики, двойственных направлений) используют различные аппроксимационные матрицы, и соответствующие методы изучаются независимо друг от друга, что затрудняет выявление каких-либо общих идей и свойств различных методов.

В данной работе предлагается достаточно общий подход к минимизации квадратичных функций, позволяющий строить и обосновывать различные по характеру алгоритмы минимизации (главным образом, будут рассматриваться методы, использующие те или иные процессы ортогонализации). В задачах нелинейной оптимизации результаты работы дают возможность строить различные аппроксимационные матрицы, которые используются при разработке соответствующих алгоритмов. В области точных методов решения линейных систем уравнений результаты дают возможность обосновывать свойства методов, базируясь на некоторых общих теоретических результатах, справедливых для различных процессов ортогонализации.

Под точными методами минимизации квадратичных функций подразумеваются процессы, дающие точное решение задачи за конечное число шагов (итераций).

ПРЕОБРАЗОВАНИЕ ЗАДАЧИ

Пусть

$$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c\,, \tag{1}$$

где $x \in E^n$ — n-мерное евклидово пространство $\langle Ax, x\rangle \ge 0$. Пусть ранг матрицы A равен $0 < t \le n$. Будем считать, что функция (1) достигает минимального значения, т.е. уравнение

$$f'(x) \equiv Ax + b = 0 \tag{2}$$

имеет решение. Обозначим подпространство вырождения матрицы A через E^{n-t} (подпространство, образованное такими линейно независимыми векторами r_i, что $Ar_i = 0$, $1 \le i \le n-t$). Ортогональное дополнение к E^{n-t} обозначим E^t. Тогда $f'(x) \in E^t$,

$$e = f'(x) - f'(x-r) = Ar \in E^t \text{, если } r \notin E^{n-t}. \tag{3}$$

Задачу минимизации функции (1) можно по-разному свести к решению системы линейных уравнений, определение которой не предполагает использования матрицы вторых производных A. (Отметим, что при минимизации конкретно функции (1) это чистая формальность, однако для нелинейных функций, где $f''(x)$ — матрица с переменными коэффициентами, использование лишь градиентов функции играет принципиальную роль.)

Рассмотрим некоторые из подходов.

Будем далее считать, что системы векторов r_i и $e_i = Ar_i$, $0 \le i \le t-1$ линейно независимы.

А. Определим систему уравнений

$$\langle Ax + b, r_i \rangle = 0,\ 0 \le i \le t-1 \text{, т.е. с учетом (3)}$$

$$\langle e_i, x \rangle = -\langle b, r_i \rangle,\ 0 \le i \le t-1. \tag{4}$$

Лемма 1. Любое решение уравнения (2) является решением системы (4) и наоборот, т.е. уравнение (2) и система (4) эквивалентны.

Доказательство. Первая часть утверждения очевидна: если x_* — произвольное решение уравнения (2), то оно удовлетворяет и системе (4).

Пусть теперь $\widetilde{x}$ — некоторое решение системы (4), т.е. $\langle e_i, \widetilde{x} \rangle = -\langle b, r_i \rangle$, $0 \le i \le t-1$. Если x_* — произвольное решение уравнения (2), то $\langle e_i, x_* \rangle = -\langle b, r_i \rangle$, $0 \le i \le t-1$, и поэтому $\langle e_i, \widetilde{x} - x_* \rangle = 0$, $0 \le i \le t-1$. Отсюда, с учетом линейной независимости векторов $e_i \in E^t$, $0 \le i \le t-1$ вытекает, что $\widetilde{x} - x_* = w \in E^{n-t}$, т.е. $Aw = 0$. Следовательно, $A\widetilde{x} + b = A(x_* + w) + b = Ax_* + b = 0$, т.е. $\widetilde{x}$ — решение уравнения (2).

Лемма доказана.

Б. Функцию $f(x)$ (1) представим в виде

$$f(x) = f(x_*) + \frac{1}{2}\langle A(x - x_*), x - x_* \rangle$$

или

$$f(x) = f(x_*) + \frac{1}{2}\langle f'(x), x - x_* \rangle. \tag{5}$$

Используя выражение (5), можно построить систему уравнений

$$f(x_i) = f(x) + \frac{1}{2}\langle f'(x_i), x_i - x\rangle, \ 0 \le i \le t \tag{6}$$

(x_i, $i = 0,...,t$ — произвольные точки).

Очевидно, любое решение уравнения (2) является и решением системы (6). С другой стороны, если система векторов $f'(x_{i+1}) - f'(x_i) = e_i = A(x_{i+1} - x_i)$, $0 \le i \le t-1$ линейно независима, то любое решение системы (6) определяет некоторую точку x_* (решение (2)) и значение $f(x_*)$. Действительно, пусть $(\widetilde{x}, f(\widetilde{x}))$ — произвольное решение системы (6). Вычитая уравнения (6) друг из друга, получаем

$$f(x_{i+1}) - f(x_i) = \frac{1}{2}\langle f'(x_{i+1}) - f'(x_i), x_{i+1} - \widetilde{x}\rangle + \frac{1}{2}\langle f'(x_i), x_{i+1} - x_i\rangle =$$

$$= \frac{1}{2}\langle e_i, x_{i+1} - \widetilde{x}\rangle + \frac{1}{2}\langle f'(x_i), x_{i+1} - x_i\rangle .$$

Поскольку в то же время

$$f(x_{i+1}) - f(x_i) = \langle f'(x_i), x_{i+1} - x_i\rangle + \frac{1}{2}\langle A(x_{i+1} - x_i), x_{i+1} - x_i\rangle =$$

$$= \langle f'(x_i), x_{i+1} - x_i\rangle + \frac{1}{2}\langle e_i, x_{i+1} - x_i\rangle, \text{ устанавливаем}$$

$$\frac{1}{2}\langle e_i, x_i - \widetilde{x}\rangle = \frac{1}{2}\langle f'(x_i), x_{i+1} - x_i\rangle . \tag{7}$$

Но $f'(x_i) = A(x_i - x_*)$, т.е. из (7) следует

$$\frac{1}{2}\langle e_i, x_i - \widetilde{x}\rangle = \frac{1}{2}\langle x_i - x_*, e_i\rangle, \ 0 \le i \le t-1 .$$

Отсюда так же, как в А, устанавливается

$\widetilde{x} = x_* + w$, $w \in E^{n-t}$, т.е. $\widetilde{x}$ — решение уравнения (2). Следовательно, $f(\widetilde{x}) = f(x_*)$.

Приведенные рассуждения показывают, что справедлива

Лемма 2. Если система векторов $f'(x_{i+1}) - f'(x_i) = e_i = A(x_{i+1} - x_i)$, $0 \le i \le t-1$, линейно независима, то уравнение (2) и система (6) эквивалентны в том смысле, что любое решение x_* уравнения (2) удовлетворяет системе (6), и наоборот.

Тот факт, что решение системы (6) позволяет определять значение $f(x_*)$, при решении задачи (1) не играет особой роли, однако при минимизации нелинейных функций с использованием вспомогательных квадратичних задач это может иметь существенное значение.

Отметим, что системы вида (6), помимо минимизации квадратичных функций, использовались также для минимизации функций более общего вида — однородных [8].

В. Пусть $e_i = Ar_i$, $i = 0,1,...,\lambda$, где $0 < \lambda \le n-1$ — линейно независимая система векторов, и

$$f'(x_0) \equiv A(x_0 - x_*) = \sum_0^{\lambda} \beta_i e_i \ . \quad (8)$$

Тогда

$$x_0 - x_* = \sum_0^{\lambda} \beta_i r_i \ . \quad (9)$$

КОНЕЧНОШАГОВЫЕ ПРОЦЕССЫ РЕШЕНИЯ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

Итак, для отыскания точки минимума функции $f(x)$ (1) достаточно решить одну из систем уравнений (4), (6). Далее будем рассматривать конечношаговые методы решения систем линейных уравнений. Разработка таких методов основывается на следующих соображениях (берем за основу систему (4)).

Пусть x_0 — произвольная точка, а точки $x_1,...,x_t$ строятся таким образом, чтобы выполнялись условия

$$\langle x_{k+1}, e_i \rangle = -\langle b, r_i \rangle, \quad 0 \le i \le k, \quad k = 0,1,...,t-1 . \quad (10)$$

Тогда, как вытекает из сравнения систем (10) и (4), точка x_t будет решением системы (4), т.е. $x_t = x_*$. Если выбирать точки $x_1,...,x_t$ по формуле

$$x_{k+1} = x_k + \alpha_k p_k, \quad k = 0,1,...,t-1, \quad (11)$$

где α_k — скалярный множитель, а p_k — вектор, то систему (10) можно записать в виде

$$\alpha_k \langle p_k, e_i \rangle = -\langle b, r_i \rangle - \langle x_k, e_i \rangle, \quad 0 \le i \le k . \quad (12)$$

При $k \ge 1$ точка x_k должна удовлетворять условиям (10): $\langle x_k, e_i \rangle = -\langle b, r_i \rangle$, $0 \le i \le k-1$.

Следовательно, из (12) следует

$$\alpha_k \langle p_k, e_i \rangle = 0, \quad 0 \le i \le k-1, \quad k \ge 1 . \quad (13)$$

При $i = k$ оказывается

$$-\langle b, r_k \rangle - \langle x_k, e_k \rangle = -\langle b, r_k \rangle - \langle Ax_k, r_k \rangle = -\langle Ax_k + b, r_k \rangle = -\langle f'(x_k), r_k \rangle ,$$

т.е.

$$\alpha_k \langle p_k, e_k \rangle = -\langle f'(x_k), r_k \rangle . \quad (14)$$

Условия (13), (14) будут, очевидно, выполнены, если вектор p_k удовлетворяет условиям

$$\langle p_k, e_i \rangle = 0, \quad 0 \le i \le k-1, \quad (15)$$

$$\langle p_k, e_k \rangle = c_k \neq 0 \,, \tag{16}$$

а множитель

$$\alpha_k = -\frac{\langle f'(x_k), r_k \rangle}{\langle p_k, e_k \rangle} \,. \tag{17}$$

Таким образом, точка x_{k+1}, определяемая в виде (11), будет удовлетворять условиям (10), если вектор p_k и множитель α_k определяется условиями (15), (16) и (17) (при $k = 0$ вектор p_k, очевидно, определяется лишь условием (16)).

Сформулируем некоторые утверждения, справедливые при условии, что точка x_{k+1} и векторы p_k определяются соотношениями (10), (15), (16), причем $p_k \neq 0$.

Утверждение 1. Векторы $p_0, \ldots, p_k$, удовлетворяющие условиям (15), (16), линейно независимы.

Утверждение 2. При любом $0 \le k \le t-1$ в точке x_{k+1}, удовлетворяющей условиям (10), реализуется минимум функции (1) на подпространстве, образованном векторами $r_0, \ldots, r_k$, т.е.

$$\langle f'(x_{k+1}), r_i \rangle = 0, \quad 0 \le i \le k \,. \tag{18}$$

Справедливость этого утверждения нетрудно установить, воспользовавшись условиями (13) и (14), которые можно записать в следующем виде:

$$\langle \alpha_k p_k, e_i \rangle = \langle A(x_{k+1} - x_k), r_i \rangle =$$

$$= \langle f'(x_{k+1}) - f'(x_k), r_i \rangle = 0, \quad 0 \le i \le k-1, \quad k \ge 1, \tag{19}$$

$$\langle \alpha_k p_k, e_k \rangle = \langle A(x_{k+1} - x_k), r_k \rangle =$$

$$= \langle f'(x_{k+1}) - f'(x_k), r_k \rangle = -\langle f'(x_k), r_k \rangle, \quad k \ge 0 \,. \tag{20}$$

Из (20) следует, что $\langle f'(x_{k+1}), r_k \rangle - 0$ при любом k. Учитывая это и используя (19), получаем условия (18).

Утверждение 3. Если

$$r_i = \sum_{j=0}^{i} \lambda_{ij} f'(x_j), \quad 0 \le i \le k \,, \tag{21}$$

где λ_{ij} — произвольные коэффициенты, причем $\lambda_{ii} \neq 0$, то при любом $0 \le k \le t-1$

$$\langle f'(x_{k+1}), f'(x_i) \rangle = 0, \quad 0 \le i \le k \,. \tag{22}$$

Справедливость этого утверждения можно установить, используя выражения (21) в (19), (20) и рассуждая по индукции. (Параллельно устанавливается линейная независимость векторов $f'(x_i)$, $0 \le i \le k$.)

Замечание. Если H_0 — некоторая невырожденная матрица, и

$$r_i = \sum_{j=0}^{i} \lambda_{ij} H_0 f'(x_j), \quad 0 \le i \le k, \tag{23}$$

то при любом $0 \le k \le t-1$

$$\langle f'(x_{k+1}), H_0 f'(x_i) \rangle = 0, \quad 0 \le i \le k. \tag{24}$$

Справедливость (24) при выборе r_i в виде (23) устанавливается точно так же, как (22). В частности, если $H_0 = A$, то условия (24) показывают, что градиенты $f'(x_0),\dots,f'(x_{k+1})$ оказываются A-ортогональными (или сопряженными), т.е. при выборе r_i в виде (23) (при условии $H_0 = A$) процесс (11) является методом сопряженных градиентов.

Отметим еще следующее. Система векторов $r_0,\dots,r_i$ выбирается достаточно произвольно (при условии, что она линейно независима, а также линейно независима система векторов $e_0,\dots,e_i$). Поэтому можно полагать $r_k = \beta_k p_k$, $k \ge 0$, $\beta_k \ne 0$, где вектор p_0 выбирается произвольно, а при $k \ge 1$ вектор p_k удовлетворяет условиям (15). При этом условия (15), (16) превращаются в условия A-ортогональности векторов p_i.

$$\beta_i \langle p_k, Ap_i \rangle = 0,$$

$$\beta_k \langle p_k, Ap_k \rangle = c_k \ne 0.$$

Коэффициент α_k выбирается по формуле

$$\alpha_k = -\frac{\langle f'(x_k), \beta_k p_k \rangle}{\langle p_k, \beta_k A p_k \rangle} = -\frac{\langle f'(x_k), p_k \rangle}{\langle p_k, Ap_k \rangle} = -\frac{\langle f'(x_k), p_k \rangle}{\langle p_k, f'(x_k + p_k) - f'(x_k) \rangle}.$$

Реализация процесса (11) таким способом ($r_k = \beta_k p_k$) заведомо возможна в случае, если матрица A невырожденная. Однако в случае, когда A — вырожденная матрица, процесс (11) также может вырождаться. Это происходит, если при некотором k оказывается $\langle p_k, Ap_k \rangle = 0$. Коэффициент α_k при этом определить невозможно.

Ситуация, когда процесс (11) вырождается, возможна только в случае, если выбор векторов r_k, $k = 0,1,\dots$ связан с реализацией процесса (11), т.е. зависит от вектора p_k. Если же линейно независимая система векторов r_k, $0 \le k \le t-1$, выбрана произвольно, то процесс (11) не вырождается. Ситуация, когда выполняются условия (15), (16), уже рассмотрена выше. Но может оказаться, что в некоторой точке x_k, $k \le t-1$, будет $\langle f'(x_k), r_k \rangle = 0$. (С учетом (18) это означает, что в точке $\boldsymbol{x_k}$ реализуется минимум функции (1) на подпространстве, образованном векторами $r_0,\dots,r_k$.) В этом случае можно полагать коэффициент $\alpha_k = 0$, при этом $x_{k+1} = x_k$. Вектор p_{k+1} определяется условиями

$$\alpha_{k+1}\langle p_{k+1}, e_i\rangle = 0, \quad 0 \le i \le k,$$

$$\alpha_{k+1}\langle p_{k+1}, e_{k+1}\rangle = -\langle f'(x_{k+1}), r_{k+1}\rangle.$$

Далее, в зависимости от значения $\langle f'(x_{k+1}), r_{k+1}\rangle = \langle f'(x_k), r_{k+1}\rangle$ (равна ли эта величина нулю или нет) процесс (11) продолжается так, как описано выше. По существу это означает, что процесс (11) может реализовывать точку минимума функции (1) за число шагов, меньшее t.

В частности, если определить систему векторов $r_0, \dots, r_{t-2}$ таким образом, чтобы выполнялись условия

$$\langle f'(x_0), r_i\rangle = 0, \quad 0 \le i \le t-2,$$

то можно полагать $\alpha_0 = \dots = \alpha_{t-2} = 0$, ($p_0 = \dots = p_{t-2} = 0$). В этом случае

$$x_* = x_t = x_0 + \alpha_{t-1} p_{t-1}.$$

Приведем еще одно свойство процессов (11).

Утверждение 4. Если

$$p_i = \sum_{j=0}^{i} \lambda_{ij} H_0 f'(x_j), \quad 0 \le i \le k, \quad r_i = \alpha_i p_i = x_{i+1} - x_i, \tag{25}$$

то

$$\langle f'(x_{k+1}), H_0 e_i\rangle = 0, \quad 0 \le i \le k-1, \quad k \ge 1. \tag{26}$$

Действительно, в данном случае справедливо (23), а в силу (25) $e_i = A(x_{i+1} - x_i) = f'(x_{i+1}) - f'(x_i)$. С учетом этого справедливость (26) непосредственно вытекает из (24).

Как уже отмечалось выше, в случае, когда матрица A вырожденная, соотношения (26) могут и не выполняться при некотором $k \le t-1$ (процесс (11) вырождается).

Наконец, отметим следующее.

Если определены векторы p_k, $0 \le k \le \lambda$, удовлетворяющие условиям (15), (16), то это позволяет вычислить коэффициенты β_i в уравнении (8).

$$\beta_\lambda = \frac{\langle f'(x_0), p_\lambda\rangle}{\langle p_\lambda, e_\lambda\rangle},$$

$$\beta_i = \frac{1}{\langle p_i, e_i\rangle}\left[\langle f'(x_0), p_i\rangle - \sum_{i<j\le\lambda} \beta_j \langle e_j, p_i\rangle\right], \quad i = \lambda-1, \lambda-2, \dots, 0.$$

Тем самым может быть определено решение по формуле (9). Определение решения в виде (9) также можно трактовать как конечношаговый процесс построения точек $x_{k+1} = x_k - \beta_k r_k$, $k = 0, 1, \dots, \lambda$.

В заключение отметим, что выполнения условий (15), (16) можно добиться, используя различные процессы ортогонализации: собственно ортогонализацию векторов e_i, A-ортогонализацию векторов p_i при специальном выборе векторов r_i, построение биортогональной системы векторов. Разумеется, можно использовать и другие процессы ортогонализации, известные в линейной алгебре (например, [9]). Во второй части данной работы будут подробно рассмотрены вопросы реализации различных способов построения вектора p_k.

ЛИТЕРАТУРА

1. *Пшеничный Б.Н., Данилин Ю.М.* Численные методы в экстремальных задачах. — М.: Наука, 1975. — 320 с.
2. *Полак Е.* Численные методы оптимизации. — М.: Мир, 1974. — 384 с.
3. *Гилл Ф., Мюррей У., Райт М.* Практическая оптимизация. — М.: Мир, 1985. — 510 с.
4. *Данилин Ю.М.* Методы сопряженных направлений для решения задач минимизации // Кибернетика. — 1971. — № 5. — С. 122–136.
5. *Данилин Ю.М.* Скорость сходимости методов сопряженных направлений // Кибернетика. — 1977. — № 6. — С. 97–105.
6. *Данилин Ю.М., Буланый А.П.* Квазиньютоновские методы минимизации, основанные на построении сопряженных систем векторов // Журн. вычисл. матем. и мат. физики. — 1978. — № 4. — С. 877–885.
7. *Данилин Ю.М.* Методы сопряженных направлений, не требующие решения одномерных задач минимизации // Докл. АН СССР. — 1974. — **218**, №3. — С. 513–516.
8. *Данилин Ю.М.* Об одном классе алгоритмов со сверхлинейной сходимостью // Журн. вычисл. матем. и мат. физики. — 1974. — № 3. — С. 598–609.
9. *Фаддеев Д.К., Фаддеева В.Н.* Вычислительные методы линейной алгебры. — М.: Наука, 1960. — 300 с.

Поступила 19.06.2007