УДК 519.8

ОПЕРАТОРЫ ПРЕОБРАЗОВАНИЯ ПРОСТРАНСТВА В КВАЗИНЬЮТОНОВСКИХ МЕТОДАХ И МЕТОДАХ СОПРЯЖЕННЫХ ГРАДИЕНТОВ

Журбенко Н. Г.

Институт кибернетики им. В.М. Глушкова НАН Украины

E-mail: zhurbnick@yandex.ua

Аннотация. Приводится интерпретация квазиньютоновского условия и условия сопряженности направлений при построении квазиньютоновских методов и методов сопряженных направлений на основе использования операторов преобразования пространства.

Ключевые слова. Квазиньютоновский метод, сопряженные направления, оператор преобразования, минимизация.

Введение

Эффективным и часто используемым на практике средством решения задач минимизации гладких функций являются алгоритмы, относящиеся к классу квазиньютоновских методов и методов сопряженных направлений. Имеются различные интерпретации и схемы построения этих методов [1 – 3]. Общая схема квазиньютоновских методов и методов сопряженных направлений безусловной минимизации функции $f(x)$ в R^n определяется следующим итеративным процессом генерации приближений $x_k \in R^n$:

$$x_k = x_{k-1} - h_k H_{k-1} g_{k-1}, k = 1,2,\ldots, \tag{1}$$

где h_k – шаговый множитель,

$$h_k = \arg\min_h f(x_{k-1} - hH_{k-1} g_{k-1}); \tag{2}$$

g_{k-1} – градиент функции $f(x)$ в точке x_{k-1}; H_{k-1} – симметричная положительно определенная матрица.

После шага k матрица H_k определяется в форме аддитивной поправки матрицы H_{k-1}: $H_k = H_{k-1} + \nabla H_k$, где ∇H_k – матрица ранга 1 или 2. Выбранная корректировка ∇H_k матрицы H_{k-1} определяет конкретный алгоритм (1).

Известны следующие «градиентные» интерпретации метода (1). Первая из них обычно приводится для квазиньютоновских методов и состоит в следующем. Шаг метода (1) соответствует шагу градиентного метода в пространстве с определяемой матрицей H_{k-1} метрикой, т.е. в пространстве со скалярным произведением $(\cdot,\cdot)_H$: $(x,z)_H = (H_{k-1}^{-1}x, z)$. Поэтому квазиньтоновские методы часто называют методами переменной метрики.

Другая интерпретация метода (1) связана с операцией преобразованием пространства переменных линейным оператором. Представим матрицу H_k в виде $H_k = B_k B_k^T$, где B_k – невырожденная матрица $n \times n.$. Такое представление корректно поскольку матрица H_k невырождена и положительно определена. Положим $A_k = B_k^{-1}$. Пусть Y – «преобразованное» соответствующим матрице A_k линейным оператором пространство: $Y = A_k X$. Градиент $\widetilde{g}(y)$ функции $\widetilde{f}(y) = f(B_k x)$, соответствующей функции $f(x)$ в пространстве Y, равен $B_k^T g(x)$. Поэтому шаг метода (1) соответствует (в исходном пространстве X) шагу градиентного алгоритма в «преобразованном» пространстве Y.

Построение методов рассматриваемого класса обычно основано на основе анализа задачи минимизации квадратичной функции

$$\min\{f(x) = (Cx, x) + (b, x) + c, \tag{3}$$

где $C -$ положительно определенная симметричная матрица (для методов сопряженных направлений можно рассматривать случай $C \geq 0$, однако, для краткости, мы будем считать, что $C > 0$).

Квазиньтоновские методы и методы сопряженных направлений обеспечивают (и должны обеспечивать по определению) решение задачи (3) не более чем за n шагов. В дальнейшем, для краткости, будем считать, что решение обеспечивается за n шагов.

Для квазиньютоновских методов (по определению) должно выполняться условие: $H_n = C^{-1}$. Отсюда следует так называемое квазиньютоновское условие рекуррентного пересчета матрицы H_k:

$$H_k(g_k - g_{k-1}) = -h_k H_{k-1} g_{k-1}. \tag{4}$$

Для методов сопряженных направлений выполнение условия $H_{n+1} = C^{-1}$ не обязательно (если это условие выполняется, то такой метод относится к классу квазиньютоновских). Так в численных схемах простейших алгоритмов сопряженных направлений матрица H_k вообще отсутствует (например, алгоритмы Флетчера-Ривса [4], Полака-Рибьера [4], Поляка Б.Т. [5]). Варианты таких алгоритмов и алгоритмов с несимметричными матрицами мы рассматривать не будем.

Основой построения методов сопряженных направлений является требование $C-$ортогональности (сопряженности) направлений $p_k = H_k g_k$:

$$(Cp_k, p_j) = 0, j = 0,1,\ldots k-1. \tag{5}$$

Наиболее известная методика построения методов сопряженных направлений разработана Х. Хуангом (H. Y. Huang) [1]. Согласно приведенной в работе [3] описанию этой методики, условие $C-$ортогональности (5) будет выполнено (с учетом (2)), если выполнены соотношения:

$$H_k C p_j = \lambda p_j, j = 0,1,\ldots k-1, \tag{6}$$

где λ – произвольная неотрицательная константа.

Также как и для квазиньютоновских методов, после шага k матрица H_k определяется в форме аддитивной поправки.

После описания приведенных выше хорошо и давно известных фактов, нашей задачей является интерпретация квазиньютоновского условия (4) и условия $C-$ортогональности (6) с позиции использования операторов преобразования пространства: матрица H_k представляется в виде $H_k = B_k B_k^T$.

Квадратичной функции $f(x)$ задачи (3) в преобразованном оператором $A_k = B_k^{-1}$ пространстве $Y = A_k X$ имеет вид $\widetilde{f}(y) = (\widetilde{C}_k y, y) + (B_k^T b, x) + c,$ где $\widetilde{C}_k = B_k^T C B_k$ – матрица квадратичной части функции в преобразованном пространстве. Положим: $p_k = H_{k-1} g_{k-1}, \quad \widetilde{p}_k = A_k p_k$ ($\widetilde{p}_k$ – направление в преобразованном пространстве, соответствующее направлению «движения» p_k алгоритма (1) в исходном пространстве).

Пусть матрица преобразования B_k удовлетворяет следующему условию (вектор $\widetilde{p}_k$ является собственным вектором матрицы квадратичной функции в преобразованном пространстве; собственное значение вектора равно λ_k)

$$\widetilde{C}_k \widetilde{p}_k = \lambda_k \widetilde{p}_k. \tag{7}$$

Утверждение 1. *Квазиньютоновское условие* (4) *и условие* (7)*, в котором* $\lambda_k = 1$, *эквивалентны.*

Естественная форма корректировки матрицы преобразования пространства A_k имеет вид

$$A_k = R_k A_{k-1} \ (B_k = B_{k-1} T_k, \text{ где } T_k = R_k^{-1}). \tag{8}$$

Такая форма соответствует последовательному преобразованию пространства: $Y_k = R_k Y_{k-1}$ ($Y_0 = X$).

Отметим, что представление известных вариантов квазиньютоновских методов (Давидона-Флетчера-Пауэлла, Бройдена-Флетчера-Шенно) в такой форме методов с преобразованием пространства приведено в работе [6].

Шаг (1) соответствует шагу обычного градиентного метода в пространстве Y_{k-1}:

$$y_k = y_{k-1} - h_k \widetilde{g}_{k-1},\ k = 1,2,\ldots, \tag{9}$$

где $\widetilde{g}_{k-1} = B_{k-1}^T g_{k-1}$ – градиент функции $\widetilde{f}(y) = f(B_{k-1} x)$ в точке $y_{k-1} = B_{k-1} x_{k-1}$.

Пусть $\widetilde{g}_k = B_{k-1}^T g_k$ – градиент функции $\widetilde{f}(y) = f(B_{k-1}x)$ в точке $y_k = B_{k-1}x_k$ пространства Y_{k-1}, $\{\widetilde{g}_{k-1}, \widetilde{g}_k\}$ – двумерное подпространство пространства Y_{k-1} определяемое векторами $\widetilde{g}_{k-1}$, $\widetilde{g}_k$ Выделим подкласс методов с преобразованием пространства следующим ограничением на выбор операторов R_k: R_k оператор изменяет только подпространство $\{\widetilde{g}_{k-1}, \widetilde{g}_k\}$). Для выделенного подкласса методов с преобразованием пространства справедливо следующее утверждение.

Утверждение 2. *Условия C-ортогональности* (6) *и условие* (7)*, в котором $\lambda_k = \lambda$, эквивалентны.*

Заключение

В заключение сделаем несколько замечаний.

1. Отметим содержательный смысл рассматриваемого класса квазиньютоновских методов и методов сопряженных направлений, генерируемых на основе преобразования пространства. В преобразованном пространстве Y_{k-1} выполняется шаг градиентного алгоритма (9). После этого производится преобразование пространства Y_{k-1} (изменяющее лишь подпространство $\{\widetilde{g}_{k-1}, \widetilde{g}_k\}$) с выполнением условия (7). Содержательный смысл этого условия: *образ* (*в пространстве* Y_k) *вектора смещения в пространстве* Y_{k-1} *должен являться собственным вектором матрицы квадратичной функции в пространстве* Y_k. Оказывается, что такой принцип выбора операторов преобразования соответствует квазиньютоновскому условию (4) и известному условию C-ортогональности (6) методов сопряженных направлений. Для квадратичной функции образ траектория метода (1) в пространстве Y_n соответствует движению по (взаимно ортогональным) собственным векторам матрицы $\widetilde{C}_n$.

2. Приведенная интерпретация квазиньютоновского условия и условия C-ортогональности не только интересна своей наглядностью (удивительно, но автор не обнаружил публикацию такой интерпретации). Она позволяет

строить новые модификации методов рассматриваемого класса. Так в работах [7, 8] разработано семейство алгоритмов на основе использования условия (7) и операторов растяжения пространства $R_\alpha(\eta)$ [9]. Как частный случай в это семейство входит предельный вариант r-алгоритма [10] (одного из наиболее эффективных алгоритмов негладкой оптимизации) с бесконечным коэффициентом растяжения.

Список литературы

[1] H. Y. Huang Unified approach to quadratically convergent algorithms for function minimization // J. Optim. Theory and Appl., 1970. – Vol. 5. – Pp. 405 – 423.

[2] Полак Э.Численные методы оптимизации. – М.: Мир, 1974. – 376с.

[3] Пшеничный Б.Н., Данилин Ю.М. Численные методы в экстремальных задачах. – М.: Наука, 1975. – 376 с.

[4] Fletcher R., Reeves C.M. Function minimization by conjugate gradients // Comput. J., 1964. –Vol. 7. – №2. – Pp. 149 – 154.

[5] Поляк Б.Т. Метод сопряженных градиентов в задачах на экстремум // ЖВМ и МФ, 1969. – Вып.9. – №4. – С. 807 – 821.

[6] Стецюк П.И. Линейные операторы в квазиньютоновских методах // Теория и приложения методов оптимизации. – К.: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 1998. – С. 3 – 8.

[7] Журбенко Н.Г. Построение семейства методов сопряженных направлений на основе использования оператора растяжения пространства // Теория и приложения методов оптимизации. – К.: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 1998. – С. 12 – 18.

[8] Журбенко Н.Г. Квазиньтоновские алгоритмы минимизации на основе использования оператора растяжения пространства // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 1999. – С. 45 – 50.

[9] Шор Н.З. Методы минимизации недифференцируемых функций и их приложения. – Киев: Наук. думка, 1979. – 199 с.

[10] Шор Н.З., Журбенко Н.Г. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов // Кибернетика. – Киев: Наук. Думка, 1971. – №3. – С. 51 – 59.