УДК 519.233.2 (075.8)

А.А. Светлаков, О.А. Пугачева

Синтез градиентных методов минимизации функций многих переменных с применением правых обратных матриц

Предлагается нетрадиционный подход к синтезу методов минимизации функций многих переменных. На примере решения простейшей задачи в ней синтезируется новый градиентный метод решения данной задачи, основанный на сведении ее на каждой итерации к решению недоопределенных линейных алгебраических уравнений и вычислению их решений с применением правых обратных матриц.

Ключевые слова: градиент, правая обратная матрица, матрица, минимизация функции.

doi: 10.21293/1818-0442-2016-19-4-117-121

Постановка простейшей задачи минимизации функции многих переменных

Пусть некоторая переменная величина Y является заданной и известной нам функцией $f(\cdot)$ независимых переменных $U_1, U_2, \dots, U_n$, где n – натуральное число такое, что $1 \leq n < \infty$. Представим данное предположение аналитически и соответственно запишем следующее равенство:

$$y = f(u_1, u_2, \dots, u_n). \quad (1)$$

Здесь y и $u_1, u_2, \dots, u_n$ – численные значения функции Y и независимых переменных $U_1, U_2, \dots, U_n$ соответственно. Всюду далее будем предполагать, что:

1) данная функция является непрерывной и хотя бы раз дифференцируемой по каждой из переменных $U_i, i = \overline{1, n}$;

2) она является унимодальной функцией и определена на всем векторном пространстве V^n .

Для упрощения и компактности последующих записей и выкладок всюду далее будем пользоваться вектором-строкой $\mathbf{u}$ и вектором-столбцом $\mathbf{u}^T$, полагая при этом, что они определяются следующими равенствами:

$$a) \mathbf{u} = (u_1, u_2, u_3, \dots, u_n), \quad (2a)$$

$$б) \mathbf{u}^T = (\mathbf{u})^T, \quad (2б)$$

где T – символ операции транспонирования векторов и матриц. Воспользовавшись данными векторами, представим соотношение (1) в следующем более компактном виде:

$$y = f(\mathbf{u}) = f(\mathbf{u}^T). \quad (3)$$

Для удобства и сокращения последующего изложения наряду с терминами «вектор $\mathbf{u}^T$ » и «вектор $\mathbf{u}$ » всюду далее будем использовать как их синонимы термины «точка $\mathbf{u}^T$ » и «точка $\mathbf{u}$ », отдавая при этом предпочтение тому из них, который наиболее соответствует сущности обсуждаемого вопроса.

Задача минимизации рассматриваемой нами функции $y = f(\mathbf{u}^T)$ заключается в том, чтобы найти вектор $\mathbf{u}_*^T$, удовлетворяющий равенству

$$\mathbf{u}_*^T = \arg \min \{f(\mathbf{u}^T) : \mathbf{u}^T \in V^n\}, \quad (4)$$

где символ $\arg \min \{\dots\}$ – (аргумент минимума) означает, что вектор $\mathbf{u}_*^T$ является таким вектором, при котором функция $y = f(\mathbf{u}^T)$ принимает минимальное значение y_* и при этом $\mathbf{u}_*^T \in V^n$.

Для решения данной задачи в настоящее время имеется значительное число методов и реализующих их алгоритмов, базирующихся на различных идеях и подходах [1, 2]. Однако, несмотря на это, остается актуальной проблема совершенствования уже имеющихся и создания новых методов и алгоритмов ее решения, обладающих теми или иными преимуществами по сравнению с уже имеющимися методами и алгоритмами.

Синтез градиентных методов минимизации функций многих переменных с применением правых обратных матриц

Известное в настоящее время семейство градиентных методов минимизации функций многих переменных получено эвристическими рассуждениями с использованием при этом сведений о градиенте минимизируемой функции и цели задачи ее минимизации [1]. Ниже наглядно иллюстрируется, что эти и подобные им методы можно не менее успешно синтезировать и с применением более формальных методов, если: 1) воспользоваться сведением задачи их синтеза к решению линейных алгебраических уравнений и 2) использовать для решения последних правые обратные матрицы [2].

Для получения необходимого нам уравнения воспользуемся линейной частью ряда Тейлора функции $f(\mathbf{u}^T)$:

$$y = f(\mathbf{u}^T) = f(\mathbf{u}_0^T) + \mathbf{g} \Delta \mathbf{u}^T + O_2(\Delta \mathbf{u}^T) \quad (5)$$

здесь $\mathbf{u}_0^T$ – какая-либо точка из V^n , в окрестности которой нас интересует поведение функции (3); $\Delta \mathbf{u}^T = \mathbf{u}^T - \mathbf{u}_0^T$ – приращение вектора $\mathbf{u}_0^T$, а $\Delta \mathbf{u} = (\Delta \mathbf{u}^T)^T$; $\mathbf{g}^T$ – градиент функции (3), определяемый равенством

$$\mathbf{g}^T = \left(\frac{\partial f}{\partial u_1}, \frac{\partial f}{\partial u_2}, \dots, \frac{\partial f}{\partial u_n} \right)^T, \quad (6)$$

где $\frac{\partial f}{\partial u_i}$ – частная производная функции (3) по переменной $u_i, i=\overline{1, n}$, вычисленная в точке $\mathbf{u}_0^T$; $O_2(\Delta \mathbf{u}^T)$ – сумма всех остальных членов ряда Тейлора, имеющих порядок по компонентам вектора $\Delta \mathbf{u}^T$ выше второго.

При этом будем считать, что приращения $\Delta \mathbf{u}^T$ таковы, что слагаемым $O_2(\Delta \mathbf{u}^T)$ можно без какого-либо ущерба для дальнейших рассмотрений пренебречь, и соответственно будем считать его равным нулю. В этом случае получаем более простое равенство

$$y = f(\mathbf{u}^T) = f(\mathbf{u}_0^T) + \mathbf{g} \Delta \mathbf{u}^T, \quad (7)$$

где $\mathbf{g} = (\mathbf{g}^T)^T$, а $\mathbf{g} \Delta \mathbf{u}^T = \sum_{i=1}^n g_i \Delta u_i$ – скалярное произведение векторов $\mathbf{g}^T$ и $\Delta \mathbf{u}^T$.

Воспользуемся теперь данным равенством и запишем его дискретный аналог применительно к k -й итерации, где $k=1, 2, 3, \dots$. В результате получим следующее, равенство:

$$y_{k+1} = y_k + \mathbf{g}_k \Delta \mathbf{u}_k^T, \quad (8)$$

где $y_{k+1} \equiv f(\mathbf{u}_{k+1}^T)$, $y_k \equiv f(\mathbf{u}_k^T)$, а знак $\equiv$ означает, что левая часть соотношения равна по определению его правой части.

Значения y_k и градиента $\mathbf{g}_k^T$ здесь являются фиксированными и известными, а значение y_{k+1} и приращение $\Delta \mathbf{u}_k^T$ – неизвестными переменными величинами, значения которых связаны равенством (8). Таким образом, данное равенство можно вполне обоснованно рассматривать и использовать как линейное алгебраическое уравнение относительно неизвестных значений y_{k+1} и $\Delta \mathbf{u}_k^T$. В результате получим следующее основополагающее для наших дальнейших целей уравнение:

$$y_{k+1} - \mathbf{g}_k \Delta \mathbf{u}_k^T = y_k. \quad (9)$$

Для удобства и упрощения рассмотрений и выкладок придадим данному уравнению традиционный в линейной алгебре векторно-матричный вид. С этой целью введем в рассмотрение $(1 \times (n+1))$ -мерную матрицу $\mathbf{A}_k$ и $(n+1)$ -мерный вектор-столбец $\mathbf{x}_k^T$, определив их равенствами

$$\text{а) } \mathbf{A}_k = (1; -\mathbf{g}_k) \quad (10a)$$

$$\text{б) } \mathbf{x}_k^T = (y_{k+1}; \Delta \mathbf{u}_k^T)^T, \quad (10б)$$

вертикальное троеточие $\vdots$ в которых отделяет первые компоненты «1» и « y_{k+1} » от векторов $\mathbf{g}_k$ и $\Delta \mathbf{u}_k$. Воспользовавшись данными равенствами, представим уравнение (9) в следующем виде:

$$\mathbf{A}_k \mathbf{x}_k^T = y_k. \quad (11)$$

Анализируя данное уравнение, нетрудно видеть, что:

- 1) оно совместимо при любой правой части y_k ;
- 2) если $\mathbf{g}_k = \mathbf{0}_n$, где $\mathbf{0}_n$ – нулевой n -мерный вектор-строка, то оно имеет единственное решение $\mathbf{x}_k^T = (y_k; \mathbf{0}_n)^T$;
- 3) если $\mathbf{g}_k \neq \mathbf{0}_n$, то оно имеет несчетное множество решений.

В наиболее общем виде множество его решений $\mathbf{x}_k^T$ можно представить следующим равенством:

$$\mathbf{x}_k^T = \mathbf{A}_k^{-\Pi} y_k. \quad (12)$$

Здесь $\mathbf{A}_k^{-\Pi}$ – правая обратная к $\mathbf{A}_k$ матрица, вычисляемая согласно равенству [2]

$$\mathbf{A}_k^{-\Pi} = \mathbf{z}_k^T (\mathbf{A}_k \mathbf{z}_k^T)^{-1}, \quad (13)$$

где $\mathbf{z}_k^T$ – некоторый заданный $(n+1)$ -мерный вектор-столбец такой, что произведение $\mathbf{A}_k \mathbf{z}_k^T \neq 0$. Умножив слева обе части данного равенства на матрицу $\mathbf{A}_k$, нетрудно видеть, что $\mathbf{A}_k \mathbf{A}_k^{-\Pi} = \mathbf{1}$, и, таким образом, можно заключить, что матрица $\mathbf{A}_k^{-\Pi}$, определяемая равенством (13), действительно является правой обратной к $\mathbf{A}_k$ матрицей.

Множество векторов $\mathbf{z}_k^T$, удовлетворяющих данному неравенству, несчетно и, таким образом, задавая различные векторы $\mathbf{z}_k^T$ и используя их в (13), можно получить несчетное множество различных правых обратных к $\mathbf{A}_k$ матриц. При этом, как непосредственно видно из данного равенства, каждому вектору $\mathbf{z}_k^T$ соответствует единственная матрица $\mathbf{A}_k^{-\Pi}$, а, как видно из (12), каждая такая матрица определяет единственное решение уравнения (11). Приведем четыре примера, наглядно иллюстрирующих рассматриваемые возможности решения уравнения (11) с применением правых обратных матриц.

Пример 1. Зададим вектор $\mathbf{z}_1^T$ в соответствии с равенством $\mathbf{z}_1^T = (1, 0; \mathbf{0}_n)^T$, где $\mathbf{0}_n$ – нулевой n -мерный вектор-строка. Выполнив все необходимые вычисления согласно равенствам (13) и (12), учитывая при этом (9), получим, что решение $\mathbf{x}_1^T$ уравнения (11) в этом случае определяется равенством

$$\mathbf{x}_1^T = (y_k; \mathbf{0}_n)^T \quad (14)$$

и оказывается $(n+1)$ -мерным вектором, у которого отлична от нуля только первая компонента x_1 , а все остальные n его компонент равны нулю.

Пример 2. Пусть вектор $\mathbf{z}_2^T$ удовлетворяет равенству $\mathbf{z}_2^T = (\mathbf{0}_n; 1)^T$. В этом случае, выполнив представленную выше последовательность операций, получаем, что решение $\mathbf{x}_2^T$ уравнения (11) удовлетворяет равенству

$$\mathbf{x}_2^T = (\mathbf{0}_n; g_{nk}^{-1})^T y_k \quad (15)$$

и является $(n+1)$ -мерным вектором с отличной от нуля $(n+1)$ -й компонентой и равными нулю его первыми n компонентами.

Здесь g_{nk}^{-1} – обращенная n -я компонента градиента $\mathbf{g}_k^T$.

Пример 3. Выберем в качестве вектора $\mathbf{z}_3^T$ вектор $\mathbf{z}_3^T = (1; \mathbf{g}_k)^T$. Подставив его в (13) и выполнив очевидные операции, получим следующее равенство:

$$\mathbf{A}_k^{-\Pi} = (1; \mathbf{g}_k)^T (1 + \mathbf{g}_k \mathbf{g}_k^T)^{-1}, \quad (16)$$

в котором слагаемое $\mathbf{g}_k \mathbf{g}_k^T$ определяется равенством

$$\mathbf{g}_k \mathbf{g}_k^T = \sum_{i=1}^n g_{ik}^2, \quad (17)$$

где g_{ik} – i -я компонента градиента $\mathbf{g}_k^T$, $i = \overline{1, n}$, и таким образом, по определению является ничем иным, как квадратом евклидовой нормы градиента $\mathbf{g}_k^T$, имеющим строго положительное значение при $\mathbf{g}_k \neq \mathbf{0}_n$ и равным нулю при $\mathbf{g}_k = \mathbf{0}_n$. Решение $\mathbf{x}_3^T$ уравнения (11) в данном случае вычисляется согласно следующему равенству:

$$\mathbf{x}_3^T = (1; \mathbf{g}_k)^T (1 + \mathbf{g}_k \mathbf{g}_k^T)^{-1} y_k. \quad (18)$$

Пример 4. Будем считать, что вектор $\mathbf{z}_4^T$ определяется равенством $\mathbf{z}_4^T = \mathbf{e}_{n+1}^T = (1, 0; 1, 0; \dots 1, 0)^T$, т.е. является $(n+1)$ -мерным вектор-столбцом, все компоненты которого равны 1, 0. Поставив его в (13), получаем, что правая обратная к $\mathbf{A}_k$ матрица $\mathbf{A}_k^{-\Pi}$ в данном случае определяется равенством

$$\mathbf{A}_k^{-\Pi} = \mathbf{e}_{n+1}^T (1 + \sum_{i=1}^n g_{ik})^{-1}. \quad (19)$$

Решение $\mathbf{x}_4^T$ уравнения (11), вычисленное с использованием данной матрицы, получается в соответствии со следующим равенством:

$$\mathbf{x}_4^T = \mathbf{e}_{n+1}^T (1 + \sum_{i=1}^n g_{ik})^{-1} y_k. \quad (20)$$

Непосредственной проверкой, т.е. поставив в (11) решения (14), (15), (18) и (20) и выполняя соответствующие арифметические операции, можно убедиться в том, что векторы $\mathbf{x}_1^T$, $\mathbf{x}_2^T$, $\mathbf{x}_3^T$ и $\mathbf{x}_4^T$ действительно являются решениями уравнения (11). Используя данные решения и равенство (10б), можно соответственно синтезировать следующие 4 итерационных метода минимизации функции $f(\mathbf{u}^T)$:

$$1. \text{ а) } \mathbf{u}_{k+1}^T = \mathbf{u}_k^T, \quad (21a)$$

$$\text{б) } \mathbf{y}_{k+1} = \mathbf{y}_k; \quad (21б)$$

$$2. \text{ а) } \mathbf{u}_{k+1}^T = \mathbf{u}_k^T - (\mathbf{0}_{n-1}; g_{nk}^{-1})^T y_k, \quad (22a)$$

$$\text{б) } \mathbf{y}_{k+1} = 0; \quad (22б)$$

$$3. \text{ а) } \mathbf{u}_{k+1}^T = \mathbf{u}_k^T - \mathbf{g}_k^T (1 + \mathbf{g}_k \mathbf{g}_k^T)^{-1} y_k, \quad (23a)$$

$$\text{б) } \mathbf{y}_{k+1} = \mathbf{y}_k (1 + \mathbf{g}_k \mathbf{g}_k^T)^{-1}; \quad (23б)$$

$$4. \text{ а) } \mathbf{u}_{k+1}^T = \mathbf{u}_k^T - \mathbf{e}_k^T (1 + \mathbf{e}_k \mathbf{g}_k^T)^{-1} y_k, \quad (24a)$$

$$\text{б) } \mathbf{y}_{k+1} = \mathbf{y}_k (1 + \mathbf{e}_k \mathbf{g}_k^T)^{-1}. \quad (24б)$$

Здесь $\mathbf{0}_{n-1}$ – нулевой $(n-1)$ -мерный вектор-строка, а $\mathbf{e}_k^T$ – n -мерный вектор-столбец, каждая из компонент которого равна 1, 0.

Приведем краткие комментарии, более полно раскрывающие наиболее существенные особенности синтезированных методов минимизации функции $f(\mathbf{u}^T)$, и ограничения, которые необходимо учитывать при выборе вектора $\mathbf{z}^T$, фигурирующего в определении (13) матрицы $\mathbf{A}_k^{-\Pi}$.

Во-первых, как вытекает из определения вектора $\mathbf{z}_1^T$ и равенств (14), (21a), (21б), использование вектора $\mathbf{z}_1^T$ предельно упрощает вычисление матрицы $\mathbf{A}_k^{-\Pi}$ и решения $\mathbf{x}_1^T$ уравнения (11). Однако получаемый при этом итерационный метод (21a), (21б) оказывается совершенно бесполезным, так как его применение никак не решает задачу минимизации функции $f(\mathbf{u}^T)$.

Во-вторых, метод (22a), (22б) также оказывается малоприменимым для минимизации функции $f(\mathbf{u}^T)$, так как его применение позволяет изменять только n -ю компоненту u_n вектора $\mathbf{u}^T$, что, очевидно, явно недостаточно для отыскания минимума функции $f(\mathbf{u}^T)$. Более того, поскольку $y_{k+1} = 0$, то на $(k+1)$ -й и последующих интеграциях данная компонента также не будет в этом случае изменяться и соответственно не будет в этом случае никакого поиска минимума функции $f(\mathbf{u}^T)$ по первым $(n+1)$ компонентам $u_1, u_2, \dots, u_{n-1}$ вектора $\mathbf{u}^T$.

Однако вполне очевидно, что если вместо вектора $\mathbf{z}_2^T$ использовать вектор $\mathbf{z}_j^T$, определяемый равенством вида

$$\mathbf{z}_j^T = (\mathbf{0}_{j-1}; 1; \mathbf{0}_{n+1-j})^T, \quad j = \overline{2, n+1} \quad (25)$$

то можно без каких-либо больших усилий синтезировать итерационный метод покоординатной минимизации функции $f(\mathbf{u}^T)$. Для этого необходимо и достаточно организовать цикл по j и на каждой из итераций вычислять правую обратную к $\mathbf{A}_k$ матрицу $\mathbf{A}_{kj}^{-\Pi}$, решение $\mathbf{x}_{2j}^T$ и вычисление вектора $\mathbf{u}_{k+1}^T$ и значения y_{k+1} в соответствии с равенствами

$$а) \mathbf{u}_{k+1,j}^T = \mathbf{u}_{k,j-1}^T + \mathbf{z}_j^T y_j, \quad (26a)$$

$$б) y_{j+1} = y_j. \quad (26б)$$

В-третьих, как видно из (23а), (23б), представленный данными равенствами итерационный метод является градиентным методом минимизации функции $f(\mathbf{u}^T)$. Направление приращения $\Delta \mathbf{u}_k^T$ в данном случае совпадает с направлением её антиградиента $(-\mathbf{g}_k^T)$ в точке $\mathbf{u}_k^T$, а его евклидова норма $\|\Delta \mathbf{u}_k^T\|$ зависит как от градиента $\mathbf{g}_k^T$ и квадрата его евклидовой нормы $\|\mathbf{g}_k^T\|^2$, так и от значения y_k функции $f(\mathbf{u}^T)$. В окрестности точки $\mathbf{u}^*$ – минимума функции $f(\mathbf{u}^T)$, где выполняется приближенное равенство $\|\mathbf{g}_k^T\|^2 \approx 0$, направление приращения $\Delta \mathbf{u}_k^T$ также совпадает с направлением антиградиента $(-\mathbf{g}_k^T)$ в точке $\mathbf{u}_k^T$, а его евклидова норма $\|\Delta \mathbf{u}_k^T\|$ зависит от нормы $\|\mathbf{g}_k^T\|$ градиента $\mathbf{g}_k^T$ и от значения y_k функции $f(\mathbf{u}^T)$ в точке $\mathbf{u}_k^T$, монотонно уменьшаясь и стремясь к нулю с увеличением k . Как непосредственно видно из (23б), значение y_{k+1} функции $f(\mathbf{u}^T)$ в точке $\mathbf{u}_{k+1}^T$ оказывается в $(1 + \mathbf{g}_k^T \mathbf{g}_k^T)$ раз меньше её значения в точке $\mathbf{u}_k^T$ и с увеличением k монотонно и с монотонно уменьшающейся скоростью стремится к её значению $f((\mathbf{u}^*)^T)$.

В-четвёртых, из определения вектора $\mathbf{e}^T$ и равенства (24а) видно, что все компоненты u_{ik} вектора $\mathbf{u}_k^T$ в данном случае изменяются на одну и ту же величину, пропорциональную значению y_k и множителю $(1 + \mathbf{e}_k^T \mathbf{g}_k^T)^{-1}$. При этом направление перемещения точки $\mathbf{u}_k^T$ и его величина существенно зависят от того, каков угол φ между вектором $\mathbf{e}^T$ и градиентом $\mathbf{g}_k$. Так, если данный угол равен $\pi/2$, то имеют место равенства $\Delta \mathbf{u}_k^T = \mathbf{e}_n^T y_k$ и $y_{k+1} = y_k$, таким образом, никакого уменьшения значения функции $f(\mathbf{u}^T)$ в данном случае не будет. Оно будет максимальным при $\mathbf{e}_n^T = -\mathbf{g}_k^T$, т.е. тогда, когда угол φ между векторами $\mathbf{e}_n^T$ и $\mathbf{g}_k^T$ равен $-\pi$. Во всех других случаях, когда имеют неравенства $-\pi < \varphi < \pi/2$, оно будет изменяться пропорционально $\cos \varphi$.

Приведённые выше примеры наглядно иллюстрируют широчайшие возможности синтеза итерационных методов минимизации функции $f(\mathbf{u}^T)$ с

применением правых обратных к $\mathbf{A}_k$ матриц. Они же показывают, что: 1) используя данные матрицы, можно получить как совершенно бесполезные, так и представляющие значительный интерес итерационные методы минимизации функции $f(\mathbf{u}^T)$; 2) наибольший интерес представляют методы, полученные с применением вектора $\mathbf{z}_3^T$ не только потому, что его применение позволяет получить градиентный метод минимизации функции $f(\mathbf{u}^T)$, но и синтезировать нечётное семейство градиентных методов подобного назначения, заданных с точностью до диагональной положительно определённой матрицы $\mathbf{D}$ порядка k . Для получения какого-либо конкретного метода из данного семейства необходимо и достаточно вместо вектора $\mathbf{z}_3^T$ использовать вектор $\mathbf{z}_k^T$, определяемый равенством

$$\mathbf{z}_k^T = \mathbf{D}_k \mathbf{g}_k^T, \quad (27)$$

где $\mathbf{D}_k$ – диагональная порядка n матрица.

Равенства (8)–(27) полностью и вполне однозначно представляют совокупность операций, которые необходимо выполнить на k -й интеграции $k=1, 2, 3, \dots$. Однако они никак не представляют начало и окончание процесса минимизации функции $f(\mathbf{u}^T)$, т.е. не представляют никаких действий и правил, связанных с заданием начальной точки $\mathbf{u}_0^T$ и указанием условий останова данного процесса. Вместе с тем вполне очевидно, что и то и другое не только оказывают существенное влияние на продолжительность (число интеграций, необходимых для получения решения) процесса минимизации функции $f(\mathbf{u}^T)$, но и являются необходимыми условиями для его реализации.

Восполняя данный пробел, отметим, во-первых, что в любом случае точку $\mathbf{u}_0^T$ необходимо выбирать только из области определения функции $f(\mathbf{u}^T)$. Во-вторых, если рассматривать только сходящиеся методы минимизации функции $f(\mathbf{u}^T)$, которые прежде всего и представляют наибольший интерес с точки зрения их пригодности и полезности для решения задачи, то нетрудно видеть, что выбор начальной точки $\mathbf{u}_0^T$, необходимой для запуска процесса, вообще говоря, не является актуальной задачей, т.к. какую бы начальную точку $\mathbf{u}_0^T$ из множества отмеченных выше точек не выбрать, вычисляемая последовательность точек $\mathbf{u}_1^T, \mathbf{u}_2^T, \mathbf{u}_3^T, \dots$ будет удовлетворять соотношению вида

$$\lim_{k \rightarrow \infty} \mathbf{u}_k^T = \mathbf{u}_*^T, \quad (28)$$

т.е. будет сходиться к искомой точке $\mathbf{u}_*^T$. В-третьих, в тех случаях, когда актуальной является не только

сходимость данной последовательности к точке $\mathbf{u}_*^T$, но и её скорость, выбор начальной точки $\mathbf{u}_0^T$ оказывается актуальной задачей. Совершенно ясно, что в любом из подобных случаев начальную точку $\mathbf{u}_0^T$ необходимо выбирать как можно ближе к искомой точке $\mathbf{u}_*^T$, используя при этом все имеющиеся и доступные сведения о данной точке и поведении функции $f(\mathbf{u}^T)$ в её окрестности. В самом деле, такой выбор точки $\mathbf{u}_0^T$ при прочих одинаковых условиях минимизирует число интеграций, необходимых для попадания в достаточно малую окрестность точки $\mathbf{u}_*^T$, и соответственно, сокращает суммарный объём вычислений и время их реализации.

В-четвёртых, не менее очевидно и то, что во многих случаях оказывается вполне оправданным выбирать начальную точку $\mathbf{u}_0^T$ в соответствии с равенством

$$\mathbf{u}_0^T = 0,5(\mathbf{u}_{\min}^T + \mathbf{u}_{\max}^T), \quad (29)$$

т.е. использовать в качестве точки $\mathbf{u}_0^T$ центральную точку области задания функции $f(\mathbf{u}^T)$.

И, наконец, завершая синтез предлагаемого градиентного метода минимизации функции $f(\mathbf{u}^T)$, отметим, что для остановки процесса поиска точки $\mathbf{u}_*^T$ можно использовать какое-либо одно из следующих двух неравенств:

$$\text{а) } \|\Delta \mathbf{u}_k^T\| \leq \Delta_u \quad (30a)$$

$$\text{б) } \|\mathbf{g}_k^T\| \leq \Delta_g \quad (30б)$$

и завершать поиск данной точки при его выполнении. Здесь $\|\bullet\|$ – какая-либо, например евклидова, норма n -мерных векторов; Δ_u и Δ_g – некоторые заданные достаточно малые положительные числа, значения которых выбираются с учётом желаемой точности определения точки $\mathbf{u}_*^T$, а также длительности процесса её отыскания.

Заключение

Представленные выше результаты, позволяют заключить, что использование правых обратных матриц открывает широкие возможности синтеза различных методов и алгоритмов минимизации функций многих переменных и создания на их основе вычислительного алгоритма, реализующего не один, а целое семейство методов подобного назначения.

Литература

1. Крылов В.И. Вычислительные методы высшей математики. – Минск: Высшая школа, 1972. – 584 с.
2. Светлаков А.А. Обобщенные обратные матрицы: некоторые вопросы теории и применения в задачах управления процессами. – Томск: Изд-во НТЛ, 2003. – 388 с.

Светлаков Анатолий Антонович

Д-р техн. наук, профессор каф. компьютерных систем в управлении и проектировании (КСУП) ТУСУРа
Тел.: +7 (382-2) 90-01-73
Эл. почта: svetlakov.38@mail.ru

Пугачева Оксана Анатольевна

Ст. преподаватель каф. математики ТУСУРа
Тел.: +7 (382-2) 70-15-98
Эл. почта: Pugacheva_OA@main.tusur.ru

Svetlakov A.A., Pugacheva O.A.

Synthesis of Gradient minimization methods of functions of several variables with the right inverse matrix

It offers an unconventional approach to the synthesis methods of minimizing functions of many variables. For example, the simplest solution of the problem it is synthesized a new gradient method for solving this problem, based on intelligence it on each iteration to solve underdetermined linear algebraic equations and computation of their solutions with the right inverse matrix.

Keywords: gradient, right inverse matrix, matrix, minimization of the function.