

Ингрид МАУЭР

ОБ ОДНОМ СПОСОБЕ ПОСТРОЕНИЯ ТЕСТОВЫХ ЗАДАЧ ДЛЯ АЛГОРИТМОВ МИНИМИЗАЦИИ

(Представил Н. Алумяэ)

Результаты практических и теоретических исследований показывают, что скорость локальной сходимости алгоритмов при решении задач безусловной оптимизации значительно зависит от числа обусловленности гессиана оптимизируемой функции в точке экстремума, причем трудности решения, как правило, увеличиваются с ростом этого числа. Поэтому понятие числа обусловленности широко используется при построении тестовых задач для многих алгоритмов безусловной оптимизации. Но для составления условных тестовых задач наряду с целевой функцией необходимо учитывать и ограничения задачи. Так, в [1] при составлении условных задач, аналогично числу обусловленности в безусловной оптимизации, предлагают использовать число обусловленности некоторого уменьшенного гессиана функции Лагранжа составляемой задачи в точке ее решения и рассматривать это число как меру трудности этой задачи.

В данной работе описан базирующийся на этой мере способ построения условных и безусловных тестовых задач минимизации. Составляемые этим способом задачи содержат параметр, с ростом которого к бесконечности стремится к бесконечности и отмеченное число обусловленности. Таким образом, данный способ характеризуется целенаправленностью, без которой, как аргументируется в [2], составляемые условные задачи имеют тенденцию быть по предложенной в [1] мере сравнительно хорошо обусловленными.

Построим задачу

$$\min\{f(x) \mid g_1(x) \leq 0, g_2(x) = 0\}, \quad (1)$$

где $x \in E^n$ и функция $f(x)$ и вектор-функции $g_1(x)$ (m_1 -мерная) и $g_2(x)$ (m_2 -мерная) дважды непрерывно дифференцируемые. Обозначим через $\bar{f}_x(\bar{x})$ и $\bar{f}_{xx}(\bar{x})$ соответственно градиент и гессиан функции $f(x)$ в точке \bar{x} .

Пусть для этой задачи известны решение x^* и точка $u^* = (u_1^*, u_2^*)$, $u_1 \in E^{m_1}$, $u_2 \in E^{m_2}$, $u \in E^{m_1+m_2}$, такая, что пара (x^*, u^*) удовлетворяет необходимым условиям Куна—Таккера, т. е. $g_1(x^*) \leq 0$, $g_2(x^*) = 0$, $u_1^* \geq 0$, $u_1^{*T} g_1(x^*) = 0$, $L_x(x^*, u^*) = 0$, где $L(x, u)$ — функция Лагранжа задачи (1).

Пусть в точке x^* опорные гиперплоскости всех активных ограничений задачи (1), за исключением тех из $g_1(x) \leq 0$, которым в u_1^* соответствуют нулевые компоненты, выражаются системой уравнений

$$A(x - x^*) = 0.$$

Строками матрицы A являются градиенты соответствующих компонент $g_1(x)$ и $g_2(x)$ в точке x^* .

Предположим, что строки матрицы A линейно независимые, причем их число $m < n$, а пара (x^*, u^*) удовлетворяет условию второго порядка: для любого вектора $z \neq 0$ ($z \in E^n$), $Az = 0$ имеет место

$$z^T L_{xx}(x^*, u^*) z > 0.$$

При выполнении этих условий x^* — изолированное локальное решение задачи (1) и u^* определена единственным образом (см. [1, 3]). Отметим, что такую задачу можно построить, например, изложенными в [4, 5] способами.

Составим теперь новую задачу с параметром, слагая к целевой функции построенной задачи (1) некоторую функцию,

$$\min_x \left\{ f(x) + \frac{1}{2} t \|v(x) - v(x^*)\|^2 \mid g_1(x) \leq 0, g_2(x) = 0 \right\}. \quad (2)$$

Здесь $v(x)$ — дважды непрерывно дифференцируемая s -мерная вектор-функция, t — неотрицательный скалярный параметр и норма — евклидова пространства.

При любом значении $t \leq 0$, как нетрудно убедиться, точка x^* является изолированным локальным решением задачи (2) и пара (x^*, u^*) удовлетворяет необходимым условиям Куна—Таккера этой задачи, причем u^* единственно определена.

Обозначим функцию Лагранжа задачи (2) через $\mathcal{L}(x, u; t)$.

Не ограничивая общности, можно считать, что первые m столбцов матрицы A линейно независимые. Пусть матрица $A = (B, C)$, где B — $(m \times m)$ -мерная и C — $(m \times (n - m))$ -мерная матрицы. Так как строки матрицы $Z^T = ((-B^{-1}C)^T, I)$, где I — $((n - m) \times (n - m))$ -мерная единичная матрица, линейно независимые и $AZ = 0$, то с помощью этих строк можно получить $n - m$ ортонормированных векторов, определяющих пространство $D = \{z \mid Az = 0\}$. Пусть эти векторы являются столбцами $(n \times (n - m))$ -мерной матрицы Q .

Запишем теперь матрицу $\mathcal{L}_{xx}(x^*, u^*; t)$, уменьшенную на пространство D ,

$$Q^T \mathcal{L}_{xx}(x^*, u^*; t) Q$$

и обозначим через $p(t)$ функцию, значение которой при произвольной величине $t \geq 0$ аргумента t — число обусловленности этой матрицы при $t = \bar{t}$.

По предложенному в [1], мерой трудности задачи (2) со значением \bar{t} параметра t можно рассматривать число $p(\bar{t})$ и, следовательно, тестовыми задачами подходят такие задачи (2) с фиксированным параметром, для которых это число сравнительно большое. Таким образом, для получения хороших тестовых задач предложим в задаче (2) определить вектор-функцию $v(x)$ такой, чтобы для этой задачи $p(t) \rightarrow \infty$ при $t \rightarrow \infty$. В последующем приведем необходимые и достаточные условия, при выполнении которых этот предельный переход имеет место.

Далее используем замену переменных $x = Qy + x^*$, $y \in E^{n-m}$, которая определяет однозначное соответствие между точками множества $\bar{D} = \{x \mid A(x - x^*) = 0\}$ и пространства E^{n-m} и, в частности, точке x^* соответствует точка $y^* = 0$. Но, если $D = \bar{D}$, т. е. $Ax^* = 0$, то можно использовать и замену $x = Qy$, в случае которой $x^* = Qy^*$.

Введем вспомогательную функцию

$$\Phi(y; t) = \mathcal{L}(Qy + x^*, u^*; t),$$

для которой

$$\Phi_{yy}(y; t) = Q^T \mathcal{L}_{xx}(Qy + x^*, u^*; t) Q,$$

и, в частности,

$$\Phi_{yy}(y^*; t) = Q^T \mathcal{L}_{xx}(x^*, u^*; t) Q.$$

Обозначим через V якобиан вектор-функции $v(x)$ в точке x^* и покажем, что при любом значении $t \geq 0$ матрица $\Phi_{yy}(y^*; t)$ положительно определена. Действительно, для любой точки $y \neq 0$ величина

$$\begin{aligned} y^T \Phi_{yy}(y^*; t) y &= y^T Q^T L_{xx}(x^*, u^*) Q y + t y^T Q^T V^T V Q y = \\ &= z^T L_{xx}(x^*, u^*) z + t z^T V^T V z \end{aligned}$$

положительная; это по условию второго порядка задачи (1), поскольку $z \neq 0$, $Az = 0$, и симметричности матрицы $V^T V$.

Теперь можно рассмотреть функцию

$$p(t) = \|\Phi_{yy}(y^*; t)\| \|(\Phi_{yy}(y^*; t))^{-1}\|.$$

В этой функции при $t \geq 0$ обратная матрица от $\Phi_{yy}(y^*; t)$ всегда существует, так как эта матрица при любом значении $t \geq 0$ положительно определена.

Введем обозначение

$$W = Q^T V^T V Q.$$

Так как абсолютная величина определителя не изменится при перестановке его строк (столбцов) и при сложении к одной строке (столбцу) умноженной на произвольное число другой строки (столбца), то этими преобразованиями наряду с матрицей $\Phi_{yy}(y^*; t) = \Phi_{yy}(y^*; 0) + tW$ можно найти и другую матрицу того же типа, в которой вместо $\Phi_{yy}(y^*; 0)$ имеется другая независимая от t матрица, а вместо W — диагональная матрица (если, разумеется, матрица W сама уже не является диагональной), и такую, что их определители совпадают до знака (см. [6]). Число отличных от нуля элементов в полученной вместо W диагональной матрице является рангом матрицы W . Следовательно, определитель матрицы $\Phi_{yy}(y^*; t)$ — многочлен от t , степень которого равна рангу матрицы W .

Теорема. Для того, чтобы

$$p(t) \rightarrow \infty$$

при $t \rightarrow \infty$, необходимо и достаточно, чтобы матрица V удовлетворяла условиям

а) $V \neq 0$,

б) W невырожденная.

Доказательство. Необходимость. Для того, чтобы $p(t) \rightarrow \infty$ при $t \rightarrow \infty$, необходимо, чтобы $W \neq 0$, но это невозможно, если $V = 0$. Также необходимо, чтобы W была невырожденной. В самом деле, если матрица W вырожденная, то при $t \rightarrow \infty$ функция $p(t)$

$$\left\| \frac{1}{t} \Phi_{yy}(y^*; t) \right\| \left\| \left(\frac{1}{t} \Phi_{yy}(y^*; t) \right)^{-1} \right\| \rightarrow \|W\| \|W^{-1}\|,$$

которое конечная величина.

Достаточность. Пусть выполняются условия а) и б) теоремы. Из-за $V \neq 0$ можно найти точку \bar{z} такую, что $V\bar{z} \neq 0$. Определяя еще точку \bar{y} , для которой $\bar{z} = Q\bar{y}$, получим, что $\bar{z}^T V^T V \bar{z} = \bar{y}^T W \bar{y} > 0$, откуда вытекает, что $W \neq 0$.

Знаменатель, т. е. определитель матрицы $\Phi_{yy}(y^*; t)$, и числитель, т. е. соответствующее рассматриваемому элементу алгебраическое дополнение той же матрицы всех элементов матрицы $(\Phi_{yy}(y^*; t))^{-1}$,

являются многочленами от t . Определитель матрицы $\Phi_{yy}(y^*; t)$, т. е. знаменатель всех элементов $(\Phi_{yy}(y^*; t))^{-1}$, как мы уже показали, — многочлен, степень которого равна рангу матрицы W . Но по величине ту же степень имеет и числитель хотя одного элемента матрицы $(\Phi_{yy}(y^*; t))^{-1}$; в этом можно аналогично убедиться, учитывая, что матрица W вырожденная. По такому анализу элементов матрицы $(\Phi_{yy}(y^*; t))^{-1}$ приходим к выводу, что предельная матрица здесь при $t \rightarrow \infty$ не является нулевой и, следовательно, $\|(\Phi_{yy}(y^*; t))^{-1}\|$ стремится к отличному от нуля пределу при $t \rightarrow \infty$. В то же время $\|1/t\Phi_{yy}(y^*; t)\| \rightarrow \|W\|$, которая также отличается от нуля, так как $W \neq 0$. Следовательно, функция $p(t)$

$$t \left\| \frac{1}{t} \Phi_{yy}(y^*; t) \right\| \|(\Phi_{yy}(y^*; t))^{-1}\| \rightarrow \infty$$

при $t \rightarrow \infty$. Теорема доказана.

Так как ранг произведения матриц не превосходит рангов множителей этого произведения, то для выполнения условия б) теоремы достаточно выбрать числа m , n и s составляемой задачи такими, чтобы имело место неравенство $n - m > s$.

Однако не все задачи (2), вектор-функция $v(x)$ которых удовлетворяет условиям теоремы, с одинаковым успехом применимы в практике. Поэтому обратим внимание на функцию $\Phi(y; t)$. Так как $\Phi_y(y^*; t) = Q^T L_x(x^*, u^*) = 0$, поскольку $L_x(x^*, u^*) = 0$, и матрица $\Phi_{yy}(y^*; t)$ при любом значении $t \geq 0$ положительно определена, то для всех $t \geq 0$ точка y^* — изолированный локальный минимум этой функции. Опираясь на это обстоятельство, можно предложить в функции $\Phi(y; t)$ при фиксированном значении $t > 0$ выбрать вектор-функцию $v(Qy + x^*)$ такой, чтобы условия теоремы были выполнены и одновременно функция $\Phi(y; t)$ имела в окрестности точки y^* такую овражность, которая с ростом t стала сразу быстро увеличиваться, а затем эту вектор-функцию перенести в виде $v(x)$ в задачу (2).

Проиллюстрируем этот способ примерами безусловных задач (2), для которых $x^1 = (x_1, x_2)^T \in E^2$, $x^* = 0$, $Q = Z = I (m = 0)$, $D = \bar{D} = E^2$, $x = y$. Рассмотрим задачи

$$a) \min_x \frac{1}{2} [x_1^2 + x_2^2 + t(x_1 + x_2)^2], \quad t \geq 0,$$

где $V = (1, 1)$, $V^T V$ вырожденная, $p(t) = 2t + 1$;

$$б) \min_x \frac{1}{2} [x_1^2 + x_2^2 + 100(x_1 + x_2)^2 + t(x_1 - x_2)^2], \quad t \geq 0,$$

где $V = (1, 1)$, $V^T V$ вырожденная, $p(t) = 2t + 1$;

$$p(t) = \begin{cases} \frac{201}{2t+1}, & t \in [0, 100] \\ \frac{2t+1}{201}, & t \in (100, \infty). \end{cases}$$

В примере а) с ростом t овражность минимизируемой функции увеличивается, но в примере б) овражность сначала уменьшается, а начиная со значения $t = 100$ увеличивается.

Автор благодарит О. Ваарманна за ценные замечания.

ЛИТЕРАТУРА

1. Hock, W., Schittkowski, K. Lect. Notes Econ. and Math. Syst., № 187 (1981).
2. Stewart, G. H. Math. Programming, **21**, № 2, 235—238 (1981).
3. Фиакко А., Мак-Кормик Г. Нелинейное программирование. Методы последовательной безусловной минимизации. М., «Мир» 1972.
4. Schittkowski, K. Lect. Notes Econ. and Math. Syst., № 183, (1980).
5. Rosen, J. B., Suzuki, S. Commun. ACM, **8**, № 2, 113, (1965).
6. Kangro, G. Kõrgem algebra. Tallinn, ERK. 1962.

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
17/XI 1982

Ingrid MAUER

ÜHEST MINIMEERIMISALGORITMIDELE TESTÜLESANNETE KOOSTAMISE VIISIST

Artiklis on koostatud parameetrit sisaldav ülesanne (2) ning näidatud, et teatud tingimustel parameetri kasvades lõpmatusse suureneb lõpmatusse ka selle ülesande raskusmõõduna vaadeldav arv.

Ingrid MAUER

ABOUT A WAY OF CONSTRUCTING TEST PROBLEMS FOR ALGORITHMS OF MINIMIZATION

In this paper a problem, including the parameter, has been constructed (problem (2)). It is shown that under certain conditions, if the parameter is increasing to infinity, the condition figure considered as a measure of difficulty for the given problem, increases to infinity, too.