8. Влияние слабых магнитных полей на динамику изменения микротвердости кремния, индуцируемых малоинтенсивным бета-облучением / Ю.И. Головин [и др.] // ФТТ. – 2007. – Т.49, вып.5. – С. 822–823.
9. Песчанская, Н.Н. Скачкообразная ползучесть при сжатии монокристаллов цинка в магнитном поле / Н.Н. Песчанская, Б.И. Смирнов, В.В. Шпейзман // ФТТ. – 2008. – Т. 50, вып 6. – С. 997–1001.
10. Пинчук, А.И. Двойникование в кристаллах висмута при одновременном воздействии постоянного магнитного поля и сосредоточенной нагрузки / А.И. Пинчук, С.Д. Шаврей // Металлофизика. Новейшие технологии. – 2000. – Т.22, №12. – С. 43–46.
11. Пинчук, А.И. Магнитопластический эффект в случае двойникования кристаллов висмута под воздействием сосредоточенной нагрузки / А.И. Пинчук, С.Д. Шаврей // ФТТ. – 2001. – Т.43, вып.1. – С. 39–41.
12. Shavrey, S.D. A Decrease in the Mobility and Multiplication of Twinning Dislocations in Bismuth Crystals Exposed to Constant Magnetic Field / S.D. Shavrey, A.I. Pinchook // Technical Physics Letters. – 2003. – V. 29, №8. – P. 632–633.
13. Шаврей, С.Д. Влияние постоянного магнитного поля и сосредоточенной нагрузки на двойникование в кристаллах висмута / С.Д. Шаврей, А.И. Пинчук // Машиностроение. Республиканский межведомственный сборник научных трудов. Вып.18 / УП «Технопринт»; под ред. И.П. Филонова. – Минск, 2002. – С. 521–524.
14. Пинчук, А.И. Пороговый характер магнитопластического эффекта при двойниковании в кристаллах висмута / А.И. Пинчук, С.Д. Шаврей // ФТТ – 2004. – Т.46, вып.9. – С. 1603–1604.
15. Влияние постоянного магнитного поля и механической нагрузки на скопления полных и частичных дислокаций в кристаллах висмута: Отчет о НИР / Мозырский гос. пед. ун-т; рук. С.Д. Шаврей. – Мозырь, 2003. – 22 с. – № ГР 20031846.
16. Пинчук, А.И. Объемная упругая энергия двойников кристаллов висмута и поверхностная энергия границы раздела двойник-матрица в магнитном поле / А.И. Пинчук, С.Д. Шаврей // ФТТ – 2005. – Т.47, вып.11. – С. 1964–1966.
17. Исследование магнитопластического эффекта в монокристаллах цинка/ В.И.Альшиц [и др.] // Кристаллография. – 1990. – Т.35, вып.4. – С. 1014–1016.
18. Абраимов, В.В. Влияние магнитного поля на низкотемпературную пластическую деформацию некоторых нормальных ГЦК металлов / В.В. Абраимов // ФНТ. – 1980. – Т.6, №10. – С. 1334–1343.
19. Дацко, О.И. Внутреннее трение в магнитообработанном материале с дислокациями / О.И. Дацко, В.И. Алексеенко // ФТТ. – 1997. – Т. 39, №7. – С. 1234–1236.
20. Грабко, Д.З. Механические свойства полуметаллов типа висмута / Д.З. Грабко, Ю.С. Боярская, М.П. Дынту. – Кишинев: Штиинца, 1982. – 132 с.
21. Остриков О.М. Влияние скорости нагружения на механизм пластической деформации в висмуте / О.М. Остриков, С.Н. Дуб // ЖТФ. – 2001. – Т.71, вып.5. – С. 44–46.
22. Косевич, А.М. Дислокационная теория упругого двойникования / А.М. Косевич, В.С. Бойко // УФН. – 1971. Т.104, вып.2 – С. 201–255.
23. Лифшиц, И.М. О макроскопическом описании явления двойникования кристаллов / И.М. Лифшиц // ЖЭТФ. – 1948. – Т.18, вып.12. – С. 1134–1143.
24. Бойко, В.С. Силы трения и поверхностного натяжения двойникующих дислокаций / В.С.Бойко, Р.И. Гарбер, Л.Ф. Кривенко // ФТТ. – 1967. – Т.9, вып.2. – С. 435–443.
25. Солдатов, В.П. О равновесной форме двойника, затормозившего у препятствия / В.П. Солдатов, В.И. Старцев // ДАН СССР. – 1966. – Т.166, вып.3. – С. 588–591.

Материал поступил в редакцию 14.11.11

PINCHUK A.I., SHAVREJ S.D. The behavior of twin-matrix boarders in bismuth crystals when constant magnetic field and concentrated applied simultaneously

It has been established that the simultaneous application of constant magnetic field and concentrated load to bismuth crystals changes substantially the behavior of twin-matrix boarders. The lengths of twins is smaller. The average density of twinning dislocation in the area of twins' tip is sharply increased. As a result, the geometry of wedge-shaped twins changes and twin boarder is shifted noticeable from twinning plane.

УДК 519.853.3

Ракецкий В.М.

К МИНИМИЗАЦИИ ВЫПУКЛЫХ ФУНКЦИЙ С ПРОСТЫМИ ОГРАНИЧЕНИЯМИ

1. Постановка задачи. Рассмотрим задачу выпуклого программирования

$$f(x) \to \min, \qquad (1)$$

при простых ограничениях

$$d_* \le x \le d^*, \qquad (2)$$

где $f(x)$ – выпуклая дважды дифференцируемая функция, $x, d_*, d^* \in R^n$. Для решения задачи (1) (без ограничений) успешно используются различные методы сопряженных направлений [1–3]. Однако наличие простых (в буквальном смысле этого слова) ограничений (2) существенно снижает эффективность этих методов, так как при каждом выходе на границу области допустимых точек процедуру построения сопряженных направлений приходится начинать заново.

Ниже предлагается метод, который естественным образом учитывает структуру ограничений (2) и не требует «обнуления» итерационной процедуры при выходе на границу допустимой области.

2. Простая задача квадратичного программирования. Допустим, что

$$f(x) = \frac{1}{2} x' D x + c' x, \qquad (3)$$

где $D = D(J, J) = \{d_{ij}, i, j \in J\} \ge 0$ – постоянная матрица, $J = \{1, 2, \ldots, n\}$. Используем для решения этой задачи прямой опорный метод [4–6].

Вектор $x = x(J)$, удовлетворяющий ограничениям задачи (1)–(3), назовем планом. Множество индексов $J_{оц} \subset J$, возможно пустое, назовем опорой целевой функции, если $\det D_{оц} \ne 0$ при $J_{оц} \ne \varnothing$. Совокупность $\{x, J_{оц}\}$ из плана и опоры целевой функции назовем опорным планом задачи (1)–(3).

Введем вектор оценок (градиент целевой функции)

$$\Delta = Dx + c.$$

Опорный план назовем согласованным (СОП), если

Ракецкий Валерий Михайлович, *зав. кафедрой информатики и прикладной математики Брестского государственного технического университета.*
Беларусь, БрГТУ, 224017, г. Брест, ул. Московская, 267.

$\Delta_{оц} = \Delta(J_{оц}) = 0$. СОП $\{x,\ J_{оц}\}$, по определению, невырожден, если

$$d_{*j} < x_j < d_j^*, \quad j \in J_{оц}.$$

В том случае, когда условие (5) не выполняется, СОП $\{x,\ J_{оц}\}$ называется вырожденным. Компоненту плана x_j, $j \in J$ назовем критической, если $x_j = d_{*j}$ или $x_j = d^*_j$.

Из общей теории необходимых и достаточных условий минимума и сделанных определений следует:

Критерий оптимальности. Для оптимальности СОП $\{x,\ J_{оц}\}$ необходимо и достаточно выполнение соотношений:

$$\Delta_j \ge 0 \text{ при } x_j = d_{*j};\ \Delta_j \le 0 \text{ при } x_j = d_j^*;$$

$$\Delta_j = 0 \text{ при } d_{*j} < x_j < d_j^*,\ j \in J_н = J \setminus J_{оц}. \qquad (4)$$

Направление l назовем допустимым для СОП $\{x,\ J_{оц}\}$, если найдется такое $\Theta^*>0$, что для любого Θ, $0<\Theta\le\Theta^*$, пара $\{\tilde{x}, J_{оц}\}$, $\tilde{x} = x + \theta l$, является СОП. Каждое допустимое для СОП $\{x,\ J_{оц}\}$ направление (при $J_{оц} \ne \varnothing$) удовлетворяет [5,6] равенству:

$$l_{оц} = -D_{оц}^{-1} D(J_{оц}, J_н) l_н. \qquad (5)$$

В силу определения, для любого невырожденного СОП $\{x,\ J_{оц}\}$ существует допустимое направление.

Допустимое для СОП $\{x,\ J_{оц}\}$ направление назовём подходящим, если

$$\frac{\partial F(x)}{\partial l} < 0.$$

Для каждого неоптимального невырожденного СОП существует [5, 6] подходящее направление l, неопорная компонента $l_{оц}$ которого строится по формулам:

$$l_{j_0} = -sign\Delta_{j_0}, \quad l_j = 0, \quad j \in J_н \setminus j_0, \qquad (6)$$

а опорная компонента $l_{оц}$ (при $J_{оц} \ne \varnothing$) находится из соотношения (5).

На основании введенных понятий и определений опишем итерацию прямого метода [4–6] для решения задачи (1)–(3). Допустим, что к началу итерации известен СОП $\{x,\ J_{оц}\}$, для которого не выполняются соотношения оптимальности (4), а также построена матрица $G_{оц} = D_{оц}^{-1}$ (если $J_{оц} \ne \varnothing$). Для упрощения рассуждений предположим, что СОП $\{x,\ J_{оц}\}$ невырожден. Тогда итерация состоит из следующих этапов:

1) среди оценок Δ_j, $j \in J_н$, не удовлетворяющих условиям оптимальности, выберем оценку Δ_{j_0} с максимальной по модулю величиной. Подходящее для СОП $\{x,\ J_{оц}\}$ направление построим по формулам (5), (6). В силу (6) из (5) получаем:

$$l_{оц} = G_{оц}\, p_{оц}\, sign\,\Delta_{j_0} \text{ при } J_{оц} \ne \varnothing, \qquad (7)$$

где $p_{оц} = D(J_{оц}, j_0)$.

2) вычислим шаг θ^0 вдоль направления l:

$$\theta^0 = \min\{\theta_{оц}, \theta_\alpha, \theta_{j_0}\}, \qquad (7)$$

где $\theta_{оц}$ – ограничение на шаг, вытекающее из учета ограничений по опорным компонентам:

$$\theta_{оц} = \begin{cases} \infty, \text{ если } J_{оц} = \varnothing, \\ \min\limits_{j \in J_{оц}} \theta_j, \text{ если } J_{оц} \ne \varnothing, \end{cases}$$

$$\theta_j = \begin{cases} (d_{*j} - x_j)/l_j, \text{ если } l_j < 0, \\ (d_j^* - x_j)/l_j, \text{ если } l_j > 0, \\ \infty, \text{ если } l_j = 0. \end{cases} \qquad (8)$$

θ_α – ограничение на шаг, вытекающее из условия минимума целевой функции вдоль направления $f(x + \theta_\alpha l) = \min\limits_\theta f(x + \theta l)$:

$$\theta_\alpha = \begin{cases} \left|\Delta_{j_0}\right| / \alpha, \text{ если } \alpha > 0, \\ \infty, \text{ если } \alpha = 0, \end{cases} \qquad (9)$$

$$\alpha = l'Dl = d_{j_0 j_0}, \text{ если } J_{оц} = \varnothing, \qquad (10)$$

и

$$\alpha = l'Dl = l'_{оц} p_{оц} sign\,\Delta_{j_0} + d_{j_0 j_0}, \text{ если } J_{оц} \ne \varnothing; \qquad (11)$$

θ_{j_0} – ограничение на шаг, вытекающее из учета ограничений по компоненте x_{j_0}.

3) вычислим новый план $\tilde{x} = x + \theta^0 l$. По построению пара $\{\tilde{x}, J_{оц}\}$ СОП, т.е. $\overline{\Delta}(J_{оц}) = 0$,

где $\overline{\Delta}$ – вектор оценок, соответствующий $\{\tilde{x}, J_{оц}\}$.

4) в зависимости от результатов вычисления максимально допустимого шага построим новую опору $\tilde{J}_{оц}$ и соответствующую ей матрицу $\tilde{G}_{оц} = \left[D[\tilde{J}_{оц}, \tilde{J}_{оц}]\right]^{-1}$:

а) $\theta^0 = \theta_{j_0}$. Полагаем $\tilde{J}_{оц} = J_{оц}$ и переходим к следующей итерации. Поскольку опора целевой функции не изменилась, $\tilde{G}_{оц} = G_{оц}$ (при $J_{оц} \ne \varnothing$).

б) $\theta^0 = \theta_\alpha$. Полагаем $\tilde{J}_{оц} = J_{оц} \cup j_0$. Если $J_{оц} = \varnothing$, то

$$\tilde{G}_{оц} = \left\{\frac{1}{d_{j_0 j_0}}\right\}; \qquad (12)$$

в противном случае $\tilde{G}_{оц}$ вычисляется по формуле [4]:

$$\tilde{G}_{оц} = \begin{pmatrix} G_{оц} + \dfrac{l_{оц} l'_{оц}}{\alpha} & -\dfrac{l_{оц}}{\alpha} sign\Delta_{j_0} \\ -\dfrac{l'_{оц}}{\alpha} sign\Delta_{j_0} & \dfrac{1}{\alpha} \end{pmatrix}. \qquad (13)$$

в) $\theta^0 = \theta_{оц} = \theta_{i_0}$. Полагаем $\tilde{J}_{оц} = J_{оц} \setminus i_0$. При этом, если $\tilde{J}_{оц} \ne \varnothing$, то

$$\tilde{G}_{оц} = G(\tilde{J}_{оц}, \tilde{J}_{оц}) - \frac{G(i_0, \tilde{J}_{оц}) G(\tilde{J}_{оц}, i_0)}{G(i_0, i_0)}. \qquad (14)$$

В [4–6] показано, что построенная по правилам 1)–4) пара $\{\tilde{x}, \tilde{J}_{оц}\}$ представляет собой согласованный опорный план. При этом $f(\tilde{x}) < f(x)$.

Итерационное применение правил 1)–4) к задаче (1)–(3) позволяет за конечное число итераций найти ее оптимальный план.

3. Простая задача выпуклого программирования. Вернемся к задаче (1), (2) с произвольной выпуклой функцией $f(x)$. Адаптируем определения и построения, приведенные в п.2, к новой ситуации. Вектору оценок Δ вернем изначальное название – градиент целевой функции – и определение:

$$\Delta = \frac{\partial f(x)}{\partial x}.$$

Зададим величину $\varepsilon>0$ – точность отыскания оптимального плана задачи (1)–(2). Будем считать, что точка x – решение задачи (1)–(2) с заданной точностью, если

$$\left|\Delta_j\right| < \varepsilon,\ j \in J.$$

Матрицу D определим как матрицу вторых производных целевой функции.

Множество индексов $J_{оц} \subset J$, возможно пустое, назовем опорой целевой функции, если $\det D_{оц} \neq 0$ при $J_{оц} \neq \varnothing$. Совокупность $\{x,\ J_{оц}\}$ из плана и опоры целевой функции назовем согласованным опорным планом задачи (1)–(2), если

$$\left|\Delta_j\right| < \varepsilon,\ j \in J_{оц}. \quad (15)$$

Пусть $\{x,\ J_{оц}\}$ – СОП. Применим к нему преобразования 1)–4) из п.2. Поскольку $f(x)$ – выпуклая неквадратичная функция, то 1) новая пара $\{\tilde{x}, \tilde{J}_{оц}\}$ может оказаться несогласованной; 2) марица $\tilde{G}_{оц}$ лишь приблизительно соответствует матрице $\tilde{D}_{оц}$.

Проблема 1) решается достаточно просто. Дополним метод этапом 5). На этом этапе будем проверять согласованность (15) опорного плана $\{\tilde{x}, \tilde{J}_{оц}\}$. Те индексы $j \in \tilde{J}_{оц}$, для которых $\left|\Delta_j\right| > \varepsilon$, выведем из опоры целевой функции, получив опору $\tilde{\tilde{J}}_{оц}$. Вопрос о пересчете матрицы $G_{оц}$ требует дополнительного обсуждения. Поскольку на этапе 5 возможно одномоментное выведение из опоры целевой функции нескольких индексов, то целесообразно оценить трудоемкость а) получения матрицы $\tilde{\tilde{G}}_{оц}$ из матрицы $\tilde{G}_{оц}$ по формулам типа (14); б) вычисления матрицы $\tilde{\tilde{G}}_{оц}$ посредством обращения матрицы $\tilde{\tilde{D}}_{оц} = D[\tilde{\tilde{J}}_{оц}, \tilde{\tilde{J}}_{оц}]$, например, по формулам типа (12), (13). При сопоставимой трудоемкости вариант б) по понятным причинам предпочтительнее.

Проблема 2) более сложна, так как только для ее распознавания (соответствует или не соответствует $\tilde{G}_{оц}$ матрице $\tilde{D}_{оц}$) требуются достаточно сложные и трудоемкие вычисления. Поэтому предлагается контролировать это соответствие косвенно на этапе вычисления максимально допустимого шага. Несоответствие матрицы $G_{оц}$ матрице $D_{оц}$ может проявиться следующим образом:

а) при вычислении величины α окажется, что $\alpha < 0$.

б) $\alpha \geq 0$, но $f(x + \theta_\alpha l) \geq f(x)$.

Для случая а) модифицируем правила (7)–(10) вычисления максимально допустимого шага: если $\alpha < 0$, то положим $\theta_\alpha = \infty$.

Независимо от знака α после расчета θ^0 проверим дополнительное условие

$$f(x + \theta^0 l) < f(x) - \eta, \quad (16)$$

где η – параметр метода. Если оно выполняется, то продолжим вычисления по обычным правилам.

В противном случае вернемся к уточнению максимально допустимого шага θ^0. Очевидно, что в рассматриваемом случае точка минимума функции $f(x)$ вдоль луча $(x + \theta l)$ находится внутри допустимой области, однако формулы (9)–(11) не обеспечивают необходимой точности. Поэтому величину θ_α найдем из условия минимума целевой функции вдоль направления

$$f(x + \theta_\alpha l) = \min_\theta f(x + \theta l), \quad (17)$$

использовав для этого один из известных методов одномерного поиска экстремума [1, 2]. Необходимый для продолжения вычислений параметр α пересчитаем по «обратной» к (10) формуле

$$\alpha = \left|\Delta_{j_0}\right| / \theta_\alpha.$$

Проверка соотношения (16) позволяет добиться гарантированного уменьшения целевой функции от итерации к итерации на величину, превосходящую η. Понятно, что параметр η не может оставаться постоянным и должен постепенно уменьшаться в процессе работы метода. Предлагается пересчитывать параметр η всякий раз, когда нарушается соотношение (16), по правилу

$$\eta = (f(x) - f(x + \theta_\alpha l)) / 2. \quad (18)$$

Формула (18) носит адаптивный характер и практически снимает проблему начального значения параметра η: в процессе работы метода будет откорректировано как слишком большое, так и слишком маленькое значение начальное значение η.

4. Некоторые аспекты численной реализации. Описанный в п. 2.3 метод можно отнести к группе методов сопряженных направлений, поскольку используемые в процессе его работы направления (6), (7) обладают свойством сопряженности. В силу использования в процессе вычислений специальной матрицы $G_{оц}$, он напоминает собой квазиньютоновские методы [1-3]. Для задачи (1), (2) предложенный метод в сравнении с известными методами сопряженных направлений (квазиньютоновскими методами) обладает важным преимуществом: он естественным образом учитывает выход на границу допустимой области и не требует «обнуления» системы сопряженных направлений при выходе на границу. Обработка выхода на границу осуществляется на этапе 4) по правилам случая в).

В определении понятия согласованного опорного плана содержится требование (15). Это требование позволяет при проверке достаточных условий оптимальности вычислять только неопорные компоненты градиента $\Delta_j,\ j \in J_н$ и тем самым уменьшать объем вычислений. Проверку согласованности опорного плана $\{x,\ J_{оц}\}$ следует выполнять в двух случаях 1) при нарушении условия (16); 2) при завершении работы метода, когда неопорные компоненты градиента удовлетворяют условиям оптимальности.

В соответствии с традиционной классификацией численных методов оптимизации предложенный метод относится к классу методов 2-го порядка: при построении допустимого направления используются элементы матрицы вторых производных. Понизить класс метода можно, если для вычисления элементов матрицы вторых производных (а также, если необходимо, и градиента) использовать численное дифференцирование.

СПИСОК ЦИТИРОВАННЫХ ИСТОЧНИКОВ

1. Поляк, Б.Т. Введение в оптимизацию. – М.: Наука, 1983. – 384 с.
2. Васильев, Ф.П. Методы оптимизации. – М.: Факториал Пресс, 2002. – 824 с.
3. Ветошкин, А.М. Структура квазиньютоновских методов минимизации // Журнал вычислительной математики и математической физики. – 2010. – Т. 50. – № 5. – С. 817–831.
4. Габасов, Р. К методам решения общей задачи выпуклого квадратичного программирования. / Р. Габасов, Ф.М. Кириллова, В.М. Ракецкий // Докл. АН СССР. – 1981. – Т. 258. – N 6. – С. 1289–1293.
5. Ракецкий, В.М. Прямой опорный метод квадратичного программирования. // Проблемы оптимального управления: сб. научн. тр. – Минск: Наука и техника, 1981. – С. 318–335.
6. Габасов, Р. Конструктивные методы оптимизации. / Р. Габасов, Ф.М. Кириллова, О.И. Костюкова, В.М. Ракецкий – Мн.: Изд-во "Университетское", 1987. – Ч. 4: Выпуклые задачи – 223 с.

Материал поступил в редакцию 03.01.12

RAKETSKI V.M. To minimization of convex functions with simple limitations.

A new method for minimization of convex functions with simple limitations is described.