УДК 519.68: 681.513.7

ПРИМЕНЕНИЕ МЕТОДА ПОСЛЕДОВАТЕЛЬНЫХ ПРИБЛИЖЕНИЙ С МНЕРЦИЕЙ К РЕШЕНИЮ ЗАДАЧИ 3-SAT

Файзуллин Р.Т., Салаев Е.В.

Омский государственный университет им. Ф.М.Достоевского,
факультет компьютерных наук
пр-т Мира, 55, г. Омск, Россия, 644077
e-mail: RTF@UNIVER.OMSK.SU

Abstract

In this paper we consider general SAT problem as a problem of global optimization for associated functional. Without loss of generality it can be considered for 3-SAT problem too. Our aim was not local search but applying Newton's methods and Newton-like procedures for direct solution. We show so for the naive Newton method has a nontrivial kernel and there are no any kind of convergence. But, we have success on the way of modification for simple iteration method with 'inertia'.

Одной из наиболее интересных задач дискретной математики является задача поиска решающего набора в задаче ВЫПОЛНИМОСТЬ [1]. После классической работы Кука [2] *усилия многих исследователей были направ-лены на построение прямых эвристических алгоритмов решения КНФ* [3]. Наиболее перспективным подходом представляется сведение SAT к непрерывному аналогу,– к задаче поиска точек глобального минимума ассоциированного функционала. *В данной работе обосновывается выбор функционала специального вида и предлагается применить к решению системы нелинейных уравнений, задающих стационарные точки функционала, модифицированный метод последовательных приближений.*

Рассмотрим переход от задачи ВЫПОЛНИМОСТЬ (SAT):

$$L = C_1 \wedge C_2 \wedge \ldots \wedge C_M,$$

$$C_j = (\vee_\xi x_l^r \in \xi, r \in \zeta, \xi \in \Xi, \zeta \in \Theta$$

$$\Xi \subset 2^N, \zeta \in B^\xi(0,1)$$

$$x_i^0 = x_i, x_i^1 = \overline{x}_i)$$

к задаче поиска глобального минимума функционала вида:

$$F(y_1, y_2, \ldots, y_N) = \sum_{s=1}^{M} c_j$$

$$c_s = \prod_{i=1}^{N} q_{i,j}$$

$$\min_{y \in E^N} F(y)$$

Здесь $q_{i,j}$ равно $|y_i \mp 1|^p$ если литерал x_i, или его дополнение принадлежит скобке C_j и единице в противном случае. Соответствие между булевыми и вещественными переменными следующее: булевой переменной 1 отвечает вещественное число 1, булевой переменой 0 отвечает вещественное значение -1. Легко заметить, что минимум равен 0, и достигается на наборе, компоненты которого равны и соответствует достижению значения ИСТИНА на исходной КНФ. В работах [4], [5] рассмотрен подход, основанный на локальном поиске экстремума, но как показал опыт, наличие локальных экстремумов существенно ограничивает данную методику.

Рассмотрим систему нелинейных уравнений определяющую стационарные точки при $p = 2$:

$$\frac{\partial F}{\partial y_i} = 0 \tag{1}$$

или

$$A(y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_N) y_i = B(y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_N)$$

Коэффициенты A и B получаются суммированием вкладов по всем скобкам содержащим yi. Но отметим то, что тогда как вклады в коэффициент A всегда положительны, вклады в коэффициент B зависят от знака ∓ 1, т.е. при суммировании большого числа вкладов правая часть будет стремиться к нулю, например на произвольном векторе y^0, который мы выбираем в качестве начального приближения. Отсюда следует, что использовать уравнение (1) для определения решения не представляется возможным.

Чтобы обойти данную техническую трудность рассмотрим другое представление:

$$x_i \vee x_j \longrightarrow y_i + y_j$$

$$x_i \wedge x_j \longrightarrow y_i^2 y_j^2$$

$$\overline{x_i} \longrightarrow (1 - y_i)^2$$

Без потери общности можно рассмотреть следующий функционал:

$$J = \sum_{s=1}^{M} \prod_{\xi} z_i^2 z_j^2 z_k^2$$

где суммирование происходит по всем M триплетам составляющим 3-ДНФ и индекс ξ равен $(i, j, k, 1, 1, 1)$ при вхождении x_i, x_j, x_k в какой-либо триплет без отрицания, при этом $z_i = y_i, z_j = y_j, z_k = y_k$,соответственно $\xi = (i, j, k, -1, 1, 1)$ и $z_i = (1 - y_i)$ если в исходной 3-ДНФ имеется триплет $\overline{x}_i x_j x_k$ и т.д.

Дифференцируя J по переменным y_i мы получаем систему нелинейных алгебраических уравнений вида:

$$\sum_{\xi\in\,\Xi} z_j^2 z_k^2 y_i = \sum_{\xi\in\Lambda} z_j^2 z_k^2 \qquad (2)$$
$$i = 1, 2, \ldots, N$$

где множества индексов определяются, как $\Xi = (\xi, i \in \xi), \quad \Lambda = (\xi, i \in \xi, \xi_{i+3} = -1)$

В данном представлении соответствующие коэффициенты также связаны соотношением $A(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_N)y_i = B(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_N)$, но в отличие от (1) вклады в правую часть уже не сокращаются, а суммируются. Также, следует пояснить выбор представления исходной КНФ именно в виде эквивалентной 3-ДНФ. Дело в том, что при попытке решения (**??**), каким угодно способом, мы будем вынуждены проводить арифметические операции с вкладами в обе части (**??**), имеющими мантиссу фиксированной длины, но при произвольной длине скобок, эта процедура будет естественным образом приводить к большим ошибкам округления. Ограничивая число переменных в скобках, мы исключаем и эту техническую трудность.

Рассмотрим систему (2), как нелинейное операторное уравнение:

$$\Phi(y) = 0$$

и попытаемся применить к его решению метод Ньютона. Производная оператора вычисляется согласно формуле:

$$\lim_{t\to 0} \frac{\Phi(y + th) - \Phi(y)}{t} = \Phi_y^{'}(h)$$

а итерацию метода Ньютона проводятся согласно:

$$y_{p+1} = y_p - \Phi_{y_p}^{'-1}(\Phi(y_p))$$

Вычислительные эксперименты показали крайнюю неустойчивость данной итерационной процедуры, что объясняется вырожденностью производного оператора на решении, например, легко проверяется, что само решение входит в ядро производного оператора:

$$(\sum_{\xi\in\,\Xi} z_m^2 z_n^2)h_l + 2((\sum_{\xi\in\,\Xi} h_m z_n^2)z_l + (\sum_{\xi\in\,\Xi} z_m^2 h_n)z_l) = (\sum_{\xi\in\,\Xi} z_m^2 z_n^2)z_l$$

Более того, решения входит и в ядро второй производной оператора:

$$(\sum_{\xi\in\,\Xi} z_m z_n)z_l = 0$$

Данное обстоятельство делает практически неприменимыми локальные градиентные методы. Как альтернатива методу Ньютона может быть предложен метод последовательных приближений:

$$\sum_{\xi\in\,\Xi} z_j^2(t) z_k^2(t) y_i(t+1) = \sum_{\xi\in\Lambda} z_j^2(t) z_k^2(t)$$

где после нескольких зейделевских итераций происходит проекция вектора $y(t)$ на $B^N(0,1)$ и осуществляется проверка набора $x_1,\ldots,x_N$ на выполнимость для исходной КНФ.

Сходимость в этом случае гарантирована ограниченностью решения и естественно выполняющимся условием $\| \nabla\frac{B}{A} \| \leq 1$. Вычислительные эксперименты, проводившиеся для случайного наполнения наборов скобок SAT и 3-SAT, показали, что при соотношении $\vartheta = \frac{N}{M} \leq 0.01$, где N это число переменных, M число скобок в 3-SAT, итерационная процедура всегда сходится к решению. Но также было выяснено, что при увеличении ϑ скорость сходимости резко падает и, хотя большинство компонент решения формируется верно, и быстро, оставшаяся часть оказывается практически недостижимой в силу накопления ошибок округления. Для исправления данного обстоятельства была предложена модификация метода последовательных приближений, так называемый метод последовательных приближений с инерцией, хорошо зарекомендовавший себя при решении систем нелинейных уравнений гидравлики [6]:

$$\sum_{\xi\in\,\Xi}(\sum_{p=0}^{K}\alpha_p z_j^2(t-p)z_k^2(t-p))y_i(t+1) = \sum_{\xi\in\,\Xi} z_j^2(t)z_k^2(t) \tag{3}$$

$$\sum_{p=0}^{K}\alpha_p = 1$$

Смысл модификации заключается в том, что формирование приближений происходит не так быстро, как в исходном методе, и мы избегаем областей притяжения паразитных аттракторов. Как показали численные эксперименты наилучшие результаты получаются при выборе в качестве α_i сверхубывающей последовательности (каждый член последовательности больше чем сумма всех последующих). Отметим, что предыдущее выражение можно рассматривать как аппроксимацию системы интегральных уравнений, с экспоненциально убывающей мерой.

$$(\sum_{\xi\in\,\Xi}\frac{1}{T}\int\limits_{t-T}^{t}\phi(u)y_m^2(u)y_n^2(u)y_i(t+1) = \sum_{\xi\in\,\Xi} y_m^2(t)y_n^2(t)$$

Численные эксперименты со случайно сгенерированными КНФ (вплоть до $N\sim 10^6, M\sim 10^6, \vartheta\approx 1/2$) показали практическую эффективность алгоритма и практически линейную зависимость времени выполнения, как от числа скобок, так и от числа переменных. Расчеты проводились на неспециализированном компьютере Пентиум 4 с тактовой частотой 3ГГц, объемом оперативной памяти 1ГБ, и, например, время выполнения алгоритма для $M = 2*10^6, N = 10^6$ равнялось 40 минутам. Также, было продемонстировано, что для случайных КНФ имеющих решение, всегда удается найти решение и отклонений с экспоненциальным ростом времени выполнения алгоритма не наблюдается.

Были предприняты попытки применения алгоритма к решению КНФ, ассоциированных с задачей факторизации. При этом задача нахождения простых сомножителей p и q для $n = pq$ предварительно сводилась к задаче ВЫПОЛНИМОСТЬ для специальной 3-КНФ (В.И.Дулькейт, частное сообщение) Для длины представления n равной 1024 бит получены следующие характеристики: $M \approx 1.2*10^7, N \approx 5*10^5$. Ситуация оказалось обратной предыдущей, т.е. для малых отвечающих общей задаче решение найти не удается при любом обозримом числе рестартов, но при определенных условиях, т.е. когда мы берем всего часть скобок КНФ и $\vartheta \geq 1/6$, частное решение легко находится. Данное обстоятельство позволяет надеяться, что распараллеливание исходной процедуры позволит в дальнейшем добиться некоторого успеха и в этой важной для приложений задаче.

Если рассматривать исходную логическую форму L как совокупность линейных алгебраических уравнений, например, ассоциируя с C_j уравнение $\sum x_i = l_j$, где $l_j \geq 1$, мы можем построить перемежающийся итерационный процесс, т.е. после нескольких итераций (3) мы подставляем x ассоциированные с y в линейные уравнения и корректируем правые части так, чтобы все компоненты правой части были ненулевыми. Полученная систему линейных алгебраических уравнений решается для определения нового значения y^0. Это не решает задачу факторизации принципиально, но позволяет несколько продвинуться в сторону уменьшения ϑ.

Отметим, что здесь наблюдается некоторая аналогия с задачами гидравлики для сетей, так линейные уравнения можно интерпретировать, как условия выполнения «первого закона Кирхгофа», а систему нелинейных уравнений, как аналог аппроксимации «второго закона Кирхгофа»с «напорами»(правыми частями), зависящими от «расходов»(компонент вектора y).

Заключение

Основным результатом данной статьи является построение итерационной процедуры, позволяющей решать большие системы нелинейных уравнений, ассоциированные со случайно сгенерированными КНФ.

Представляется перспективным дальнейшее изучение и разработка аналогичных процедур, с учетом распараллеливания вычислений и применительно к КНФ, отвечающих логической форме задачи факторизации.

Список литературы

1. *Стайглиц, Пападмитриу А.Я.* Комбинаторная оптимизация. - М.: Наука, 1974. - 432 с.
2. *Кук С. А.* Сложность процедру вывода теорем. Кибернетический сборник: Новая серия,- 1975. - Вып. 12.- С. 5-15.
3. *Gu J., Purdom P.W., Franco J.*, Wah B. Algorithms for the Satisfiability problem: A survey DIM ACS Series on Discrete Mathematics and Theoretical Computer Science. V.35. - American Mathematical Society, 1997. - P. 19-151.

4. *Gu J.* How to solve Very Large -Scale Satisfiability problems. Theoretical report UUCS-TR-88-032, 1988.
5. *Gu J. UniSAT* problem models (appendix). IEEE Trans, on Pattern Analisys and Machine Intelligence, 14 (8):865, Aug. 1992.
6. *Файзуллин Р.Т.* О решении нелинейных алгебраических систем гидравлики. Сибирский журнал индустриальной математкики - 1999. - Т. 2, №2. - С.36-38.