

УДК: 519.85

## Подход к разработке алгоритмов ньютоновских методов безусловной оптимизации, программная реализация и сравнение эффективности

Г. А. Зеленков<sup>1</sup>, А. Б. Хакимова<sup>2,а</sup>

<sup>1</sup> ФГБОУ ВПО «ГМУ имени адмирала Ф. Ф. Ушакова»,  
Россия, 353911, г. Новороссийск, пр. Ленина 93

<sup>2</sup> ФГБОУ ВПО «Кубанский государственный университет» филиал в г. Новороссийске,  
Россия, 353922, г. Новороссийск, ул. Героев-Десантников 87

E-mail: <sup>а</sup> puma\_home@mail.ru

Получено 4 марта 2013 г.,  
после доработки 2 мая 2013 г.

Предложен подход к увеличению эффективности алгоритма Гилла и Мюррея к построению ньютоновских методов безусловной оптимизации с регулировкой шага, основанных на факторизации Холецкого. Доказано, что стратегия выбора направления спуска определяет и решение проблемы масштабирования шагов при спуске, и аппроксимацию не квадратичными функциями, и интеграцию с методом доверительной окрестности.

Ключевые слова: информационные технологии, алгоритмы, метод Ньютона.

## Approach to development of algorithms of Newtonian methods of unconstrained optimization, their software implementation and benchmarking

G. A. Zelenkov<sup>1</sup>, A.B. Khakimova<sup>2</sup>

<sup>1</sup> Admiral Ushakov State Maritime University, 93 Lenin's avenue, Novorossiysk, 353911, Russia

<sup>2</sup> FSEI of HPE "Kuban State University" branch in Novorossiysk, 87 Geroev-Desantnikov street, Novorossiysk, 353922, Russia

**Abstract.** — The approach to increase efficiency of Gill and Murray's algorithm of Newtonian methods of unconstrained optimization with step adjustment creation is offered, rests on Cholesky's factorization. It is proved that the strategy of choice of the descent direction also determines the solution of the problem of scaling of steps at descent, and approximation by non-quadratic functions, and integration with a method of a confidential vicinity.

Keywords: information technology, algorithm, Newton's method

Citation: *Computer Research and Modeling*, 2013, vol. 5, no. 3, pp. 367–377 (Russian).

## Вычислительная схема алгоритма Гилла и Мюррея

Пусть  $h^k$ ,  $H^k$  — градиент, и гессиан, вычисленные на итерации  $k$  в точке  $x^k$  процесса безусловной минимизации гладкой функции  $F(x)$ , тогда в приведенных обозначениях принцип построения большинства ньютоновских методов с регулировкой шага опишем так. На каждой итерации  $k$  сначала строится некоторая, связанная с исходной матрицей вторых производных  $H^k$ , положительно определенная матрица  $\bar{H}^k$ , а затем направление спуска  $p^k$  вычисляется как решение системы  $\bar{H}^k p^k = -h^k$ . Поскольку положительная определенность  $\bar{H}^k$  гарантирована, то вектор  $p^k$  будет направлением спуска. При этом процедуру построения  $\bar{H}^k$  организуют так, чтобы  $\bar{H}^k$  совпадала с матрицей  $H^k$ , если последняя сама является положительно определенной. Это означает, что на каждой итерации выясняется, все ли собственные числа положительны. Причем выяснение определенности  $H^k$  и построение  $\bar{H}^k$  осуществляется параллельно в рамках одной процедуры на основе некоторых матричных разложений, которые позволяют выявить знаки собственных чисел  $H^k$  и приспособиться для генерации  $\bar{H}^k$ . Одна из наиболее эффективных вычислительных схем получается, если отталкиваться от разложения  $H^k$  по Холецкому [Гилл и др., 1985].

Подход Гилла и Мюррея состоит в построении факторов Холецкого  $D^k$  и  $L^k$ , подчиняющихся двум требованиям: все диагональные элементы  $D^k$  должны быть существенно положительными, модули всех элементов треугольного фактора  $L^k$  должны быть равномерно ограничены сверху. Точнее говоря, требуется, чтобы для всех  $j=1,2,\dots,n$  и некоторых заданных положительных  $\delta$  и  $\beta$  выполнялись неравенства

$$d_{jj}^k \geq \delta, \quad |r_{ij}^k| \leq \beta, \quad i > j, \quad (1)$$

где  $r_{ij}^k$  — вспомогательные величины, по определению равные  $u_{ij}^k \sqrt{d_{jj}^k}$ . Выбор  $\beta$  обсуждается позже. При реализации алгоритма на конкретной ЭВМ, в которой под запись мантиссы отводится  $t$  битов, величину  $\delta$  предлагается вычислять по формуле  $\delta = \max\{2^{-t} \|H^k\|_{\infty}, 2^{-t}\}$ .

Процедура расчета модифицированных факторов  $L^k$ ,  $D^k$  фактически представляет собой обычный алгоритм факторизации Холецкого с попутным увеличением (по мере необходимости) диагонали исходной матрицы с целью добиться выполнения неравенств (1).

Рассмотрим  $j$ -ый шаг предлагаемой процедуры. Пусть первые  $j-1$  столбцов модифицированных факторов Холецкого известны и для  $\mu=1,\dots,j-1$  условия (1) выполнены. Вычислим величину

$$\gamma_j = \left| \xi_j - \sum_{s=1}^{j-1} d_{ss}^k (l_{js}^k)^2 \right|, \quad (2)$$

где через  $\xi_j$  обозначен диагональный элемент  $h_{jj}^k$  матрицы  $H^k$ . В качестве значения  $\bar{d}_{jj}^k$  для  $j$ -го диагонального элемента  $D^k$  возьмем  $\bar{d}_{jj}^k = \max\{\gamma_j, \delta\}$ , где  $\delta$  — малое положительное число из (1). Если же неравенства (1) при  $d_{jj}^k = \bar{d}_{jj}^k$  нарушаются, иными словами, какие-то из  $r_{ij}^k$  оказываются больше, чем  $\beta$ , окончательное значение  $d_{jj}^k$  определим по формуле, получающейся из (2) заменой  $\xi_j$  на  $h_{jj}^k + e_{jj}^k$ , где число  $e_{jj}^k$  подбирается так, чтобы максимальная из величин  $r_{ij}^k$  совпала с  $\beta$ . Матрицы  $L^k$  и  $D^k$ , полученные по окончании описанной процедуры,

будут факторами Холецкого для положительно определенной матрицы  $\overline{H}^k$ , связанной с  $H^k$  следующим образом:  $L^k D^k (L^k)^T = H^k + E^k = \overline{H}^k$ , где  $E^k$  — неотрицательная диагональная матрица,  $j$ -ый элемент которой равен  $e_{jj}^k$ . Таким образом, положительно определенная матрица  $\overline{H}^k$  может отличаться от исходной матрицы  $H^k$  только диагональными элементами.

Направление спуска  $p^k$  вычисляется как решение системы  $\overline{H}^k p^k = -h^k$ , причем  $p^k = -(\overline{H}^k)^{-1} h^k$  определяют последовательным решением двух систем линейных уравнений с треугольными матрицами:  $L^k y^k = -h^k$ ,  $D^k (L^k)^T p^k = y^k$ .

При заданной матрице  $H^k$  диагональная поправка  $E^k$  зависит от величины  $\beta$ . Желательно, чтобы эта величина была достаточно большой, так как заниженное значение  $\beta$  приведет к неоправданной модификации. Для положительно определенной матрицы  $H^k$  справедливо неравенство  $(l_{ij}^k)^2 d_{ii}^k \leq h_{ii}^k$ . Значит, чтобы обеспечить равенство нулю поправки  $E^k$ , если матрица  $H^k$  существенно положительно определена, в качестве  $\beta$  надо взять величину, удовлетворяющую неравенству  $\beta^2 \geq \gamma$ , где  $\gamma$  — значение максимального из диагональных элементов  $H^k$ . Когда у матрицы  $H^k$  есть отрицательные собственные значения, поправка  $E^k$  будет ненулевой независимо от выбора  $\beta$ . При этом для  $n > 1$  справедлива оценка

$$\|E^k(\beta)\|_\infty \leq \left( \frac{\xi}{\beta} + (n-1)\beta \right)^2 + 2(\gamma + (n-1)\beta^2) + \delta. \text{ Здесь } \xi \text{ — максимальный модуль недиаго-$$

нального элемента  $H^k$ . Правая часть неравенства достигает минимума при  $\beta^2 = \xi\sqrt{n^2 - 1}$  и неограниченно возрастает с увеличением  $\beta$ . Это говорит о том, что в общем случае слишком большие значения  $\beta$  столь же нежелательны, как и слишком малые. Помимо того, что завышенное значение  $\beta$  чревато потерей численной устойчивости процедуры построения  $\overline{H}^k$ , оно может еще и привести к неоправданно большому отклонению  $\overline{H}^k$  от  $H^k$ . Выбор бесконечного  $\beta$ , грубо говоря, означает, что диагональными элементами модифицированного фактора  $D^k$  будут модули диагональных элементов обычного (если обычное разложение возможно). Учитывая приведенные соображения, величину  $\beta$  вычисляют по формуле  $\beta^2 = \max\{\gamma, \xi / \sqrt{n^2 - 1}, \varepsilon_M\}$ . Здесь  $\varepsilon_M$  — машинная точность. Она вводится в формулу расчета  $\beta$ , чтобы обеспечить устойчивость вычислений, когда норма  $H^k$  очень мала.

Процедура модифицированной факторизации Холецкого — это численно устойчивый алгоритм, генерирующий положительно определенную матрицу, которая может отличаться от исходной матрицы только диагональными элементами. Характеризуют ее и как оптимизированный алгоритм в том смысле, что параметр  $\beta$  подбирается в нем путем минимизации априорной поправки  $E^k$  при условии сохранения существенно положительно определенной матрицы неизменной. Следует также отметить, что реальная величина нормы  $E^k$  почти всегда оказывается меньше априорной оценки. Фактическое значение нормы  $E^k$  можно дополнительно уменьшить, если использовать симметричные перестановки столбцов и строк  $H^k$ . На очередном,  $j$ -ом шаге факторизации в качестве  $j$ -ой строки и  $j$ -го столбца надо брать ту из нетронутых пар, для которой величина  $\gamma_j$  в (2) максимальна. Такая стратегия приведет к разложению вида  $P^t H^k P + E^k = L^k D^k (L^k)^T$ , где  $P$  есть некоторая перестановочная матрица. В методе

с перестановками поправка  $E^k$  инвариантна относительно нумерации переменных. Ниже приводится детальное описание всех операций, выполняемых по ходу построения модифицированного разложения Холецкого с перестановками. Дана наиболее эффективная схема организации расчетов. В процессе вычисления  $j$ -го столбца матрицы  $L^k$  участвуют вспомогательные величины  $c_{is}^k = l_{is}^k d_{ss}^k, s = 1, \dots, j, i = j, \dots, n$ .

Шаг 0. (Расчет порога для элементов). Вычислить  $\beta^2 = \max\{\gamma, \xi / \nu, \varepsilon_M\}$ , где  $\nu = \max\{1, \sqrt{n^2 - 1}\}$ , а числа  $\gamma$  и  $\xi$  суть максимальные значения модулей диагонального и недиагонального элементов  $H^k$ .

Шаг 1. (Инициализация). Присвоить индексу столбца  $j$  значение 1. Положить  $c_{ii}^k = g_{ii}^k, i = 1, \dots, n$ .

Шаг 2. (Перестановка строк и столбцов). Найти индекс  $q$ , такой, что  $|c_{qq}^k| = \max_{j \leq i \leq n} |c_{ii}^k|$ . Поменять местами все данные, отвечающие столбцам матрицы  $H^k$  с номерами  $q$  и  $j$ , а затем проделать то же самое с данными, отвечающими ее  $q$ -ой и  $j$ -ой строкам.

Шаг 3. (Поиск максимальной по модулю величины  $l_{ij}^k d_{ii}^k$ ). Вычислить  $c_{ij}^k = g_{ij}^k - \sum_{s=1}^{j-1} l_{js}^k c_{is}^k$  для  $i = j+1, \dots, n$  и найти  $\theta_j = \max_{j+1 \leq i \leq n} |c_{ij}^k|$  (если  $j = n$ , взять  $\theta_j = 0$ ).

Шаг 4. (Расчет  $j$ -го диагонального элемента фактора  $D^k$ ). Вычислить  $d_{jj}^k = \max\{\delta, |c_{jj}^k|, \theta_j^2 / \beta^2\}$  и поправку  $e_{jj}^k = d_{jj}^k - c_{jj}^k$ . Если  $j = n$ , вычисления прекратить.

Шаг 5. (Расчет  $j$ -ой строки  $L^k$ ). Вычислить  $l_{js}^k = c_{js}^k / d_{ss}^k$  для  $s = 1, \dots, j-1$ . Для  $i = j+1, \dots, n$  пересчитать диагональные элементы  $c_{ii}^k$  по формуле  $c_{ii}^k = c_{ii}^k - l_{ij}^k c_{ij}^k$ . Присвоить  $j$  значение  $j+1$  и вернуться к шагу 2.

Модифицированная факторизация Холецкого помимо прочего позволяет определить и направление отрицательной кривизны, решая уравнение  $(L^k)^T p^k = e_s$ , где индекс  $s$  выбирается из условия  $d_{ss}^k - e_{ss}^k \leq d_{jj}^k - e_{jj}^k, j = 1, \dots, n$ .

## Предлагаемая вычислительная схема алгоритма

Направление спуска  $p^k$ , в конечном счете, вычисляется как решение системы  $(L^k)^T p^k = u^k$ , где  $u^k = (D^k)^{-1} y^k$ , а  $y^k$  есть решение системы  $L^k y^k = -h^k$ . Величина элементов  $|u_j^k|$  зависит от способа задания направления спуска. Следуя подходу Гилла и Мюррея, для численной устойчивости расчета элементов  $D^k, L^k, u^k$  достаточно потребовать изменения способа задания  $p^k$  так, чтобы для всех  $j = 1, 2, \dots, n$  и некоторых заданных положительных  $\delta, \beta$  и  $\omega$  выполнялись неравенства:  $d_{jj}^k \geq \delta, |r_{ij}^k| \leq \beta, i > j, |u_j^k| \leq \omega$ . Это может стать ключом к решению проблемы масштабирования шагов при спуске [Гилл и др., 1985], но такое предположение ошибочно — подход Гилла и Мюррея к вычислению направления спуска  $p^k$  исключает расчет элементов вектора  $u^k$  по ходу построения модифицированного разложения Холецкого. Нужен альтернативный подход, который опишем так. Равенства  $H^k = (U^k)^T D^k U^k$ ,

$U^k = (L^k)^T$  достаточно для определения элементов матриц  $D^k, U^k$ . В скалярной форме это равенство выглядит следующим образом:  $h_{ij}^k = \sum_{\mu=1}^i u_{\mu i}^k u_{\mu j}^k d_{\mu\mu}^k$ , отсюда, полагая  $u_{ii}^k = 1$ , построим соотношения

$$d_{ii}^k = h_{ii}^k - \sum_{s=1}^{i-1} (u_{si}^k)^2 d_{ss}^k, \quad u_{ij}^k = (h_{ij}^k - \sum_{s=1}^{i-1} u_{si}^k u_{sj}^k d_{ss}^k) / d_{ii}^k, \quad j > i \quad (3)$$

для расчета элементов  $D^k, U^k$ . Построение факторов Холецкого математически эквивалентно применению метода исключения Гаусса к системе уравнений  $H^k p^k = -h^k$  в прямом порядке,

при этом техника исключения Гаусса позволяет получить соотношение

$$u_i^k = -(h_i^k - \sum_{s=1}^{i-1} u_s^k u_{si}^k d_{ss}^k) / d_{ii}^k \quad (4)$$

для расчета элементов вектора  $u^k$  и систему уравнений  $U^k p^k = u^k$  для вычисления направления спуска  $p^k$ . Интеграция техники исключения Гаусса и факторизации Холецкого приводит к соотношениям (3), (4), которые с помощью элементарных арифметических операций позволяют вычислить элементы  $u^k$  по ходу построения факторов Холецкого. Порядок расчета удобно пояснить следующей диаграммой:  $d_{ii}^k \Rightarrow u_{i,i+1}^k, \dots, u_{in}^k, u_i^k$ .

Интеграция техники исключения Гаусса и факторизации Холецкого порождает множество численно устойчивых способов задания направления спуска  $p^k$ , в основе которых лежит требование, чтобы на очередном шаге вычислений коэффициентов факторов Холецкого соответствующий диагональный элемент матрицы  $D^k$  и соответствующий элемент вектора  $u^k$  сначала рассчитывались по вычисленным ранее значениям этих коэффициентов. Затем диагональный элемент  $D^k$  увеличивается настолько, насколько необходимо, чтобы все диагональные элементы  $D^k$  были существенно положительными, модули всех элементов  $U^k, u^k$  были равномерно ограничены сверху. Например, достаточно потребовать, чтобы для всех  $i = 1, 2, \dots, n$  выполнялись неравенства

$$d_{ii}^k \geq \delta, \quad |u_{ij}^k| \leq 1, \quad j > i, \quad |u_i^k| \leq 1, \quad (5)$$

положив  $d_{ii}^k = \max\{\delta, |c_{ii}^k|, \theta_i^2 / \beta^2, |c_i^k|, \theta_i\}$ , поскольку элементы  $d_{ii}^k, c_{ij}^k / d_{ii}^k$  и  $c_i^k / d_{ii}^k$  определяют  $i$ -ые строки матриц  $D^k, U^k, u^k$  и удовлетворяют неравенствам (5). Здесь  $c_{ii}^k, c_{ij}^k, c_i^k$  — вспомогательные величины, определяемые следующим образом:

$$c_{ii}^k = h_{ii}^k - \sum_{s=1}^{i-1} (u_{si}^k)^2 d_{ss}^k, \quad c_{ij}^k = h_{ij}^k - \sum_{s=1}^{i-1} u_{si}^k u_{sj}^k d_{ss}^k, \quad j > i, \quad c_i^k = -(h_i^k - \sum_{s=1}^{i-1} u_s^k u_{si}^k d_{ss}^k). \quad (6)$$

Следующую стратегию опишем так. В процессе построения  $D^k, U^k, u^k$  вычислим элементы  $g_i^k = d_{ii}^k / l_i^k$  (если  $d_{ii}^k > l_i^k \gamma$ , взять  $g_i^k = l_i^k / \gamma$ ), где  $l_i^k = \max\{\delta, |d_{ii}^k|\}$ , вычислить  $\gamma^k = \max_{1 \leq i \leq n} g_i^k$ , и определить направление спуска  $p^k$  из системы

$$\frac{1}{\gamma^k} \bar{H}^k p^k = -h^k. \quad (7)$$

Здесь параметр  $\gamma$  задается пользователем, число  $\gamma^k$  ( $\gamma^k \geq 1$  по построению) характеризует степень однородности, при этом (7) есть простейшая форма аппроксимации, отличная от квадратичной. Пусть  $\bar{H}^k p^k = -\gamma^k h^k = -h^k - (\gamma^k - 1)h^k = -h^k - G^k p^k$ , где  $G^k$  есть диагональная матрица с элементами  $g_{ii}^k = \frac{(\gamma^k - 1)h_i^k}{p_i^k}$  на диагонали, или  $(\bar{H}^k + G^k)p^k = -h^k$ . Это означает сдвиг на  $g_{ii}^k$

всех  $i$ -ых собственных значений матрицы  $\bar{H}^k$ . В методах с регулировкой шага матрицу  $H^k$  модифицируют так, чтобы изменения не затрагивали подпространства, натянутого на ее собственные векторы с положительными собственными значениями [Гилл и др., 1985]. Если же замена осуществляется в методе доверительной окрестности, то она отражается на всех векторах, так как результатом замены  $H^k$  на матрицу  $\bar{H}^k = H^k + \lambda^k I$  будет сдвиг на  $\lambda^k$  всех собственных значений матрицы  $H^k$  [Гилл и др., 1985]. Из вышесказанного следует, что стратегия выбора направления спуска определяет интеграцию с методом доверительной окрестности.

Ниже дана наиболее эффективная схема организации расчетов. Все фигурирующие в ней величины при реализации на ЭВМ могут размещаться в памяти, первоначально выделяемой для записи матриц  $H^k$ ,  $h^k$ . При этом коэффициенты рассчитываемых факторов занимают места ее использованных элементов. В процессе вычисления  $i$ -ой строки фактора  $U^k$  участвуют вспомогательные величины (6). В процессе построения системы уравнений  $U^k p^k = u^k$  на каждой итерации  $k$  выявляются знаки собственных чисел, и вычисляется количество отрицательных  $n_o$  и нулевых  $n_0$  собственных значений, число переходов к направлению отрицательной кривизны  $n_s$ .

Шаг 0. Вычислить  $\beta^2 = \max\{\gamma^0, \xi / \nu, \varepsilon_M\}$ , где  $\nu = \max\{1, \sqrt{n^2 - 1}\}$ , а числа  $\gamma^0$  и  $\xi$  суть максимальные значения модулей диагонального и недиагонального элементов  $H^k$ . Положить  $\varepsilon_0 = 2^{-\frac{\tau_F}{2}}$ ,  $\varepsilon_s = 2^{-\frac{\tau_F}{3}}(1 + \|F(x^{k-1}) - F(x^k)\|)$ . Здесь  $\tau_F$  — число правильных разрядов  $F(x^k)$ , которые хотелось бы получить.

Шаг 1. Положить  $i = 1$ , положить  $c_{jj}^k = h_{jj}^k$ ,  $j = 1, \dots, n$ ,  $c_j^k = -h_j^k$ ,  $j = 1, \dots, n$ .

Шаг 2. Найти индекс  $q$ , такой, что  $|c_{qq}^k| = \max_{i \leq j \leq n} |c_{ij}^k|$ , и поменять местами все данные, отвечающие строкам матрицы  $H^k$  с номерами  $q$  и  $i$ , а затем проделать то же самое с данными, отвечающими ее  $q$ -му и  $i$ -му столбцам.

Шаг 3. Если  $c_{ii}^k < \varepsilon_0$ , то положить  $l_i^k = \delta$ ; иначе положить  $l_i^k = c_{ii}^k$ .

Шаг 4. Если  $|c_{ii}^k| \leq \varepsilon_0$ , то положить  $n_0 = n_0 + 1$  и перейти к шагу 8; иначе перейти к шагу 5.

Шаг 5. Если  $c_{ii}^k > 0$ , то перейти к шагу 8; иначе перейти к шагу 6.

Шаг 6. Положить  $n_o = n_o + 1$ . Если  $\|h^k\| > \varepsilon_s$ , то перейти к шагу 8; иначе перейти к шагу 7.

Шаг 7. Построить систему уравнений  $U^k p^k = e_i$  для поиска направления отрицательной кривизны по формулам  $u_j^k = 0, j = 1, \dots, n$ ,  $u_i^k = 1$ , положить  $n_s = n_s + 1$  и остановиться.

Шаг 8. Если  $i = n$ , то положить  $\theta_i = 0$ ; иначе вычислить вспомогательные величины  $c_{ij}^k, j = i + 1, \dots, n$ , по формуле  $c_{ij}^k = h_{ij}^k - \sum_{s=1}^{i-1} u_{si}^k u_{sj}^k d_{ss}^k$  и найти  $\theta_i = \max_{i+1 \leq j \leq n} |c_{ij}^k|$ .

Шаг 9. Вычислить диагональный элемент фактора  $D^k$  и элемент вектора  $u^k$  по формулам  $d_{ii}^k = \max\{\delta, |c_{ii}^k|, \theta_i^2 / \beta^2, |c_i^k|, \theta_i\}$ ,  $u_i^k = c_i^k / d_{ii}^k$ . Вычислить  $g_i^k = d_{ii}^k / l_i^k$  (если  $d_{ii}^k > l_i^k \gamma$ , взять  $g_i^k = l_i^k / \gamma$ ). Если  $i = n$ , то перейти к шагу 11.

Шаг 10. Вычислить строку  $u_{ij}^k = c_{ij}^k / d_{ii}^k, j = i+1, \dots, n$ , фактора  $U^k$ , пересчитать вспомогательные величины по формулам  $c_{jj}^k = c_{jj}^k - u_{ij}^k c_{ij}^k, j = i+1, \dots, n$ ,  $c_j^k = c_j^k - u_{ij}^k c_i^k, j = i+1, \dots, n$ , положить  $i = i+1$  и перейти к шагу 2.

Шаг 11. Найти  $\gamma^k = \max_{1 \leq i \leq n} g_i^k$ , положить  $u^k = \gamma^k u^k$  и остановиться.

## Вычислительные схемы конечно-разностной аппроксимации

Для получения быстроходящегося алгоритма не обязательно пользоваться точными значениями производных, его можно построить и на основе их конечно-разностной аппроксимации [Гилл и др., 1985]. Во многих случаях такой подход оказывается самым подходящим. Имеются в виду ситуации, когда вычисление аналитических значений первых и вторых производных минимизируемой функции  $F(x)$  затруднено или просто невозможно [Гилл и др., 1985].

За оценку  $i$ -го столбца матрицы  $H^k$  обычно принимают его правую конечно-разностную аппроксимацию  $y_i^k = \frac{1}{l}(h^k(x^k + l e_i) - h^k(x^k))$ . Здесь  $l$  есть число, именуемое конечно-разностным интервалом, а  $e_i$  суть  $i$ -ый единичный вектор. Ради простоты изложения будем полагать, что все конечные разности вычисляются на одном и том же интервале  $l$ . На самом же деле приходится строить набор интервалов  $\{l_i\}, i = 1, \dots, n$ , и технически это делается так, как будет описано ниже.

Для аппроксимации матрицы  $H^k$  берут полусумму  $\tilde{H}^k = \frac{1}{2}(Y^k + (Y^k)^t)$ , которая симметрична по построению. Метод, получающийся из обычного ньютоновского метода заменой в уравнении  $H^k p^k = -h^k$  матрицы  $H^k$  на матрицу  $\tilde{H}^k$ , принято называть дискретным методом Ньютона [Гилл и др., 1985]. Подобно своему классическому прототипу, он порождает целый класс модифицированных методов. Предлагаемая техника модификации, разработанная для обобщения метода Ньютона на случаи со знаконеопределенными  $H^k$ , применима и к его конечно-разностному аналогу.

Достоинства дискретных ньютоновских методов те же, что и у обычных методов ньютоновского типа: высокая скорость сходимости и способность «ощущать» приближение седловой точки и уходить от нее. К недостаткам относится необходимость вычисления на каждой итерации  $n$  значений градиента, которые требуются для построения  $n$  столбцов аппроксимации матрицы Гессе, когда в качестве конечно-разностных направлений используются столбцы единичной матрицы. Поэтому при значениях  $n > 10$  или около того дискретные ньютоновские методы нередко начинают проигрывать другим методам первого порядка. Однако они могут сохранить первенство и при больших значениях  $n$ , если матрица Гессе разрежена, и известная структура расположения ее ненулей позволяет экономно организовать процедуру построения конечно-разностной аппроксимации.

Замена матрицы  $H^k$  на матрицу  $\tilde{H}^k$  ставит вопрос о выборе величины интервала  $l$ . Рассмотрим процедуру вычисления конечно-разностных интервалов, предназначенных для оптимизационных методов с оцениванием производных по конечно-разностным формулам, предложенную Полаком в работе [Полак, 1974].

Шаг 0. Выбрать  $\beta > 0, l > 0$ . Попробуйте взять  $\beta \in [5, 10], l \in [10^{-3}, 10^{-2}]$ .

Шаг 1. Положить  $i = 1$ .

Шаг 2. Положить  $l_i = l$ .

Шаг 3. Вычислить вектор  $h^k(l_i, x^k) \in R^n$ ,  $i$ -ая компонента  $h_i^k(l_i, x^k)$  которого определяется равенством  $h_i^k(l_i, x^k) = -\frac{1}{l_i}(F^k(x^k + l_i e_i) - F^k(x^k))$ ,  $i = 1, \dots, n$ , где  $e_i$  есть  $i$ -ый столбец единичной матрицы размера  $n \times n$ , то есть  $e_1 = (1, 0, \dots, 0)^t$ ,  $e_2 = (0, 1, \dots, 0)^t$  и так далее.

Шаг 4. Вычислить  $\Delta(l_i, x^k) = F^k(x^k + l_i \beta h_i^k(l_i, x^k)) - F^k(x^k)$ .

Шаг 5. Если  $\Delta(l_i, x^k) < 0$ , то перейти к шагу 6; иначе положить  $l_i = \frac{1}{2}l_i$  и перейти к шагу 3.

Шаг 6. Положить  $i = i + 1$ . Если  $i > n$ , то остановиться; иначе перейти к шагу 2.

Применять процедуру оценивания конечно-разностных интервалов на каждой итерации алгоритма поиска минимума функции  $F(x)$  слишком накладно, поэтому Гилл, Мюррей и Райт [Гилл и др., 1985] рекомендуют обращаться к процедуре один раз в точке  $x^0$ , затем учитывать изменения  $l_i$  в разных точках, получающихся по ходу минимизации так. По окончании работы процедуры оценивания конечно-разностных интервалов, значения  $l_i$  пересчитываются по формуле  $l_i = \frac{l_i}{1 + |x_i^0|}$ , а в процессе конечно-разностной аппроксимации первых и вторых производных значения  $l_i$  пересчитываются по формуле  $l_i = l_i(1 + |x_i^k|)$ . Рекомендации Гилла, Мюррея и Райт лежат в основе предлагаемой модификации подхода Полака.

Шаг 0. Выбрать  $\beta > 0$ ,  $l > 0$ . Попробуйте взять  $\beta \in [5, 10]$ ,  $l \in [10^{-7}, 10^{-6}]$ .

Шаг 1. Положить  $i = 1$ .

Шаг 2. Положить  $l_i = l$ ,  $y_i^0 = x_i^0$ .

Шаг 3. Вычислить  $x_i^0 = y_i^0 - l_i$ ,  $f_l = F(x^0)$ ,  $x_i^0 = y_i^0 + l_i$ ,  $f_p = F(x^0)$ ,  $l_e = -\frac{f_p - f_l}{2l_i}$ ,  $x_i^0 = y_i^0 + l_i l_e$ ,  $f_e = F(x^0)$ .

Шаг 4. Если  $f_e < F(x^0)$ , то положить  $x_i^0 = y_i^0$  и перейти к шагу 5; иначе положить  $l_i = \frac{1}{2}l_i$  и перейти к шагу 3.

Шаг 5. Положить  $i = i + 1$ . Если  $i > n$ , то перейти к шагу 6; иначе перейти к шагу 2.

Шаг 6. Для всех  $i = 1, \dots, n$  вычислить  $l_i = \frac{l_i}{1 + |x_i^0|}$  и остановиться.

Рекомендации Гилла, Мюррея и Райт лежат в основе предлагаемой процедуры конечно-разностной аппроксимации первых и вторых производных.

Шаг 1. ( $1 \leq i \leq n$ ). Положить  $y_i^k = x_i^k$ , вычислить  $\eta_i^k = l_i(1 + |y_i^k|)$ ,  $x_i^k = y_i^k + \eta_i^k$ ,  $f_p = F(x^k)$ , положить  $z_i = f_p$ , вычислить  $x_i^k = y_i^k - \eta_i^k$ ,  $f_l = F(x^k)$ ,  $h_i^k = \frac{f_p - f_l}{2\eta_i^k}$ ,  $x_i^k = y_i^k + 2\eta_i^k$ ,

$f_l = F(x^k)$ ,  $H_{ii}^k = \frac{f_l + f_k - 2f_p}{(\eta_i^k)^2}$ , положить  $x_i^k = y_i^k$ .

Шаг 2. Положить  $i = 1$ .

Шаг 3. Положить  $y_i^k = x_i^k$ , вычислить  $\eta_i^k = l_i(1 + |y_i^k|)$ ,  $x_i^k = y_i^k + \eta_i^k$ .

Шаг 4. ( $i + 1 \leq j \leq n$ ). Положить  $y_j^k = x_j^k$ , вычислить  $\eta_j^k = l_j(1 + |y_j^k|)$ ,  $x_j^k = y_j^k + \eta_j^k$ ,  $f_p = F(x^k)$ ,  $H_{ij}^k = \frac{f_p + f_k - z_i - z_j}{\eta_i^k \eta_j^k}$ , положить  $x_j^k = y_j^k$ .

Шаг 5. Положить  $x_i^k = y_i^k$ ,  $i = i + 1$ . Если  $i > n$ , то остановиться; иначе перейти к шагу 3.

### Тестовые задачи и сравнительная оценка алгоритмов

Всюду далее:

$\delta x = \max  x_i^* - x_i^k $ — точность решения по аргументам, $\delta F =  F(x^*) - F(x^k) $ — точность решения по функции, $\mu$ — число обусловленности $x^*$ , которое характеризует степень вытянутости линий уровня $F(x)$ в окрестности $x^*$ . (Если $\mu$ велико, то линии уровня сильно вытянуты — функция имеет овражный характер, т.е. резко возрастает по одним направлениям и слабо меняется по другим [Поляк, 1983]) $k$ — число итераций,	$k_0$ — число вычислений функции, $k_1$ — число аппроксимаций $H^k$ при наличии отрицательных собственных значений, $k_2$ — число аппроксимаций $H^k$ при наличии нулевых собственных значений, $k_3$ — число переходов к направлению отрицательной кривизны, $\tau_F$ ( $\tau_F \leq 400$ ) — число «правильных разрядов» $F(x^k)$ , которые хотелось бы получить, $\gamma$ — характеризует степень однородности модельной квадратичной функции.
---	--

**Задача 1** (функция Розенброка,  $n=2$ ):

$$F(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \quad x^0 = (-1.2, 1)^T, \quad x^* = (1, 1)^T, \quad F(x^*) = 0.$$

Функция плохо обусловленная, не выпуклая, с параболическим оврагом, точка  $x^0$  далека от точки  $x^*$ .

**Задача 2** (функция Пауэлла,  $n=4$ ):

$$F(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4, \quad x^0 = (3, -1, 0, 1)^T, \quad x^* = (0, 0, 0, 0)^T, \\ F(x^*) = 0. \text{ Функция не выпуклая, точка минимума — вырожденная.}$$

**Задача 3** (среднеквадратичная аппроксимация экспонентами,  $n=4$ ):

$$F(x) = \sum_{j=1}^{10} [\exp(-0.2j) + 2\exp(-0.4j) - x_1 \exp(-0.2jx_2) - x_3 \exp(-0.2jx_4)]^2, \quad x^0 = (0.5, 0, 2.5, 3)^T, \\ x^* = (1, 1, 2, 2)^T. \text{ Функция не выпуклая, с искривленным оврагом, обусловленность велика.}$$

**Задача 4** (функция Вуда,  $n=4$ ):

$$F(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 + 90(x_4 - x_3^2)^2 + (1 - x_3)^2 + 10.1((x_2 - 1)^2 + (x_4 - 1)^2) + 19.8(x_2 - 1)(x_4 - 1), \\ x^0 = (3, -1, -3, -1)^T, \quad x^* = (1, 1, 1, 1)^T, \quad F(x^*) = 0. \text{ Функция имеет несколько локальных минимумов, это обстоятельство может вызвать преждевременное окончание процесса.}$$

**Задача 5** (функция Степенная,  $n=2$ ):

$$F(x) = (10(x_1 - x_2)^2 + (x_1 - 1)^2)^4, \quad x^0 = (-1.2, 0)^T, \quad x^* = (1, 1)^T, \quad F(x^*) = 0.$$

Таблица 1. Сравнение предлагаемых методов для задачи 1

Метод	$k$	$\delta F$	$\delta x$	$k_0$	$k_1$	$k_2$	$k_3$	$\tau_F$	$\gamma$
«Mnb»	10	$2.5 \cdot 10^{-24}$	$1.5 \cdot 10^{-12}$	24	0	0	7	48	10
«Mnb»	10	$2.5 \cdot 10^{-24}$	$2 \cdot 10^{-14}$	24	0	0	7	96	10
«Mnb»	11	0	0	26	0	0	8	130	10
«MnbApp»	17	$8.8 \cdot 10^{-20}$	$2.9 \cdot 10^{-10}$	244	0	0	14	48	—
«MnbApp»	20	0	0	313	0	0	17	96	—

Таблица 2. Сравнение предлагаемых методов для задачи 2

Метод	$k$	$\delta F$	$\delta x$	$k_0$	$k_1$	$k_2$	$k_3$	$\tau_F$	$\gamma$
«Mnb»	4	0	0	6	0	1	1	48	10
«MnbApp»	4	0	0	108	0	1	1	48	–

Таблица 3. Сравнение предлагаемых методов для задачи 3

Метод	$k$	$\delta F$	$\delta x$	$k_0$	$k_1$	$k_2$	$k_3$	$\tau_F$	$\gamma$
«MnbApp»	36	$1.2 \cdot 10^{-11}$	$4 \cdot 10^{-5}$	1176	4	0	0	48	–
«MnbApp»	41	$1.1 \cdot 10^{-21}$	$5 \cdot 10^{-10}$	1319	4	0	0	96	–
«MnbApp»	41	$4.2 \cdot 10^{-33}$	0	1272	4	0	0	150	–

Таблица 4. Сравнение предлагаемых методов для задачи 4

Метод	$k$	$\delta F$	$\delta x$	$k_0$	$k_1$	$k_2$	$k_3$	$\tau_F$	$\gamma$
«Mnb»	13	$3.9 \cdot 10^{-28}$	$2 \cdot 10^{-15}$	45	3	0	9	96	10
«Mnb»	13	0	0	47	3	0	10	150	10
«Mnb»	14	0	0	47	0	0	10	150	20
«MnbApp»	13	$3.2 \cdot 10^{-22}$	$3 \cdot 10^{-11}$	365	0	0	10	96	–
«MnbApp»	20	$1.7 \cdot 10^{-26}$	$4 \cdot 10^{-14}$	711	0	0	15	150	–

Таблица 5. Сравнение предлагаемых методов для задачи 5

Метод	$k$	$\delta F$	$\delta x$	$k_0$	$k_1$	$k_2$	$k_3$	$\tau_F$	$\gamma$
«Mnb»	12	0	0	163	0	2	9	48	10
«Mnb»	12	0	0	163	0	2	9	150	10
«MnbApp»	24	$6.6 \cdot 10^{-8}$	$3 \cdot 10^{-2}$	201	0	0	0	48	–

В таблицах 1–5 приводятся результаты вычислений с помощью двух версий предложенного ньютоновского метода безусловной минимизации: «Mnb» и «MnbApp». «MnbApp» отличается от своего классического прототипа «Mnb» (ввод градиента и гессиана производится вручную) конечно-разностной аппроксимацией первых и вторых производных. Все версии алгоритмов реализованы на языке Visual Basic .NET.

Таблица 6. Сравнение общеизвестных методов для задач 1–3

Метод	$k$	$k_0$	$\delta x$	$\delta F$
Задача 1				
«Градиент»	–	7657	$2 \cdot 10^{-6}$	$10^{-9}$
«Симплекс»	–	200	–	$10^{-8}$
«Барицентр»	–	674	$10^{-9}$	$10^{-17}$
«Конград-2»	30	–	$10^{-5}$	$10^{-9}$
«ДФП»	20	–	$10^{-5}$	$10^{-9}$
«БФШ»	22	–	$10^{-6}$	$10^{-11}$
«Шор-1»	39	286	$10^{-7}$	$10^{-15}$
Задача 2				
«Симплекс»	–	209	–	$7 \cdot 10^{-8}$
«Барицентр»	–	924	$10^{-4}$	$10^{-18}$
«Конград-2»	27	–	$10^{-2}$	$10^{-4}$
«ДФП»	16	–	$10^{-3}$	$10^{-9}$
«БФШ»	16	–	$10^{-3}$	$10^{-9}$
«Шор-1»	50	695	$10^{-6}$	$10^{-13}$
Задача 3				
«Конград-2»	179	781	$10^{-4}$	$10^{-13}$
«Конград-3»	47	205	$10^{-8}$	$10^{-18}$
«Шор-1»	83	400	$10^{-7}$	$6 \cdot 10^{-5}$

В таблице 6 приведены результаты вычислений с помощью методов, описанных в монографии Поляка [Поляк, 1983], краткое описание которых приводится ниже. Все столбцы разбиты на два: в первом указаны результаты, полученные соответствующим методом, а во втором — предложенным алгоритмом («Mnb» или «MnbApp»).

Метод Шора «Шор-1» основан на эвристическом подборе матрицы преобразования пространства. Преобразования сводятся к последовательным растяжениям в специально подбираемых направлениях. Такие методы называют  $R$ -алгоритмами. Они по структуре близки к квазиньютоновским методам переменной метрики, но основаны не на оценке матрицы вторых производных, а на построении матрицы преобразования, определяющей возврат от некоторых новых координат к исходным координатам.

Метод Дэвидона–Флетчера–Пауэлла «ДФП» и метод Бройдена–Флетчера–Шенно «БФШ» относятся к классу методов переменной метрики, называемых также квазиньютоновскими или градиентными с большим шагом. В этих методах в процессе поиска осуществляется аппроксимация матрицы вторых производных или обратной к ней. Причем для этого используются только первые производные. В методе «ДФП» направление градиента является отклоненным в результате умножения на положительно определенную симметрическую матрицу, обратную матрице вторых производных. На следующем шаге такая матрица представляется в виде суммы предыдущей и двух симметрических матриц ранга один каждая. Метод «БФШ» отличается от метода «ДФП» формулой пересчета положительно определенной симметрической матрицы, обратной матрице вторых производных.

Метод барицентрических координат «Барицентр» относится к классу методов прямого поиска. В их основе лежит анализ определенного набора точек вокруг текущей точки, причем ищется такая точка, в которой значение целевой функции меньше, чем значение в текущей точке. Методы прямого поиска для решения задач оптимизации можно использовать тогда, когда отсутствует какая-либо информация о дифференцируемости целевой функции или для случая прерывистой функции.

Метод Пауэлла «Пауэлл» использует значения функции в определенных точках для аппроксимации функции полиномом второй степени.

Разностный аналог градиентного метода «Градиент» относится к классу методов прямого поиска, стратегия которого заключается в использовании значений функции для конечно-разностной аппроксимации первых производных.

Метод «Конград-2» есть версия метода сопряженных направлений, а «Конград-3» является сочетанием метода сопряженных градиентов с методом переменной метрики.

Метод «Симплекс» есть версия симплексного метода, идея которого состоит в сравнении значений функции в  $n + 1$  вершинах симплекса (правильный многогранник, имеющий  $n + 1$  вершину) и перемещении в направлении оптимальной точки с помощью итерационной процедуры.

## Список литературы

- Гилл Ф., Мюррей У., Райт М. Практическая оптимизация. — М.: Изд. «Мир», 1985.
- Лачинов В. М., Поляков А. О. Информодинамика или путь к Миру открытых систем. — СПб: Изд-во СПбГТУ, 1999.
- Полак Э. Численные методы оптимизации. Единый подход. — М.: Изд. «Мир», 1974.
- Поляк Б. Т. Введение в оптимизацию. — М.: Изд.: «Наука», Главная редакция физико-математической литературы, 1983.
- Хакимов Б. Б., Маилян А. А., Маилян А. А., Хакимова А. Б. Патент на полезную модель № 51252. Система для обработки данных и управления рынка. Зарегистрировано 27 января 2006. Москва.
- Хакимова А. Б., Дикусар В. В., Зеленков Г. А. Увеличение эффективности ньютоновских методов оптимизации. Информационно-динамический подход. Труды ИСА РАН «Динамика неоднородных систем». Выпуск 14, Том 53. М.: Книжный дом «ЛИБРОКОМ», 2010. ISBN 978-5-397-00676-7