УДК 519.853:005.642.4

Г.Г. Асеев

Харьковская академия культуры, г. Харьков, Украина
gaseyev@ic.ac.kharkov.ua

СОВЕРШЕНСТВОВАНИЕ МЕТОДА МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ В ЭЛЕКТРОННОМ ДОКУМЕНТООБОРОТЕ, ОСНОВАННОГО НА ЛОКАЛЬНОЙ ГЕОМЕТРИИ МНОЖЕСТВА ПАРЕТО

Описан метод решения гладких нелинейных задач многокритериальной оптимизации с ограничениями, в электронном документообороте, позволяющий находить как одну Парето-оптимальную точку, так и восстанавливать весь Парето-фронт. Метод основан на расширении алгоритмов одномерного поиска (метод наискорейшего спуска, метод Ньютона) на многокритериальный случай для нахождения одного оптимального решения и идеях метода продолжения (решения системы нелинейных уравнений) для нахождения критической точки.

ЭЛЕКТРОННЫЙ ДОКУМЕНТООБОРОТ, ОПТИМИЗАЦИЯ, ПАРЕТО-МНОЖЕСТВО, ЦЕЛЕВАЯ ФУНКЦИЯ, МНОГОКРИТЕРИАЛЬНАЯ ЗАДАЧА, МЕТОД НАИСКОРЕЙШЕГО СПУСКА, МЕТОД НЬЮТОНА

Введение

Электронный документооборот описывает все бизнес-процессы ERP-систем[1] [1]. В описании модели данных бизнес-процессов стратегии производства и операций возникают различные оптимизационные задачи: экономисты минимизируют издержки и максимизируют прибыль; производственники оптимизируют рабочие процессы для того, чтобы добиться максимального выпуска продукции при минимальном браке, минимальных расходах энергоресурсов и комплектующих и пр. Помимо этого, в задаче часто присутствуют ограничения: процент брака не должен превышать определенной величины, выпускаемая продукция должна обладать заданным запасом прочности и заданным качеством и т.д. Таким образом, возникают задачи многокритериальной оптимизации.

Сформулируем теперь задачу оптимизации в общем виде и введем обозначения:

$$\min_{x\in R^n} \vec{f}\left(\vec{x}\right) \text{ при ограничениях } \vec{g}\left(\vec{x}\right)\le 0, \tag{1}$$

где $\vec{x}=\left(x_1,...,x_n\right)$ — вектор переменных (параметров); $\vec{f}=\left(f_1,...,f_k\right)$ — векторнозначная целевая функция; $\vec{g}=\left(g_1,...,g_m\right)$ — векторнозначная функция ограничений. Точка $\vec{x}$, удовлетворяющая всем ограничениям, называется *допустимой*. Множество всех допустимых точек обозначим $\vec{\Omega}=\left\{\vec{x}\middle|\vec{g}\left(\vec{x}\right)\le 0\right\}$. Если $\vec{f}$ состоит из одной компоненты (скалярная функция), оптимизация называется *однокритериальной*, иначе — *многокритериальной*. Используя введенные выше обозначения, перепишем (1). Для однокритериальной оптимизации

$$f\left(\vec{x}\right)\to min, \vec{x}\in\Omega, \tag{2}$$

для многокритериальной оптимизации

$$\left(f_1\left(\vec{x}\right),...,f_k\left(\vec{x}\right)\right)\to min, \vec{x}\in\Omega. \tag{3}$$

Для многокритериального случая важными являются понятия Парето-доминирования и Парето-оптимальности. Рассмотрим их несколько подробнее.

Парето-оптимальность. Основная идея Парето-оптимальности: мы не можем улучшить наше решение по одному из показателей, не ухудшив при этом по другому.

Пусть $\vec{a}=\left(a_1,...,a_n\right)$, $\vec{b}=\left(b_1,...,b_n\right)\in R^n$.

Определение 1. *$\vec{a}$ доминирует $\vec{b}$ $\left(\vec{a}\succ\vec{b}\right)$, если $a_i\le b_i$ для всех i от 1 до n и $a_i<b_i$ по крайней мере для одного из $i\in\{1,\cdots,n\}$. Иначе, считаем, что $\vec{a}$ не доминирует $\vec{b}$ $\left(\vec{a}\neq\vec{b}\right)$.*

Если $\left(\vec{a}\neq\vec{b}\right)$ и $\left(\vec{b}\neq\vec{a}\right)$, то $\vec{a}$ не сравнимо с $\vec{b}$ $\left(\vec{a}<>\vec{b}\right)$.

В дальнейшем, если говорится об отношениях доминирования между двумя множествами векторов из пространства параметров, то под этим понимаются соотношения для их образов в пространстве целевых функций.

Оптимизационные алгоритмы итеративны. Стартуя с некоторого начального приближения, алгоритм генерирует последовательность точек до тех пор, пока не будет выполнено условие остановки. Таким образом, в процессе оптимизации целевые функции и ограничения вычисляются на конечном множестве точек. При этом, множество Парето-оптимальных решений в общем случае содержит бесконечное количество точек. Соответственно, на практике задача многокритериальной оптимизации переформулируется так, чтобы ее можно было решить с помощью итеративного алгоритма.

[1] **ERP** (англ. Enterprise Resource Planning) — планирование ресурсов предприятия, организационная стратегия интеграции производства и операций, управления трудовыми ресурсами, финансового менеджмента и управления активами, ориентированная на непрерывную балансировку и оптимизацию ресурсов предприятия посредством специализированного интегрированного пакета прикладного программного обеспечения, обеспечивающего общую модель данных и процессов для всех сфер деятельности.

Решением задачи многокритериальной оптимизации является Парето-множество $\boldsymbol{S} \subset \Omega$ всех Парето-оптимальных допустимых точек, т.е. таких, которые не доминируются никакими другими допустимыми точками. Парето-множеству $\boldsymbol{S}$ соответствует Парето-фронт $\boldsymbol{P} = \boldsymbol{f}(\boldsymbol{S})$ (образ Парето-множества в пространстве целевых функций).

Таким образом, в отличие от однокритериальной задачи, ответом является не одна точка, а множество точек. В Парето-множестве точки не сравнимы между собой, т.е. все решения задачи равноценны.

1. Постановка задачи

Описанная здесь краткая методика решения задачи многокритериальной оптимизации не использует веса для целевых функций, замену целевых функций ограничениями и другие способы скаляризации. Предварительно необходимо обосновать расширение алгоритмов одномерного поиска (метод наискорейшего спуска, метод Ньютона) на многокритериальный случай для нахождения одного оптимального решения. Исходя из этого рассмотрения, надо использовать идеи метода продолжения [2] (решения системы нелинейных уравнений) для нахождения критической точки и предложить основы алгоритма, использующего веса для целевых функций.

2. Поиск Парето-оптимального решения

2.1. Однокритериальная задача

Рассмотрим задачу однокритериальной оптимизации. Требуется найти минимум функции $f(x)$. Начиная с заданной точки x_0, будем генерировать последовательность точек $x_{k+1}(x_k)$ таких, что $f(x_{k+1}) < f(x_k)$.

Пусть $x_{k+1} = x_k + \alpha\boldsymbol{v}$, где $\boldsymbol{v}$ — единичный вектор направления, а α — длина шага. Вычислим $\boldsymbol{v}$. Из теоремы Тейлора следует:

$$f\left(x_k+\alpha v\right)=f\left(x_k\right)+\alpha v^T\nabla f\left(x_k\right)+o\left(|v|\right). \qquad (4)$$

То, насколько изменится f вдоль направления v, зависит от произведения $v^T\nabla f\left(x_k\right)$. Получаем следующую задачу:

$$\min_p v^T\nabla f\left(x_k\right) \qquad (5)$$

при условии $|v|=1$.

Поскольку

$$v^T\nabla f\left(x_k\right)=|v|\left|\nabla f\left(x_k\right)\right|\cos\theta=\left|\nabla f\left(x_k\right)\right|\cos\theta,$$

то минимум (5) достигается при $\theta = \varpi$ и $\cos\theta = -1$. Таким образом, получаем:

$$v=-\nabla f_k \big/ \left|\nabla f_k\right| \qquad (6)$$

— *направление наискорейшего спуска*.

Вообще, любое направление, отличающееся от $-\nabla f_k$ меньше, чем на $\varpi/2$ является *направлением спуска*, т.е. приводит к уменьшению f при достаточно малых α. При $\nabla f_k = 0$ такого направления не существует, x_k является минимумом функции f.

Необходимое условие локального экстремума функции f: если x_k — локальный экстремум, то $f(x_k) = 0$. Теперь рассмотрим эти же вопросы для многокритериальных задач.

2.2. Многокритериальная задача с ограничениями

Определение 2. *Ограничение $g_i(x)$ в точке x^* называется активным, если $g_i(x^*) = 0$. Множество всех активных ограничений $\{i \in 1, \ldots, k \mid g_i(x^*) = 0\}$ обозначим* $\mathrm{A}\left(x^*\right)$.

Сформулируем следующую теорему (необходимое условие Парето-оптимальности Каруша—Куна-Такера) [3]:

Теорема 1. *Пусть выполнено условие: множество векторов $\left\{\nabla g_i\left(x^*\right) \mid i\in \mathrm{A}\left(x^*\right)\right\}$ линейно независимо. Если точка x^* является (локально) Парето-оптимальной для задачи (1), то существуют вектора $\lambda \in R^k$ и $\mu \in R^m$ такие, что*:

$$\begin{gathered}\sum_{i=1}^k\lambda_i\nabla f_i\left(x^*\right)+\sum_{j=1}^m\mu_j\nabla g_j\left(x^*\right)=0,\\ \mu_j g_j\left(x^*\right)=0,\quad j=1,\ldots m,\\ \lambda\geq 0,\quad \lambda\neq 0,\quad \mu\geq 0.\end{gathered} \qquad (7)$$

В ряде случаев удобно использовать формулировку необходимого условия для задачи (3), которая зависит только от геометрии множества допустимых значений Ω, а не от конкретного ее алгебраического описания с помощью функции $g(x)$. Для этого нужно ввести понятия касательного вектора и касательного конуса.

Определение 3. *Вектор $v \in R^n$ называется касательным вектором к Ω в т. $x \in \Omega$, если для всех векторных последовательностей $\{x_i\}$: $x_i \to x$, $x_i \in \Omega$ и для всех положительных скалярных последовательностей $t_i \to 0$ существует последовательность $v_i \to v$ такая, что $x_i + t_i v_i \in \Omega$ $\forall$ i.*

Определение 4. Касательный конус $T_\Omega(x)$ — множество всех касательных векторов к Ω в m, x.

Теорема 2. *Если точка x^* является (локально) Парето-оптимальной для задачи (3), то система:*

$$\left(\nabla f_i\left(x^*\right)\right)^T v<0,\quad i=1,\ldots,k,\ \ v\in T_\Omega\left(x\right). \qquad (8)$$

Определение 5. *Точка, в которой выполнены необходимые условия (локальной) Парето-оптимальности (4) или (5) называется критической.*

2.3. Многокритериальная задача без ограничений

Смысл приведенных выше теорем проще всего пояснить на задачах без ограничений.

При $m = 0$ теорема 1 утверждает, что в Парето-оптимальной точке нулевой вектор может быть записан в виде линейной комбинации градиентов компонент целевой функции с положительными коэффициентами.

Для задачи без ограничений теорема 2 переформулируется следующим образом.

$$Im(M)\left(\nabla f\left(x^*\right)\right)\cap\left(-R_{++}\right)^k=0, \qquad (9)$$

где $Im(M)$ — образ линейного отображения M; $\nabla f\left(x^*\right)$ — матрица Якоби функции $f(x^*)$: $\left(\nabla f\left(x^*\right)\right)_{i,j}=\partial f_i \big/ \partial x_j\left(x^*\right)$; R_{++} — множество положительных вещественных чисел.

Таким образом, если точка x* не является Парето-оптимальной, значит, существует вектор $v \in R^n$ такой, что

$$\nabla f\left(x^*\right)v \in \left(-R_{++}\right)^k \quad (10)$$

— направление спуска. Вдоль этого вектора функция $f(x)$ будет уменьшаться по всем своим компонентам (при достаточно малом положительном шаге).

3. Метод наискорейшего спуска

Возникает вопрос: какое из возможных направлений спуска выбрать? Очевидно, мы хотим максимально уменьшить целевую функцию по всем компонентам. Рассмотрим пока задачу без ограничений. Введем функцию $h: R^n \to R^k$ [4]:

$$h_x(v) := \max \left\{ \left(\nabla f\left(x^*\right)v\right)_i \mid i = 1,\ldots,k \right\}. \quad (11)$$

Рассмотрим следующую задачу:

$$\min h_x(v), |v|_\infty \le 1. \quad (12)$$

Вместо $|v|_\infty$ нормы можно использовать и другие, но ограничение в таком виде приводит к наиболее простой формулировке задачи. Переформулировав (12), получаем задачу линейной оптимизации, решив которую, можно найти оптимальное направление спуска:

$$\begin{gathered} \min t, \quad \left(\nabla f_i(x)\right)^T v \le t, \quad i = 1,\ldots,k, \\ |v|_\infty \le 1. \end{gathered} \quad (13)$$

Можно доказать, что итеративный алгоритм $x_{k+1} = x_k + \alpha v$, использующий данное направление спуска, при некоторых условиях, накладываемых на выбор длины шага α, сходится к критической точке [4].

При $k = 1$ (однокритериальная оптимизация) получаем $v = -\nabla f_k / |\nabla f_k|$, т.е. данный алгоритм является расширением метода наискорейшего спуска на многокритериальный случай. В однокритериальном случае алгоритм останавливается, если норма градиента достаточно мала. Но это условие остановки не подходит для многокритериальных задач, т.к. градиенты в них не обращаются в ноль в критической точке. Однако, в качестве критерия можно использовать величину t решения задачи (13), т.к. обращение ее в ноль означает, что алгоритм сошелся в критическую точку.

Теперь рассмотрим задачу с ограничениями. Воспользуемся следующей теоремой [5].

Теорема 3. *Рассмотрим множество*

$$C\left(x^*\right) := \left\{ v \in R_n \,\middle|\, \left(\nabla g_i\left(x^*\right)\right)^T v \le 0, i \in A\left(x^*\right) \right\}.$$

Если вектора $\left\{ \nabla g_i\left(x^*\right) \middle| i \in A\left(x^*\right) \right\}$ *линейно независимы, то*

$$C(x^*) = T_\Omega(x^*).$$

Таким образом, из (8) получаем: для того, чтобы точка x^* была Парето-оптимальной, необходимо, чтобы следующая система не имела решений:

$$\begin{gathered} \left(\nabla f_i(x)\right)^T v < 0, \qquad i = 1,\ldots,k, \\ \left(\nabla g_i\left(x^*\right)\right)^T v \le 0, \quad j \in A\left(x^*\right). \end{gathered} \quad (14)$$

Вместо задачи (13), соответственно, получаем [4]:

$$\min t,$$

при ограничениях

$$\begin{gathered} \left(\nabla f_i(x)\right)^T v \le t, \quad i = 1,\ldots,k, \\ \left(\nabla g_i\left(x^*\right)\right)^T v \le t, \quad j \in A\left(x^*\right), \\ |v|_\infty \le 1. \end{gathered} \quad (15)$$

4. Метод Ньютона

Главный недостаток методов первого порядка (таких, как метод наискорейшего спуска) заключается в их линейной сходимости к оптимальной точке. Обратим также внимание на то, что решение задачи (15) не дает никакой оценки длины шага, т.к. $|v|_\infty$ всегда равно 1. В этом разделе мы рассмотрим метод второго порядка (т.е. использующий вторые производные функций $f(x)$ и $g(x)$), сходящийся сверхлинейно или квадратично.

Начнем с однокритериального случая и теоремы Тейлора:

$$\begin{gathered} f\left(x_k + v\right) = f\left(x_k\right) + v^T \nabla f\left(x_k\right) + \\ + \frac{1}{2} v^T \nabla^2 f\left(x_k\right) v + o\left(|v|^2\right). \end{gathered} \quad (16)$$

Нужно найти такой v, который минимизирует $m(v) = v^T \nabla f\left(x_k\right) + \frac{1}{2} v^T \nabla^2 f\left(x_k\right) v$. Если гессиан $\nabla^2 f_k$ положительно определен, то, приравнивая $m'(v)$ к нулю, получаем:

$$v = -\nabla^2 f_k^{-1} \nabla f_k. \quad (17)$$

В многокритериальном случае для нахождения υ предлагается решать следующую задачу [6]:

$$\min_{v \in R^n} \max_{i=1,\ldots,k} \left(\nabla f\left(x_k\right)\right)^T v + \frac{1}{2} v^T \nabla^2 f_i(x) v. \quad (18)$$

Если все гессианы положительно определены, то задача (18) всегда имеет единственное решение. При $k = 1$ решением является (17).

Таким образом, разность $\max\limits_{i=1,\ldots,k} f_i(x+v) - f_i(x)$ апроксимируется максимумом локальных квадратичных моделей компонент целевой функции в точке x. Задача (18) не является гладкой, но она может быть переформулирована следующим образом:

$$\min t,$$

при условии

$$\begin{gathered} \left(\nabla f_i(x)\right)^T v + \frac{1}{2} v^T \nabla^2 f_i(x) v \le t, \ i = 1,\ldots,k, \\ \left(\nabla g_i\left(x^*\right)\right)^T v \le t, \quad j \in A\left(x^*\right). \end{gathered} \quad (19)$$

В [6] доказывается, что алгоритм, использующий решение (19) в качестве направления спуска, при некоторых условиях сходится к локально Парето-оптимальной точке сверхлинейно, если $f(x)$ дважды непрерывно дифференцируема и квадратично, если вторые производные $f(x)$ непрерывны по Липшицу.

Для оптимизации с ограничениями получаем следующую задачу:

$$\min t,$$

при ограничениях

$$\begin{gathered} \left(\nabla f_i(x)\right)^T v + \frac{1}{2} v^T \nabla^2 f_i(x) v \le t, \quad i = 1,\ldots,k, \\ \left(\nabla g_j(x)\right)^T v + \frac{1}{2} v^T \nabla^2 g_j(x) v \le t, \quad j \in A\left(x^*\right). \end{gathered} \quad (20)$$

5. Выбор длины шага

Рассмотрим теперь вопрос о нахождении длины шага. Начнем с однокритериального случая.

Оптимальное значение длины шага α — это решение следующей одномерной задачи:

$$\min_{\alpha} f\left(x_k + \alpha v\right), \quad \alpha > 0. \qquad (21)$$

Однако, точное решение задачи (21) требует большого количества вычислений целевой функции. На практике применяют другой подход: подбирают такие значения α, чтобы они удовлетворяли некоторым условиям. Вот пример таких условий (правила Вулфа [7]):

$$f\left(x_k + \alpha v\right) \le f\left(x_k\right) + c_1 \alpha \nabla f_k^T v, \qquad (22)$$

$$f\left(x_k + \alpha v\right)^T v \ge c_2 \nabla f_k^T v, \qquad (23)$$

где $0 < c_1 < c_2 < 1$.

Введем обозначение $\varphi(\alpha) = f(x_k + \alpha v)$. Первое условие (22) (также известное как правило Армихо) говорит о том, что $\varphi(\alpha)$ не должна превышать значения некоторой убывающей линейной функции, равной $f(x_k)$ в нуле. Это условие удовлетворяется для всех достаточно малых α. Второе условие (23) означает, что производная $\varphi'(\alpha)$ по крайней мере в c_2 раз больше, чем $\varphi'(0)$.

В многокритериальном случае условие (23) не работает, поскольку значения производных отдельных компонент ничего не говорят об оптимальности текущей точки. Правило Армихо можно расширить на многокритериальный случай: неравенство (22) следует понимать покомпонентно. Процедура поиска α формулируется так: начиная с $\alpha = 1$, пока неравенство (22) не выполнится, полагаем $\alpha := \alpha/2$.

Выводы

На практике якобиан и тем более гессиан задачи зачастую неизвестен. Для нахождения первых производных используются различные схемы численного дифференцирования, а для вторых производных — различные приближения гессиана.

Для того, чтобы задача (22), (23) имела решение, достаточно положительной определенности хотя бы одного из гессианов. Для приближения гессиана активных ограничений используются формула SR-1 (симметричная ранга один), для аппроксимации гессиана целевых функций – формула BFGS (Бройдена–Флетчера–Гольдфарба–Шанно) [5]. Формула BFGS гарантирует положительную определенность гессиана, если первое приближение также было положительно определенным.

Находить ответ задачи (20) можно с помощью любого метода решения задач нелинейной оптимизации, например, SQP. Она может быть достаточно сложной, однако является «внутренней», т.е. при ее решении не производится вычислений целевой функции и ограничений. Предполагается, что в реальных задачах функции являются достаточно «тяжелыми», чтобы временем решения внутренних задач можно было пренебречь.

Область применения (20), (22), (23) может быть расширена с помощью различных методик. Например, специального вида численное дифференцирование с изменяющейся длиной шага позволяет оптимизировать функции с шумом. Используя алгоритмы глобального поиска для нахождения якорей (минимумов отдельных компонент целевых функций безотносительно значения других компонент)

$$f_a^i = f\left(x_a^i\right), \quad x_a^i = \arg\min_{x \in \Omega} f_i\left(x\right) \quad i = 1, \ldots, k,$$

можно оптимизировать мультимодальные функции.

Список литературы: **1.** *Асеев Г.Г.* Электронный документооборот. Учебник / Г.Г. Асеев – К.: Кондор, 2007. – 500 с. **2.** *Eugene L.A.* Numerical continuation methods / L.A. Eugene, G. Kurt — Berlin: Springer-Verlag, 2010. — 324 s. **3.** *Miettinen K.* Nonlinear multiobjective optimization / K. Miettinen — Berlin: Springer-Verlag, 2009. — 286 s. **4.** *Fliege J.* Steepest descent methods for multicriteria optimization / J. Fliege, B. F. Svaiter // Mathematical Methods of Operations Research. – 2008. № 51(3). Pp. 479-494. **5.** *Nocedal J.* Numerical optimization / J. Nocedal, S. J. Wright — Berlin: Springer Science + Business Media, 2006. — 336 p. **6.** *Fliege J.* Newton's method for multiobjective optimization / J. Fliege, L. M. G. Drummond, B. F. Svaiter // SIAM Journal on Optimization. – 2009. № 20(2). Pp. 602–626. **7.** *Wolfe P.* Convergence conditions for ascent methods / P. Wolfe // Siam Review. – 2010. № 11(2). Pp. 226-235.

Поступила в редколлегию 21.04.2015

УДК 519.853:005.642.4

Удосконалювання методу багатокрітеріальної оптимізації в електронному документообігу, заснованого на локальній геометрії безлічі Парето / Г. Г. Асєєв // Біоніка інтелекту: наук.-техн. журнал. – 2015. – № 2 (85). – С. 58–61.

Робота розглядає бізнес-процеси систем управління виробництвом у вигляді загальної моделі даних і процесів для електронного документообігу. Розглянута ситуація оптимізації моделі даних по декількох критеріях одночасно. Запропонований метод, заснований на розширенні алгоритмів одномірного пошуку на багатокрітериальний випадок. Розвинута ідея методу продовження рішення системи нелінійних рівнянь, що відшукує критичну точку моделі даних для системи електронного документообігу організації.

Бібліогр.: 7 найм.

UDC 519.853:005.642.4

Method improvement Balatonlelle optimization in electronic document management, based on the local geometry of the Pareto set / G. G. Aseyev // Bionics of Intelligense: Sci. Mag. – 2015. – № 2 (85). – P. 58–61.

The work considers the business-processes of production management systems in the form of a common data model and processes for electronic filing. On the optimization of data models according to several criteria simultaneously. The proposed method based on the extension of linear search algorithms on multicriteria case. Developed the idea of the method continue solving the system of nonlinear equations that finds a critical point of the data model for an electronic document management system of the organization.

Ref.: 7 items.