Методы оптимизации

Лекция 5

14 апреля 2023 г.

МЕТОДЫ НЕЛИНЕЙНОЙ УСЛОВНОЙ МИНИМИЗАЦИИ

Условная оптимизация

Многие инженерные задачи связаны с оптимизацией при наличии некоторого количества ограничений на независимые переменные — с условной оптимизацией.

Такие ограничения существенно уменьшают размеры области, в которой проводится поиск оптимума. На первый взгляд, уменьшение размеров допустимой области может облегчить поиск экстремума. Между тем, напротив, процесс оптимизации становится более сложным, прежде всего потому, что описанные выше методы решения задач безусловной минимизации нельзя напрямую использовать при наличии ограничений.

При этом может нарушаться даже основное условие, в соответствии с которым оптимум должен достигаться в стационарной точке, характеризующейся нулевым градиентом. Например, минимум функции $f(x) = x^2$ с учётом ограничения $x \geqslant 2$ находится в точке $x = 2$, которая не является стационарной точкой функции $f(x)$.

Преобразование задач оптимизации

Пусть задача условной оптимизации содержит ограничения специального вида. В зависимости от вида ограничений можно преобразовывать исходную задачу.

1. *Исключение простых ограничений.* Пусть множество допустимых значений образовано простыми (интервальными) ограничениями:

$$X = \left\{x \mid x \in \mathbb{R}^n,\, x_i \geqslant 0,\, i = 1, \ldots, n\right\}.$$

В этом случае заменой переменных $x_i = u_i^2$ исходная условная задача сводится к задаче безусловной минимизации $\min\limits_{u \in \mathbb{R}^n} F(u)$, где минимизируемая функция $F(u)$ получена в результате замены переменных.

Полученная задача безусловной минимизации может оказаться сложнее исходной задачи условной минимизации вследствие большей нелинейности функции $F(u)$ по сравнению с функцией $f(x)$ из-за нелинейности преобразования.

2. *Исключение неравенств.* Для образованного неравенствами множества допустимых значений

$$X = \{x \mid x \in \mathbb{R}^n,\ g_i(x) \geqslant 0,\ i = 1, \ldots, m\}.$$

введением вспомогательных переменных неравенства превращаются в равенства:

$$g_i(x) \geqslant 0 \text{ на } g_i(x) - x_j^2 = 0,\ j = n + i,$$

где x_j, $j = n + 1, \ldots, n + m$ — вспомогательные переменные.

Цена такого преобразования состоит в повышении размерности задачи.

3. *Исключение ограничений вида равенств.* Пусть задача оптимизации содержит k ограничений в виде равенств. Эта задача в принципе может быть решена как задача безусловной оптимизации, полученная путём исключения из целевой функции k независимых переменных с помощью заданных равенств. Наличие ограничений в виде равенств фактически позволяет уменьшить размерность исходной задачи с n до $n - k$.

В качестве иллюстрации рассмотрим следующий пример. Найти $\min\limits_{x \in \mathbb{R}^n} f(x)$ при ограничении $x_n - g(x_1, \ldots, x_{n-1}) = 0$. Выразив переменную x_n через остальные, $x_n = g(x_1, \ldots, x_{n-1})$, и исключив её из функции, получим безусловную задачу оптимизации размерностью $n - 1$: $\min\limits_{x \in \mathbb{R}^{n-1}} f(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1}))$.

Метод исключения переменных применим лишь в тех случаях, когда уравнения, представляющие ограничения, можно решить относительно некоторого конкретного набора независимых переменных.

4. *Тригонометрические преобразования.* Пусть задача оптимизации содержит ограничение в виде гиперсферы:

$$X = \left\{ x \;\middle|\; x \in \mathbb{R}^n,\, \sum_{i=1}^{n} x_i^2 = 1 \right\}.$$

Преобразованием 1

$$\begin{aligned} x_1 &= \sin(y_1) \cdots \sin(y_{n-1}), \\ x_i &= \cos(y_{i-1}) \sin(y_i) \cdots \sin(y_{n-1}),\ i = 2, \ldots, n-2, \\ x_n &= \cos(y_{n-1}) \end{aligned}$$

исходная задача минимизации сводится к задаче безусловной минимизации $\min\limits_{y \in \mathbb{R}^{n-1}} F(y)$.

Недостатками такого преобразования являются периодичность новой функции $F(y)$ и её независимость от y_i, при $y_i \approx 0$.

Преобразованием 2

$$
\begin{aligned}
a &= \pm\left(1 + \sum_{i=1}^{n-1} y_i^2\right)^{1/2}, \\
x_i &= y_i/a, \quad i = 1, \ldots, n-1, \\
x_n &= 1/a
\end{aligned}
$$

исходная задача минимизации сводится к задаче безусловной минимизации $\min\limits_{y \in \mathbb{R}^{n-1}} \min\bigl\{F_p(y), F_h(y)\bigr\}$ где в формуле для вычисления a знак «$+$» для функции $F_p(y)$ и знак «$-$» для функции $F_h(y)$.

МЕТОДЫ ПОСЛЕДОВАТЕЛЬНОЙ БЕЗУСЛОВНОЙ ОПТИМИЗАЦИИ

Метод штрафных функций

Основная идея метода состоит в том, чтобы аппроксимировать исходную задачу условной оптимизации некоторой вспомогательной задачей, решение которой менее сложно, чем решение исходной.

Однако, при этом приходится решать последовательность таких задач, сходящихся к исходной. Причём, результаты решения предыдущей задачи используются в качестве начальных приближений при решении последующей задачи. Получение решений с практически необходимой точностью может быть достигнуто за конечное число шагов.

Для обоснования этого подхода рассмотрим следующую задачу с единственным ограничением $h(x) = 0$:

$$\min f(x)$$
$$h(x) = 0$$

Пусть она преобразована в задачу безусловной оптимизации:

$$\min f(x) + \mu h^2(x)$$
$$x \in \mathbb{R}^n$$

где $\mu > 0$ — некоторое большое число.

Интуитивно ясно, что на оптимальном решении x^* последней задачи значение $h^2(x^*)$ должно быть близким к нулю, так как в противном случае всегда можно сдвинуться в другую точку x', в которой приращение $f(x)$ окажется при достаточно большом μ меньше, чем $\mu h^2(x)$.

Рассмотрим теперь задачу с единственным ограничением-неравенством $g(x) \leqslant 0$:

$$\begin{aligned} &\min f(x) \\ &g(x) \leqslant 0 \end{aligned}$$

Очевидно, что форма $f(x) + \mu g^2(x)$ здесь нецелесообразна, так как при $g(x) \neq 0$ штраф будет взиматься независимо от знака $g(x)$.

Штраф желателен только в недопустимых точках, т. е. там, где $g(x) > 0$. Поэтому приемлемой эквивалентной задачей безусловной оптимизации является следующая:

$$\begin{aligned} &\min f(x) + \mu \max\{0, g(x)\} \\ &x \in \mathbb{R}^n \end{aligned}$$

Если $g(x) \leqslant 0$, то $\max\{0, g(x)\} = 0$ и штраф не берётся. Если же $g(x) > 0$, то $\max\{0, g(x)\} > 0$, т. е. взимается штраф $\mu g(x)$.

Обычно подходящая штрафная функция должна определять положительный штраф в недопустимых точках и не штрафовать допустимые точки. Если ограничения имеют форму $g_i(x) \leqslant 0$, $i = 1, \ldots, m$, и $h_i(x) = 0$, $i = 1, \ldots, l$, то целесообразна штрафная функция следующего вида:

$$\alpha(x) = \sum_{i=1}^{m} \varphi\big[g_i(x)\big] + \sum_{i=1}^{l} \psi\big[h_i(x)\big]$$

где φ и ψ — непрерывные функции, удовлетворяющие условиям:

$$\varphi(y) = 0, \text{ если } y \leqslant 0, \quad \text{и} \quad \varphi(y) > 0, \text{ если } y > 0,$$
$$\psi(y) = 0, \text{ если } y = 0, \quad \text{и} \quad \psi(y) > 0, \text{ если } y \neq 0.$$

Типичными являются следующие формы функций $\varphi(y)$ и $\psi(y)$:

$$\varphi(y) = \left[\max\{0, y\}\right]^p,$$
$$\psi(y) = |y|^p,$$

где p — целое положительное число. Таким образом, штрафная функция $\alpha = \alpha(x)$ обычно имеет вид

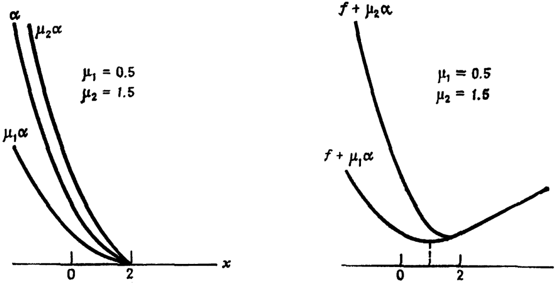
$$\alpha(x) = \sum_{i=1}^{m} \left[\max\{0, g_i(x)\}\right]^p + \sum_{i=1}^{l} |h_i(x)|^p$$

Функцию $f(x) + \mu\alpha(x)$ будем называть вспомогательной. Соответственно задачу со штрафом будем называть вспомогательной задачей.

Рассмотрим задачу

$$\min x$$

$$-x + 2 \leqslant 0$$



При $\mu \to \infty$ последовательность точек минимума функции $f(x) + \mu\alpha(x)$ стремится к $x = 2$, являющейся точкой минимума целевой функции исходной задачи.

Постановка задачи.

Требуется найти минимум функции многих переменных $y = f(x)$, то есть, такую точку $x^* \in U$, что

$$f(x^*) = \min_{x \in U} f(x),$$

где множество точек U определяется ограничениями вида

$$g_j(x) = 0, \quad j = 1, \ldots, m, \ m < n,$$
$$g_j(x) \leqslant 0, \quad j = m+1, \ldots, p.$$

Алгоритм.

1. Задать начальную точку x^0 вне области допустимых решений, начальное значение параметра штрафа $r^0 > 0$, число $C > 1$ для увеличения параметра штрафа, погрешность расчёта $\varepsilon > 0$. Принять $k := 0$.

2. Составить вспомогательную функцию

$$F(x, r^k) = f(x) + \frac{r^k}{2} \left\{ \sum_{j=1}^{m} [g_j(x)]^2 + \sum_{j=m+1}^{p} [g_j^+(x)]^2 \right\},$$

где функция штрафа $P(x, r^k) = \dfrac{r^k}{2}\left\{\sum\limits_{j=1}^{m} [g_j(x)]^2 + \sum\limits_{j=m+1}^{p} [g_j^+(x)]^2\right\}$,

$g_j^+(x)$ — срезка функции:

$$g_j^+(x) = \max\{0, g_j(x)\} = \begin{cases} g_j(x), & g_j(x) > 0, \\ 0, & g_j(x) \leqslant 0. \end{cases}$$

3. Найти точку $x^*(r^k)$ безусловного минимума функции $F(x, r^k)$ по x с помощью какого либо метода (нулевого, первого или второго порядка):

$F(x^*(r^k), r^k) = \min\limits_{x \in \mathbb{R}^n} F(x, r^k)$. При этом задать все требуемые выбранным методом параметры. В качестве начальной точки взять x^k. Вычислить функцию штрафа

$$P(x^*(r^k), r^k) = \frac{r^k}{2}\left\{\sum_{j=1}^{m} [g_j(x^*)]^2 + \sum_{j=m+1}^{p} [g_j^+(x^*)]^2\right\}.$$

4. Проверить выполнение условия окончания:

- если $P(x^*(r^k), r^k) \leqslant \varepsilon$, процесс поиска закончить: $x^* := x^*(r^k)$, $f(x^*) := f(x^*(r^k))$;
- если $P(x^*(r^k), r^k) > \varepsilon$, то принять $r^{k+1} := Cr^k$, $x^{k+1} := x^*(r^k)$, $k := k + 1$ и перейти к п. 2.

Примечание. Рекомендуемые значения параметров: $r^0 = 0.01, 0.1, 1$, $C = 4–10$.

Этот метод называется также методом внутренних штрафов, так как, в отличие от метода штрафных функций (метода внешних штрафов), функция штрафа не равна 0 даже для допустимых точек и возрастает при приближении к границе изнутри.

Барьерные функции как бы препятствуют выходу из допустимой области Если оптимальное решение оказывается на границе допустимой области, то процедура приводит к движению изнутри области к границе.

Ниже сформулируем исходную и вспомогательную (барьерную) задачи.

Исходная задача.

$$\begin{gathered} \min f(x) \\ g_i(x) \leqslant 0, \quad i = 1, \dots, m \end{gathered}$$

где f, $g_1, \dots, g_m$ — непрерывные функции, определённые в $\mathbb{R}^n$.

Вспомогательная (барьерная) задача.

$$\min \theta(\mu)$$

$$\mu \geqslant 0$$

где $\theta(\mu) = \inf\big\{ f(x) + \mu B(x) \;\big|\; g(x) < 0, x \in X \big\}$. Здесь B — *барьерная функция*, неотрицательная и непрерывная в области $\big\{x \;\big|\; g(x) < 0\big\}$ и стремящаяся к бесконечности при приближении изнутри к границе области $\big\{x \;\big|\; g(x) \leqslant 0\big\}$.

Более точно барьерная функция определяется следующим образом:

$$B(x) = \sum_{i=1}^{m} \varphi[g_i(x)],$$

где $\varphi(y)$ — функция одной переменной, непрерывная на множестве $\big\{y \;\big|\; y < 0\big\}$ и удовлетворяющая условиям

$$\varphi(y) \geqslant 0, \text{ если } y < 0 \quad \text{и} \quad \lim_{y \to 0^-} \varphi(y) = \infty.$$

Таким образом, типичная барьерная функция имеет вид

$$B(x) = \sum_{i=1}^{m} \frac{-1}{g_i(x)}$$

Функцию $f(x) + \mu B(x)$ называют вспомогательной функцией.
В идеальном случае желательна функция, которая обращается в нуль в области $\{x \mid g(x) < 0\}$ и равна бесконечности на границе этой области. Это гарантировало бы, что нет выхода за пределы множества $\{x \mid g(x) \leqslant 0\}$ при условии, что задача минимизации начинает решаться из внутренней точки.

Однако, в связи с вычислительными трудностями, идеальная конструкция функции B заменяется более реальным требованием, чтобы B была неотрицательна и непрерывна в области $\{x \mid g(x) < 0\}$ и стремилась к бесконечности при приближении из внутренней точки к границе области.

При $\mu \to 0$ функция μB стремится к идеальной барьерной функции.

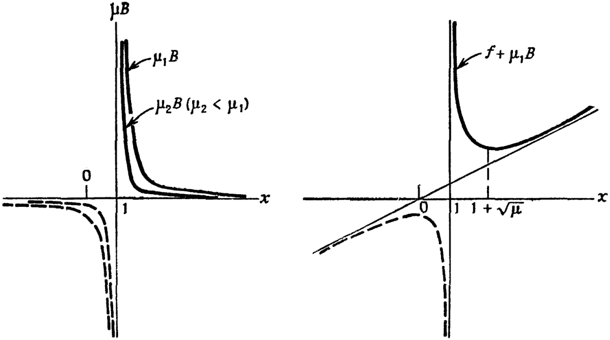
Рассмотрим задачу

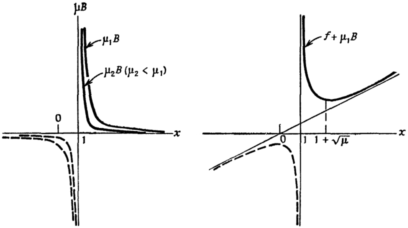
$$\min x$$

$$-x + 1 \leqslant 0$$

Рассмотрим следующую барьерную функцию:

$$B(x) = \frac{-1}{-x+1} \quad \text{для} \quad x \neq 1.$$





При приближении μ к нулю μB приближается к функции, которая при $x > 1$ равна нулю, а при $x = 1$ — бесконечности. Пунктирная кривая не имеет отношения к вычислительному процессу.

Для любого заданного $\mu > 0$ вспомогательная (барьерная) задача заключается в минимизации $x + \mu/(x-1)$ в области $x > 1$.

Функция $x + \mu/(x-1)$ выпукла при $x > 1$. Любой метод, применённый для минимизации $x + \mu/(x-1)$, при начальной внутренней точке $x > 1$ приведёт к оптимальному решению $x_\mu = 1 + \sqrt{\mu}$.

Постановка задачи.

Требуется найти минимум функции многих переменных $y = f(x)$, то есть, такую точку $x^* \in U$, что

$$f(x^*) = \min_{x \in U} f(x),$$

где множество точек U определяется ограничениями вида

$$\begin{aligned} g_j(x) &= 0, \quad j = 1, \ldots, m,\ m < n, \\ g_j(x) &\leqslant 0, \quad j = m+1, \ldots, p. \end{aligned}$$

Метод барьерных функций

Алгоритм.

1. Задать начальную точку x^0 внутри области U, начальное значение параметра штрафа $r^0 > 0$, число $C > 1$ для уменьшения величины параметра штрафа, погрешность расчёта $\varepsilon > 0$. Принять $k := 0$.

2. Составить вспомогательную функцию
$F(x, r^k) = f(x) - r^k \sum\limits_{j=1}^{m} \frac{1}{g_j(x)}$ или $F(x, r^k) = f(x) - r^k \sum\limits_{j=1}^{m} \ln(-g_j(x))$.

3. Найти точку $x^*(r^k)$ безусловного минимума функции $F(x, r^k)$ по x с помощью какого либо метода (нулевого, первого или второго порядка):

$F(x^*(r^k), r^k) = \min\limits_{x \in \mathbb{R}^n} F(x, r^k)$ с проверкой принадлежности текущей точки внутренности множества U. При этом задать все требуемые выбранным методом параметры. В качестве начальной точки взять x^k.

Вычислить функцию штрафа $P(x^*(r^k), r^k) := -r^k \sum\limits_{j=1}^{m} \frac{1}{g_j(x^*(r^k))}$ (обратная функция штрафа) или $P(x^*(r^k), r^k) := -r^k \sum\limits_{j=1}^{m} \ln[-g_j(x^*(r^k))]$ (логарифмическая функция штрафа).

4. Проверить выполнение условия окончания:

- если $\left|P(x^*(r^k), r^k)\right| \leqslant \varepsilon$, то процесс поиска закончить, приняв $x^* := x^*(r^k)$, $f(x^*) := f(x^*(r^k))$;
- если $\left|P(x^*(r^k), r^k)\right| > \varepsilon$, то принять $r^{k+1} := \dfrac{r^k}{C}$, $x^{k+1} := x^*(r^k)$, $k := k+1$ и перейти к п. 2.

Примечание. Рекомендуемые значения параметров: $r^0 = 1, 10, 100$, $C = 10, 12, 16$.

МЕТОДЫ ВОЗМОЖНЫХ НАПРАВЛЕНИЙ

Рассмотрим задачу поиска

$$\min_{x \in U} f(x)$$

где $f(x) \in C^1$, $U \subset \mathbb{R}^n$.

Применение градиентного метода в случае $U \neq \mathbb{R}^n$ может привести к затруднениям, так как точка x^{k+1} может не принадлежать U.

Это осложнение можно преодолеть, если точку x^{k+1} при каждом k проектировать на множество U. В результате придём к методу проекции градиента.

Пусть $x^0 \in U$ — некоторая начальная точка. Будем строить последовательность $\{x^k\}$ по правилу

$$x^{k+1} = P_U\big(x^k - \alpha_k \nabla f(x^k)\big), \quad \alpha_k > 0, \quad k = 0, 1, 2, \dots$$

где $P_U(z)$ — проекция точки z на множество U.

Определим понятие проекции точки на множество.

Определение.

Проекцией точки $z \subset \mathbb{R}^n$ на множество $U \subset \mathbb{R}^n$ называется ближайшая к z точка y (обозначаемая $P_U(z)$) множества U, т. е. точка $y = P_U(z) \in U$, удовлетворяющая условию

$$\|z - P_U(z)\| = \min_{x \in U} \|z - x\|$$

Если $z \in U$, то $z = P_U(z)$.

Для существования и единственности проекции требуется, чтобы множество U было компактом (выпуклым и замкнутым множеством).

Если U — компакт и определён способ выбора α_k, то последовательность $\{x^k\}$ будет определена однозначно.

При $U \equiv \mathbb{R}^n$ метод проекции градиента превращается в градиентный метод.

Если на некоторой итерации оказалось, что $x^{k+1} = x^k$, то процесс поиска минимума прекращают. В этом случае точка x^k удовлетворяет необходимому условию оптимальности для гладких функций, и если минимизируемая функция — выпуклая, то выполнено также достаточное условие оптимальности и точка x^k является точкой минимума.

В зависимости от способа выбора α_k можно получить различные варианты метода проекции градиента. Приведём несколько наиболее распространённых вариантов метода проекции градиента.

1. Выбор α_k из условия

$$f^k(\alpha_k) = \min_{\alpha \geqslant 0} f^k(\alpha) = \min_{\alpha \geqslant 0} f\left(P_U\left(x^k - \alpha \nabla f(x_k)\right)\right)$$

При $U \equiv \mathbb{R}^n$ данный вариант метода проекции градиента превращается в метод наискорейшего спуска.

2. Выбор α_k, обеспечивающего выполнение условия монотонности $f(x^{k+1}) < f(x^k)$.

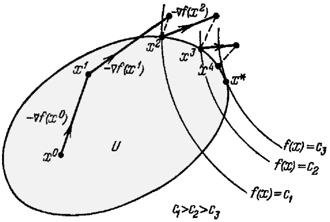
Обычно выбирают $\alpha_k \coloneqq \alpha$ и при нарушении условия монотонности уменьшают α_k до его выполнения.

3. Выбор α_k, обеспечивающего выполнение неравенства

$$f(x_k) - f\left(P_U\left(x_k - \alpha_k \nabla f(x_k)\right)\right) \geqslant \varepsilon \left\| x^k - P_X(x^k - \alpha_k \nabla f(x_k)) \right\|^2, \quad \varepsilon > 0$$

Для определения α_k берут $\alpha_k \coloneqq \alpha$ и затем дробят его, например, по закону $\alpha_k \coloneqq \lambda^i \alpha$, $i = 0, 1, \ldots$, $\lambda \in (0, 1)$, до тех пор, пока не выполнится неравенство.

Метод проекции градиента



Задача определения проекции некоторой точки x^k на множество U является задачей минимизации функции $g(x)$:

$$\min_{x\in U} g(x) = \min_{x\in U}\left\|x - x^k\right\|^2 \text{ или } \min_{x\in U} g(x) = \min_{x\in U}\left\|x - x^k\right\|.$$

Если множество U задаётся с помощью более или менее сложной системы равенств и неравенств, то метод проекции градиента практически неприменим, поскольку тогда задача определения проекции оказывается ничуть не проще исходной.

Методом проекции градиента обычно пользуются в тех случаях, когда проекция точки на множество легко определяется.

Случаи, когда задача проектирования легко решается:

а) Пусть U — шар радиуса R с центром x^0:

$$U = \left\{ x \,\middle|\, x \in \mathbb{R}^n,\ \left\| x - x^0 \right\| \leqslant R \right\}.$$

Проекция точки x на множество U определяется по формуле:

$$P_U(x) = x_0 + R(x - x^0)/\left\| x - x^0 \right\|.$$

б) Пусть U — полупространство:

$$U = \left\{ x \,\middle|\, x \in \mathbb{R}^n,\ (a, x) \leqslant b \right\},$$

Проекция точки $x \notin U$ на множество U определяется по формуле:

$$P_U(x) = x + (b - (a, x))a/\|a\|^2.$$

в) Пусть U — аффинное множество:

$$U = \left\{x \,\middle|\, x \in \mathbb{R}^n,\, (a_i, x) \leqslant b_i,\, i = 1, \dots, m\right\}, \quad m < n.$$

Проекция точки $x \notin U$ на множество U определяется по формуле:

$$P_U(x) = x - A^\mathsf{T}(AA^\mathsf{T})^{-1}(Ax - b),$$

где A — матрица, состоящая из строк a_i, $i = 1, \dots, m$; b — вектор, состоящий из b_i, $i = 1, \dots, m$.

г) Пусть U — n-мерный параллелепипед:

$$U = \left\{x \,\middle|\, x \in \mathbb{R}^n,\, a_i \leqslant x_i \leqslant b_i,\, i = 1, \dots, n\right\}.$$

Проекция точки $x \notin U$ на множество U определяется как:

$$P_U(x) = \left(P_U^1(x), \dots, P_U^n(x)\right),$$

$$P_U^i = \begin{cases} a_i, & \text{при } x_i < a_i, \\ b_i, & \text{при } x_i > b_i, \\ x_i, & \text{при } a_i \leqslant x_i \leqslant b_i. \end{cases}$$

Вновь рассмотрим задачу

$$\min_{x \in U} f(x)$$

где $f(x) \in C^1$, $U \subset \mathbb{R}^n$ — выпуклое замкнутое ограниченное множество.

Пусть $x^0 \in U$ — некоторое начальное приближение. Если известно k-е приближение, $x^k \in U$, то приращение функции $f(x)$ в точке x^k можно представить в виде

$$f(x) - f(x^k) = (\nabla f(x^k), x - x^k) + o\left(\left\|x - x^k\right\|\right).$$

Рассмотрим главную линейную часть этого приращения

$$f_k(x) = (\nabla f(x^k), x - x^k)$$

и определим вспомогательное приближение $\bar{x}^k$ из условия:

$$f_k(\bar{x}^k) = \min_{x \in U} f_k(x) = \min_{x \in U} (\nabla f(x^k), x - x^k).$$

Так как U замкнуто и ограничено, а линейная функция $f_k(x)$ непрерывна, то точка $\bar{x}_k$ всегда существует.

Если функция $f_k(x)$ достигает своей нижней грани на U более чем в одной точке, то в качестве точки $\bar{x}_k$ возьмём любую из них.

Опишем ситуации, когда решение задачи записывается в явном виде.

а) Пусть U — n-мерный параллелепипед:

$$U = \left\{x \;\middle|\; x \in \mathbb{R}^n,\, a_i \leqslant x_i \leqslant b_i,\, i = 1, \ldots, n\right\}.$$

Функция $f_k(x) = \sum\limits_{i=1}^{n} f_{x_i}(x^k)(x_i - x_i^k)$ минимальна на U в точке $\bar{x}^k = (\bar{x}_1^k, \ldots, \bar{x}_n^k)$, где

$$\bar{x}_i^k = \begin{cases} a_i & \text{при } f_{x_i}(x^k) > 0, \\ b_i & \text{при } f_{x_i}(x^k) < 0. \end{cases}$$

При $f_{x_i}(x^k) = 0$ возникает неопределённость, в качестве $\bar{x}_i^k$ можно взять любое число из отрезка $[a_i, b_i]$.

б) Пусть U — шар радиуса R с центром в точке x^0:

$$U = \left\{ x \;\middle|\; x \in \mathbb{R}^n,\ \|x - x^0\| \leqslant R \right\}.$$

Точка минимума $\bar{x}^k$ функции $f_k(x)$ определяется по формуле:

$$\bar{x}^k = x^0 - R\frac{\nabla f(x^k)}{\|\nabla f(x^k)\|}.$$

Получить вспомогательное приближение $\bar{x}^k$ в явном виде удаётся не всегда, вместо точного решения задачи приходится определять приближенное решение.

В этом случае будем предполагать, что $\bar{x}^k$ определяется из следующих условий:

$$f_k(\bar{x}^k) \leqslant \min_{x \in U} f_k(x) + \varepsilon_k,$$

$$\bar{x}^k \in U, \quad \varepsilon_k \geqslant 0, \quad \lim_{k \to \infty} \varepsilon_k = 0.$$

Следующее приближение будем искать в виде:

$$x^{k+1} = x^k + \alpha_k(\bar{x}^k - x^k), \quad \alpha_k \in [0, 1].$$

В зависимости от способа выбора величины α_k получаем различные *варианты метода условного градиента*.

Варианты метода условного градиента

1. Выбор $\alpha_k \in [0, 1]$ из решения задачи минимизации:

$$f_k(\alpha_k) = \min_{\alpha \in [0,1]} f_k(\alpha), \text{ где } f_k(\alpha) = f(x^k + \alpha(\bar{x}^k - x^k)).$$

Точное определение α_k в этом случае не всегда возможно. Величину α_k можно определять из условий:

$$f_k(\alpha_k) \leqslant \min_{\alpha \in [0,1]} f_k(\alpha) + \delta_k,$$

$$\delta_k \geqslant 0, \quad \sum_{k=1}^{\infty} \delta_k = \delta < \infty.$$

2. Выбор α_k из условия:

$$\alpha_k = \begin{cases} 0, & \text{при } f_k(\bar{x}^k) > 0, \\ \lambda^{i_0}, & \text{при } f_k(\bar{x}^k) \leqslant 0, \end{cases}$$

где i_0 — минимальное натуральное число, при котором выполняется неравенство:

$$f(x^k) - f(x^k - \lambda^i(\bar{x}^k - x^k)) \geqslant \varepsilon\lambda^i \left|f_k(\bar{x}^k)\right|, \quad \lambda > 0, \quad \varepsilon < 1.$$

3. Выбор α_k из условия монотонного уменьшения $\{f(x_k)\}$ с дроблением α при нарушении монотонности.

Вопросы?