

## Билет №2.16

*Постановка и численное решение задач оптимизации. Целевая функция, ограничения. Спуск по координатам, наискорейший спуск, случайный поиск. Условная оптимизация.*

### 1 Постановка задачи

Оптимизация занимается поиском наилучших решений. Критерий качества решения формулируется в виде функций многих переменных. Это *целевая функция*. Переменные – *параметры оптимизации*. Обычно имеются некоторые ограничения на эти переменные. Ограничения бывают двух типов: *ограничения равенства* (с жёстким допуском) и *ограничения-неравенства* (формулируются в виде диапазона).

Задача оптимизации детерминированных систем:

$$\begin{cases} f(x) \rightarrow \min \\ c(x) = 0 \\ \bar{c}(x) \geq 0 \end{cases} \quad (1)$$

### 2 Безусловная оптимизация

#### 2.1 Покоординатный спуск

Алгоритм состоит в нахождении локальных минимумов последовательно в каждом направлении. Возьмём некоторую переменную  $x_1$ , а остальные зафиксируем. Найдём минимум  $f$  по этой переменной.  $x_1^{\min}$  – 1-я координата минимума. Заменяем  $x_1$  на  $x_1^{\min}$ . Прервём к  $x_2$ . И т.д.

#### 2.2 Наискорейший спуск

Этот метод предполагает, что функция  $f$  непрерывна и имеет непрерывные первые производные. Задача нахождения минимума сводится к отысканию стационарной точки  $x^*$  ( $\nabla f(x^*) = 0$ ).

Алгоритм наискорейшего спуска состоит в следующем:

1. Выбрать начальную точку  $x^0$ .
2. На  $k$ -й итерации  $d_k = -\nabla f(x^k)$ ; найти такое  $\lambda_k \geq 0$ , что

$$\lambda_k = \operatorname{argmin} g(\lambda), \text{ где } g(\lambda) = f(x^k + \lambda d_k) \text{ (одномерная минимизация).}$$

Положить  $x^{k+1} = x^k + \lambda_k d_k$ .

3. Провести тест на остановку. Если он выполнен, то закончить процедуру; если нет, то перейти от  $k$  к  $k + 1$  и перейти к шагу 2.

В качестве теста на остановку можно взять, например, такой:

$$\|\nabla f\|^2 \leq \epsilon \text{ (при заданном заранее } \epsilon \text{).}$$

#### 2.3 Случайный поиск??

Т.к. что писать сюда непонятно, пусть будет алгоритм имитации отжига, который является вполне себе случайным.

Алгоритм имитации отжига воспроизводит процесс кристаллизации, который происходит, например, при отжиге металлов (отсюда такое название метода).

Решения задачи воспринимаются как состояния некоторой физической системы. Каждому состоянию  $s$  сопоставляется энергия  $E(s)$ . Алгоритм в ходе работы ищет состояние системы с минимальной энергией. Он начинает с некоторого стартового состояния системы  $s_0$ . На каждом шаге выбирается новое состояние для перехода. Это делается случайным образом, причём обычно так, что вероятность состояния быть выбранным зависит от расстояния до текущего состояния. После того как для перехода из состояния  $s_i$  было выбрано некоторое состояние  $s_{i+1}$  рассчитывается величина  $\Delta E = E(s_{i+1}) - E(s_i)$ . Если она оказывается отрицательна, то система переходит в состояние  $s_{i+1}$ . Если же она положительна, то переход в  $s_{i+1}$  делается с вероятностью  $p = p(\Delta E, T)$ , где  $T$  – это температура системы. В качестве  $p(\Delta E, T)$  может быть взята, например, функция  $e^{-\frac{\Delta E}{T}}$ . Температура системы постепенно понижается в ходе работы алгоритма.

### 3 Условная оптимизация??

Обычно исходную функцию как-то модифицируют, чтобы получить задачу оптимизации без ограничений, но с решением, совпадающим с решением для исходной функции. Модифицированная функция – целевая функция (пример – минимизирующая функция Лагранжа).

#### 3.1 Метод проекции градиента

Метод проекции градиента модифицирует процедуру наискорейшего спуска так, чтобы она учитывала наличие ограничений-неравенств. Алгоритм состоит в следующем:

1. Выбрать начальную точку  $x^0$ .
2. На итерации  $k$  текущая точка  $x^k$ . Определить множество индексов насыщенных ограничений  $I^0$  – тех, для которых в точке  $x^k$  достигается равенство. Пусть  $A^0$  матрица, строки которой соответствуют ограничениям из  $I^0$ . Вычислить матрицу проектирования

$$P^0 = I - A^{0T}[A^0 A^{0T}]^{-1}A^0.$$

Затем вычислить  $d_k = -P^0 \nabla f(x^k)$ .

3. Пусть  $u = [A^0 A^{0T}]^{-1}A^0 \nabla f(x^k)$ . Если  $u \geq 0$  то перейти к 4. Иначе, пусть  $u_i$  наибольшая по модулю отрицательная компонента  $u$ . Тогда можно положить  $I^0 = I^0 - \{i\}$  и вернуться к 2.
4. Если  $d_k = 0$ , то конец. Если  $d_k \neq 0$  найти  $\lambda_{max} = \max\{\lambda | x^k + \lambda d_k \in X\}$ . Найти такое  $\lambda_k \in [0, \lambda_{max}]$ , что

$$\lambda_k = \operatorname{argmin} g(\lambda), \text{ где } g(\lambda) = f(x^k + \lambda d_k).$$

Перейти от  $k$  к  $k + 1$  и перейти к шагу 2.