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 функции многих переменных. Это — *целевая функция*. Переменные — *параметры оптимизации*. Обычно имеются какие-то ограничения на эти переменные. Ограничения бывают двух типов: *ограничения-равенства* (с жёстким допуском) и *ограничения-неравенства* (формулируются в виде диапазона).

Задача оптимизации детерминированных систем:

$$f(x) \to min$$
$$c_1(x_0) = 0$$
$$c_2(x) \geq 0$$

2. Безусловная оптимизация

2.1. Покоординатный спуск

Алгоритм состоит в нахождении локальных минимумов последовательно в направлении каждой координаты. Возьмём некоторую переменную x_1, а остальные зафиксируем. Найдём минимум f по этой переменной. x_1^{min} — 1-я координата минимума. Заменим x_1 на x_1^{min}, повторим для x_2, x_3 и т.д.

2.2. Наискорейший спуск

Этот метод предполагает, что f непрерывна и везде 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 $\vec{x}^0$.
2. На k-й итерации $\vec{d}_k = -\nabla f(x^k)$; найти такое $\lambda_k \geq 0$, что: $\lambda_k = argmin\, g(\lambda)$, где $g(\lambda) = f(\vec{x}^k + \lambda \vec{d}_k)$ (задача одномерной оптимизации). Взять $\vec{x}^{k+1} = \vec{x}^k + \lambda \vec{d}_k$.
3. Проверить критерии останова. (Длина шага, изменение значения и т. п.) Если критерий выполнен, закончить. Иначе — повторить шаг 2.

В качестве критерия останова можно взять, к примеру:

$\|\nabla f\| \leq \varepsilon$ при заданном заранее ε. Или любое стандартное условие.

2.3. Случайный поиск

Есть методы, где перехож на следующий шаг осуществляется случайным образом. К ним относятся, например, метод с возвратом при неудачном шаге, метод наилучшей пробы, метод случайного поиска с постоянным радиусом и случайным направлением.

Рассмотрим метод наилучшей пробы:

1. Выбрать начальную точку $\vec{x}^0$, начальную длину шага λ^0 и ставим счётчик числа

итераций $r=0$.

2. Генерируем M случайных векторов $\vec{\psi}_i^r$, $i \in [1, M]$ и по формуле

$$\vec{x}_i^{r+1} = \vec{x}^r + \lambda^r \frac{\vec{\psi}_i^r}{\|\vec{\psi}_i^r\|}$$

находим пробные точки $\vec{x}_i^{r+1}$, $i \in [1, M]$.

3. Вычисляем значения $f(\vec{x}_i^{r+1})$, находим минимальное из этих значений.
4. Если оно меньше значения в $\vec{x}^r$ — $r=r+1$, переходим в 2. Иначе
5. Проверяем критерий останова. Выполнен — $\vec{x}\,' \approx \vec{x}_r$. Иначе, $r=r+1$, $\lambda^{r+1} = \alpha\lambda^r$, где $\alpha \in (0,1)$ — коэффициент уменьшения шага.

Возможные критерии останова:

$$\|\vec{x}^{r+1} - \vec{x}^r\| = \lambda^r \leq \varepsilon_X$$

$$|f(\vec{x}^{r+1}) - f(\vec{x}^r)| \leq \varepsilon_f$$

3. Условная оптимизация

Обычно исходную функцию как-то модифицируют, чтобы получить задачу оптимизации без ограничений, но с решением, совпадающем с решением для исходной функцией (к примеру, добавляют штраф — экспоненциальный рост функции при выходе за ограничения). Пример модифицированной функции — 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

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

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

3. Пусть $u = [A^0A^{0T}]^{-1}A^0 \bigtriangledown f(x^k)$. 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.