Математическое программирование, лекция 9

Метод Флетчера-Ривса

рассмотрим квадратичную функцию q(x)

$q(x) = a + b^T x + (1/2)x^T C x$

Одномерный поиск будем вести вдоль направлений взаимно сопряженных по отношению матрицы C. В качестве 1-ого направления поиска из начальной точки $x^{(0)}$ возьмем направление наискорейшего спуска
$d_0 = - \blacktriangledown f(x^{(0)}) = -G_0$
Вычислим $x^{(1)} = x^{(0)} - a_0^* G_0$
Далее произведем поиск в направлении d_1, сопряженном направлению d_0. Причем выберем вектор d_1 как линейную комбинации векторов d_0 и $-G_1$
$d_1 = -G_1$

Определим искомый параметр $gamma_0$. Запишем выражение для градиента функции (q(x)):
$\blacktriangledown q(x) = b^T + Cx^{(0)}, G_1 = b^T Cx^{(1)}$. Изменение градиента от точки $x^{(0)}$ к точке $x^{(1)}$ будет
$\blacktriangle G(x) = G_1 - G_0 = C(x^{(1)}) - x^{(0)} = C \blacktriangle x$

Для обеспечения сопряженности направлений d_0 и d_1 должно выполняться условие $d_1^T C d_0 = 0$, т.е.
$(-G_1 - gamma_0 G_0)^T C(-G_0) = 0$ (8.3)

На начальной итерации $x^{(1)} = x^{(0)} - a_0^* G_0$, отсюда
$-G_0 = \blacktriangle x / alpha_0^*$

Тогда 8.3 примет вид
$(-G_1 - amma_0 G_0)^T C(\blacktriangle x)/(alpha_0^*) = 0$

Используя (8.2), получим $(a_0^*! = 0)$

$G_1^T d_0 = 0$, исходя из первого поиска

$gamma_0 = (G_1^T G_1)/(G_0^T G_0)$
$gammo_0 = (||G_1||^2)/(||G_0||^2)$

Произведя поиск в направлении $d_1 = -G_1 - gamma_0 G_0$, найдем величину шага минимизирующую функцию
$q(x^{(1)} + alpha_1 d_1)$
Вычислим $x^{(2)} = x^{(1)} - alpha_1^* d_1$
Общая формула для сопряженных направлений поиска, предложенная Флетчером и Ривсом:
$d_k = -G_k + gamma_{k-1} d_{k-1},$
где $gamma_{k-1} = (||G_k||^2)/(||G_{k-1}||^2)$

4. Алгоритм метода Флетчера-Ривса
   Шаг 1. Задаем начальную точку x^{(0)}
   Определяем значение антиградиента в точке x^{(0)}
   d_0 = -G_0 = -▼f(x^{(0)})

Шаг 2. Находим значение шаги минимизирующую функцию
f(x^{(k)}) = alpha_kd_k) Т.е. проводим одномерный поиск вдоль направления d_lk

Шаг 3. Вычисляем x^{(k+1)} = x{(k)} + alphak^*d_k и f(x^{(k+1)}), G{k+1} = ▼f(x^{k+1})

Шаг 4. Проверяем является ли точка x^{(k+1)} точкой минимума функции например по условию малости градиента ||Gk+1|| <= E, если оно выполняется, о поиск закнчен x^* = x^{(k+1)}
В противном случае определяем новое направление $d_k+1 = -G_k+1 + (||G{(k+1)}||^2) / (||G_k||^2) d_k$

Полагаем k=k+1 и переходим к **шагу 2**.

Метод Полака-Рибьера

Метод Полака-Рибьера отличается от данного метода только формулой расчета параметра gammak :
$gamma{(k-1)} = (▲Gk^TG_k)/(||G{(k-1)}||)^2$

Метод гарантирует схождение выпуклой функции со сверхлинейной скоростью сходимости

Достоинства :

1. Простота вычислений и следовательно простота программирования
2. Невысокие требования к объему памяти ЭВМ, поэтому особенно полезны при решении задач большой размерности
3. Могут применяться для целевых функций, имеющих линии уровня вытянутые, изогнутые или обладающих острыми углам

4. Имеют высокую скорость сходимости
   Недостаток:
   Чувствительность к ошибкам округления, возникающих в процессе вычислений

Методы поиска оптимума функций многих переменных второго порядка

Метод Ньютона-Рафсона

1. Общая характеристика методов второго порядка. Суть состоит в том, что мы разлагаем исходную функцию в ряд Тейлора и отбрасываем все члены ряда больше 2 порядка

$$f(x) = f(x^{(k)}) + \blacktriangledown f(x^{(k)})^T \blacktriangle x + 1/2! \blacktriangle x^T H_f(x^{(k)}) \blacktriangle x + \ldots$$

Отбрасываем все члены разложения 3го порядка:
$f(x) = f(x^{(k)}) + \blacktriangledown f(x^{(k)})^T \blacktriangle x + 1/2! \blacktriangle x^T H_f(x^{(k)}) \blacktriangle x$

▼f(▲x) = 0, Т.е. $\blacktriangledown f(x^{(k)})^T + H_f(x^{(k)}) \blacktriangle x = 0$, откуда
$\blacktriangle x = x^{(k+1)} - x^{(k)} = -H_f^{-1}(x^k) \blacktriangledown f(x^k)$

Т.е. итерационный процесс для построения последовательности приближений в точке минимума функции f(x) (см. скрин 3.11.2023 14:15)

Какие условия позволяют применить метод к нашей функции (чтобы была обратная матрица матрицы Гессе)

1. Если функция не квадратичная, но выпуклая, то для обеспечения сходимости достаточно только, чтобы матрица Гессе была положительно определена на всех итерациях.
2. Если функция не выпуклая, то необходимо не только 1) но и чтобы начальная точка x_0 лежала вблизи точки x *, а последовательность длин шагов* a_k *сходилась к 1. В этом случае итерационный процесс определяется выражением: $x^{k+1} = x{k} - a_k*^H_f^{-1}(x^{(k)})▼f(x^{(k)})$

2. Алгоритм метода Ньютона-Рафсона
(см. фото 3.11.2023 14.23)

Квази-Ньютоновские методы (методы переменной метрики)

Проблема Ньютона в том, что нужно считать H и её H^{-1}

Вместо обратной матрицы Гессе возьмем матрицу A_k , которая должна каким-то образом заменять обратную матрицу Гессе.
A_k - метрика