Московский государственный университет имени М.В.Ломоносова
Факультет вычислительной математики и кибернетики
Кафедра математических методов прогнозирования

Отчет по заданию

Проекция на симплекс

Выполнил студент 517 группы

Камалов Руслан Рамилевич

Москва
2016

1 Постановка задачи

Ставится задача условной оптимизации

$$\min_{x\in Q}\frac{1}{2}\|x-v\|^2, \quad Q:\left\{\sum_{i=1}^{n}x_i=1, \quad x_i\geq 0\right\} \tag{1}$$

2 Решение

Задача (1) является выпуклой задачей оптимизации. Минимизируемая функция является сильно выпуклой, ограничения типа неравенства являются выпуклыми, ограничения типа равенства являются линейными. Выполнены условия Слейтера (например $x_i=\frac{1}{n}$). Следовательно условия ККТ являются для этой задачи необходимыми и достаточными условиями экстремума.

Выпишем Лагранжиан задачи

$$L(x,\lambda,\mu)=\frac{1}{2}\|x-v\|^2-\sum_{i=1}^{n}\lambda_i x_i+\mu\left(\sum_{i=1}^{n}x_i-1\right) \tag{2}$$

Условия ККТ принимают вид

$$\begin{aligned}
&\nabla_x L=(x-v)-\vec{\lambda}_i-\vec{\mu}=0,\\
&\lambda_i x_i\geq 0, \quad x_i\geq 0,\\
&x_i\geq 0, \quad \sum_{i=1}^{n}x_i=1
\end{aligned} \tag{3}$$

Распишем подробнее систему для градиента Лагранжиана задачи

$$x_i-\lambda_i=v_i-\mu\Rightarrow|v_i-\mu|=|x_i-\lambda_i|=x_i+\lambda_i, \tag{4}$$

последнее равенство следует из неотрицательности x_i и λ_i, а также условия дополняющей нежесткости. Действительно

$$|x_i-0|=x_i, \quad \lambda_i=0$$

$$|x_i-\lambda_i|=|0-\lambda_i|=\lambda_i, \quad \lambda_i>0$$

Сложим первое и последнее равенство из (4), получим

$$x_i=\frac{1}{2}(v_i-\mu+|v_i-\mu|)$$

Введем функцию $\phi_+(t)$ следующего вида

$$\phi_+(t)=\begin{cases}0, & t<0\\ t, & t\geq 0\end{cases}$$

Перепишем выражения для x_i, используя $\phi_+(t)$

$$x_i = \phi_+(v_i - \mu)$$

Распишем последнее ограничение из (3), получим

$$\Phi(\mu) = \sum_{i=1}^{n} \phi_+(v_i - \mu) = 1 \tag{5}$$

Мы выразили координаты решения задачи (1) через двойственную переменную μ. Для нахождения неизвестного μ необходимо решить скалярную задачу (5). Покажем, что эта задача имеет единственное решение. Действительно, отсортируем координаты вектора v по возрастанию

$$v_{i_0} \leq v_{i_1} \leq ... \leq v_{i_{n-1}}$$

При $\mu_* < v_{i_0}$ все слагаемые в (5) положительные и

$$\sum_{i=1}^{n} \phi_+(v_i - \mu_*) = \sum_{i=1}^{n} v_i - n\mu_*$$

Рассмотрим теперь $\mu' \in (v_{i_j}, v_{i_{j+1}}]$. Для всех $\{v_{i_k} : k \leq j\}$ значение $\phi_+(v_{i_k} - \mu') = 0$, а для $\{v_{i_k} : k > j\}$ значение $\phi_+(v_{i_k} - \mu') = v_{i_k} - \mu'$, отсюда получаем

$$\sum_{i=1}^{n} \phi_+(v_i - \mu') = \sum_{i \geq j}^{n} v_i - (n-j)\mu', \quad \mu' \in (v_{j-1}, v_j], \quad j = 1..n$$

Следовательно функция, стоящая в левой части уравнения (5) является кусочно-линейной, строгоубывающей(коэффициенты $-(n-j) \leq 0$) и непрерывной (как сумма непрерывных) на $t \leq v_{i_{n-1}}$, при $t > v_{i_{n-1}}$ $\phi(t) \equiv 0$, поэтому (5) имеет ровно одно решение.

Для отыскания решения уравнения (5), можно найти такой индекс j^*, что выполняется условие

$$\Phi(v_{j^*-1}) \geq 1, \quad \Phi(v_{j^*}) \leq 1, = 1..n \tag{6}$$

Это всегда возможно если $\Phi(v_{i_1}) > 1$, если это не так, то решение уравнения (5) лежит на полупрямой $(-\infty, v_{i_1}]$, где $\Phi(\mu)$ линейна, в этом случае примем $j^* = 0$.

После того как найден индекс j^*, удовлетворяющий условию (6), поиск решения уравнения (5) сводится к поиску решения линейного скалярного уравнения

$$\sum_{i \geq j^*}^{n} v_i - (n - j^*)\mu = 1,$$

отсюда

$$\mu* = \frac{1}{n - j^*}(\sum_{i \geq j^*}^{n} v_i - 1)$$

Для получения окончательного ответа достаточно подставить μ^* в выражение для x.

$$x_i^* = \frac{1}{2}(v_i - \mu^* + |v_i - \mu^*|)$$

3 Алгоритм

Algorithm 1 Проекция на симплекс

$v_{sort} \leftarrow sort(v)$ ▷ сортируем координаты вектора v по возрастанию
if $\Phi(v_{sort}[0]) \leq 1$ **then**
  $i \leftarrow v_{sort}[0]$
else
  $i \leftarrow 1$ ▷ ищем оптимальный индекс
  while $not(\Phi(v_{sort}[i]) \leq 1 \quad and \quad \Phi(v_{sort}[i-1]) \geq 1)$ **do**
    $i \leftarrow i + 1$
  end while
end if
$\mu = \frac{1}{n-i}(\sum_{j \geq i}^{n} v[j] - 1)$
for $j = 1..n$ **do**
  $x[j] = \frac{1}{2}(v[j] - \mu + |v[j] - \mu|)$
end for

Оценка времени работы

1. сортировка массива из координат вектора $v : \sim O(n \log(n))$
2. поиск оптимального индекса: $\sim O(\log(n))$
3. вычисление μ и $x : \sim O(n)$
4. итого суммарное время работы алгоритма : $\sim O(n \log(n))$