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

Отчет по бонусному заданию по МОМО: Проекция на симплекс

Студент 517 группы:
Оспанов А.М.

Москва, 2016

1 Описание задания

В данном задании предлагается придумать эффективный алгоритм вычисления евклидовой проекции $\pi_{\Delta_n}(v)$ заданной точки $v \in \mathbb{R}^n$ на симплекс:

$$\pi_{\Delta_n}(v) = \arg\min_{x \in \Delta_n} ||x - v||_2,$$

где

$$\Delta_n = \left\{ x \in \mathbb{R}^n_+ | \sum_{i=1}^n x_i = 1 \right\}$$

2 Решение

Т.к. норма всегда положительна, то $\arg\min\limits_{x \in \Delta_n} ||x - v||_2 = \arg\min\limits_{x \in \Delta_n} \frac{1}{2}||x - v||_2^2$. Таким образом можно решать задачу с квадратом нормы для упрощения выкладок.

Выпишем лагранжиан:

$$L(x; \lambda, \mu) = \frac{1}{2}||x - v||_2^2 - \lambda^T x + \mu(\sum_{i=1}^n x_i - 1)$$

и систему ККТ:

$$\begin{cases} \nabla L(x; \lambda, \mu) = 0 \\ x_i \lambda_i = 0, & i = 1, \dots, n \\ \sum\limits_{i=1}^n x_i = 1 \\ \lambda \geq 0 \\ x_i \geq 0 & i = 1, \dots, n \end{cases}$$

Так как задача выпукла и ограничения линейные, решив эту систему, получим точку минимума.

Посчитаем градиент лагранжиана и выразим λ:

$$\nabla L(x; \lambda, \mu) = x - v - \lambda + \mu\vec{1} = 0$$

$$\lambda = x - v + \mu\vec{1}$$

Преобразуем систему ККТ подставив λ:

$$\begin{cases} x_i(x_i - v_i + \mu) = 0 \\ x_i \geq 0 \\ \sum\limits_{i=1}^n x_i = 1 \\ x_i - v_i + \mu \geq 0 \end{cases}$$

где $i = 1, \dots, n$

Теперь введем следующие множества:

$$I_+(\mu) = \{i | x_i - v_i + \mu > 0\}$$

$$I_0(\mu) = \{i | x_i - v_i + \mu = 0\}$$

тогда:

$$x_i = \begin{cases} 0, & i \in I_+(\mu) \\ v_i - \mu, & i \in I_0(\mu) \end{cases} \tag{*}$$

Учитывая это, выразив $\sum\limits_{i=1}^{n} x_i$:

$$\sum_{i=1}^{n} x_i = \sum_{i \in I_+(\mu)} 0 + \sum_{i \in I_0(\mu)} (v_i - \mu) = 1$$

получим:

$$\sum_{i=1}^{n} (v_i - \mu)_+ = 1, \text{где } x_+ = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Обозначим $f(\mu) = \sum\limits_{i=1}^{n} (v_i - \mu)_+ = \sum\limits_{i=1}^{n} max\{v_i - \mu, 0\}$.

Распишем свойства этой функции:

- непрерывна
- строго убывает ($\mu_1 < \mu_2 \Rightarrow f(\mu_1) > f(\mu_2)$)
- выпукла

Докажем эти свойства:

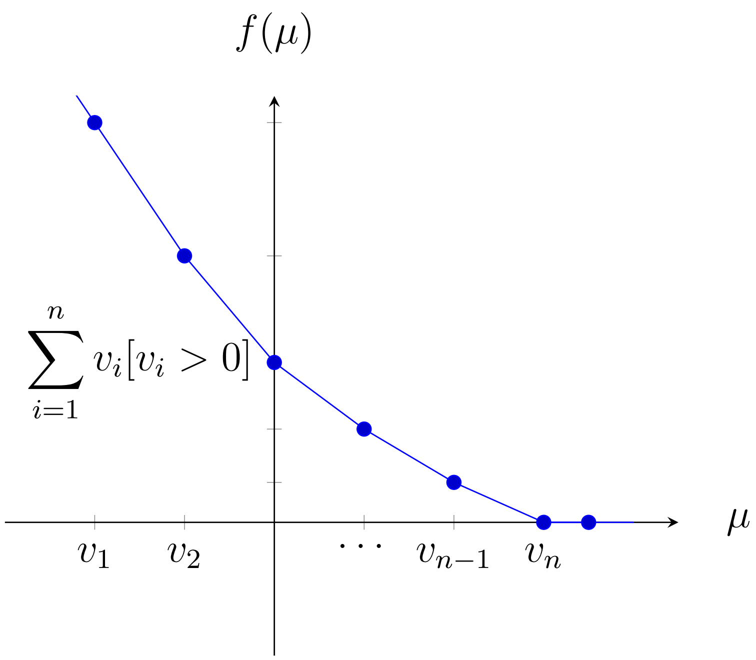
1) Непрерывность: очевидна (непрерывность линейной ф-ции и max)

2) Строгое убывание: Расположим v_i в порядке возрастания ($v_1 \leq v_2 \leq \dots \leq v_n$). Тогда видно, что при возрастании μ, ф-я убывает, причем строго.

3) Выпуклость: $v_i - \mu$ – линейная ф-я и следовательно выпукла. max – выуклая ф-я. $\Rightarrow$ Ф-я $f(\mu)$, как суперпозиция выпуклых ф-ций, – выпукла.

Далее будем считать, что v_i расположены в порядке возрастания ($v_1 \leq v_2 \leq \dots \leq v_n$).

Нарисуем график функции:



Пусть $v_i \leq \mu \leq v_{i+1}$. Тогда $f(\mu) = \sum\limits_{j=i+1}^{n} (v_j - \mu) = \sum\limits_{j=i+1}^{n} v_i - (n-i)\mu = 1$

Следовательно $\mu = \dfrac{\sum_{j=i+1}^{n} v_i - 1}{n-i}$

Чтобы вычислить x, перепишем (*), раскрыв множества:

$$x_i = \begin{cases} 0, & \mu > v_i \\ v_i - \mu, & \mu \leq v_i \end{cases}$$

Алгоритм 1 Проекция на симплекс

1: **function** proj(v_0)
2: $\quad v = sort(v_0)$
3: $\quad$ **for** $i = 1$ to $n-1$ **do**
4: $\quad\quad \mu = \dfrac{\sum_{j=i+1}^{n} v_i - 1}{n-i}$
5: $\quad\quad$ **if** $\mu \in [v_i, v_{i+1}]$ **then**
6: $\quad\quad\quad \mu_0 = \mu$
7: $\quad\quad\quad$ **break**
8: $\quad\quad$ **end if**
9: $\quad$ **end for**
10: $\quad$ **for** $i = 1$ to $n-1$ **do**
11: $\quad\quad$ **if** $\mu_0 > v_{0,i}$ **then**
12: $\quad\quad\quad x_i = 0$
13: $\quad\quad$ **else**
14: $\quad\quad\quad x_i = v_{0,i} - \mu_0$
15: $\quad\quad$ **end if**
16: $\quad$ **end for**
17: $\quad$ **return** x
18: **end function**

При некоторой оптимизации (например вместо вычисления суммы в строчке 4, вычислить полную сумму вначале, а в 4й строке отнимать v_i), то алгоритм работает за $O(n + n\log n)$ времени ($O(n\log n)$ времени тратится на сортировку), что то же самое, что и $O(n\log n)$.