

Задача распределенной оптимизации

Oleg Vasilev

March 22, 2020

1 Постановка задачи

Задан неориентированный граф коммуникации: в нем n агентов. Задана матрица смежности A и диагональная матрица степеней вершин Deg . Определена матрица Кирхгофа (Laplacian matrix) $L = Deg - A$ - т. е. сумма по каждой строчке и столбцу равна нулю.

Для каждого агента задана функция $f_i(x)$, и все агенты кооперативно действуют, чтобы оптимизировать сумму этих функций.

2 Сведение

Пусть каждый агент хранит свою копию переменных x_i . Введем функцию, в которой каждый агент считает свою функцию на своих переменных: $f'(x_1 \dots x_n) = \sum_i f_i(x_i)$. Тогда можно сформулировать исходную задачу в виде задачи с ограничениями:

$$\min f' \quad (1)$$

$$Lx = 0 \iff \sqrt{L}x = 0 \quad (2)$$

Т. е. на сумме функций достигается минимум при том, что очередной раунд коммуникации (домножение на L) не может ничего изменить.

Теперь полученную задачу можно свести к неограниченной при помощи аугментированного метода Лагранжа

$$A(x, u) = f'(x_1 \dots x_n) + \frac{\alpha}{2} \sum_{i,j} x_i \cdot x_j L_{ij} + \beta u^T \sqrt{L}x \quad (3)$$

Таким образом, переменные u можно рассматривать как двойственные - они показывают насколько то или иное ограничение выполнено.

Теперь можно решить полученную задачу прямо-двойственным алгоритмом, выполняя шаги одновременно по x и u . По x это будет шаг градиентного спуска, а по u - подъема в сторону увеличения важности насыщенных ограничений в выражении $A(x, u)$.

Можно выписать формулы пересчета переменных для i -того агента:

$$v_i = \sqrt{L}u_i \quad (4)$$

$$x_i \leftarrow x_i - \eta(\alpha \sum_j L_{ij}x_j + \beta v_i + \nabla f_i(x_i)) \quad (5)$$

$$v_i \leftarrow v_i + \eta\beta(\sum_j L_{ij}x_j) \quad (6)$$

Удивительным образом выражение для пересчета переменных зависит только от переменных, которые принадлежат "соседним" агентам - то есть тем, для которых в соответствующей строке L стоит не-ноль.

Далее, если специальным образом выбрать α и β , то можно добиться линейной скорости сходимости (т. е. очередное значение переменной в константу раз ближе к оптимуму). Константа будет зависеть от матричных свойств L , т. е. чем сложнее добраться от одной части графа до другой, тем хуже будет сходимость.

3 Альтернативное сведение

Можно использовать альтернативный способ сведения задачи с ограничениями к задаче без ограничений

$$\arg \min f(x) + \frac{R_y^2}{\varepsilon} \|Ax\|_2^2 \quad (7)$$

$$\arg \min \sum_i f(x_i) + \frac{R_y^2}{\varepsilon} \sum_i \sum_k \left(\sum_j L_{ij} x_j^{(k)} \right)^2 \quad (8)$$

$$x_i^{(k)} \leftarrow \nabla^{(k)} f(x_i) + \frac{2R_y^2}{\varepsilon} \left(\sum_j L_{ij} x_j^{(k)} \right) L_{ii} \quad (9)$$

$$x_i \leftarrow \nabla f(x_i) + \frac{2R_y^2}{\varepsilon} \left(\sum_j L_{ij} x_j \right) L_{ii} \quad (10)$$