ARTM

Виктор Панкратов, 774

10 марта 2021 г.

Напомним задачу максимизации функции на единичных симплексах. Далее Ω - набор $(w_j)_{j\in J}$

$$\begin{cases} f(\Omega) \to \max\limits_{\Omega} \\ \sum_i w_{ij} = 1,\ j \in J \\ w_{ij} \geq 0,\ i \in I_j, j \in J \end{cases} \tag{1}$$

Также напомним операцию нормировки вектора

$$p_i = \underset{i}{norm}(x_i) = \frac{\max(x_i, 0)}{\sum_k \max(x_k, 0)}$$

Лемма 1. *Пусть $f(\Omega)$ непрерывно дифференцируема по Ω. Тогда векторы w_j локального экстремума задачи (1) удовлетворяют системе:*

$$w_{ij} = \underset{i}{norm}\left(w_{ij}\frac{\partial f}{\partial w_{ij}}\right),\ \text{если } \exists i : w_{ij}\frac{\partial f}{\partial w_{ij}} > 0$$

$$w_{ij} = \underset{i}{norm}\left(-w_{ij}\frac{\partial f}{\partial w_{ij}}\right),\ \text{если } \forall i : w_{ij}\frac{\partial f}{\partial w_{ij}} \leq 0, \exists i : w_{ij}\frac{\partial f}{\partial w_{ij}} < 0$$

$$w_{ij}\frac{\partial f}{\partial w_{ij}} = 0, \text{иначе}$$

Доказательство. Запишем условия Каруша-Куна-Таккера для w_{ij}

$$\frac{\partial f}{\partial w_{ij}} = \lambda_j - \mu_{ij},\ \mu_{ij}w_{ij} = 0$$

$$w_{ij}\lambda_j = w_{ij}\frac{\partial f}{\partial w_{ij}} = A_{ij}$$

1. $\lambda_j > 0 \to w_{ij}\lambda_j = (A_{ij})_+ \underset{\sum_i}{\Rrightarrow} \lambda_j = \sum_i (A_{ij})_+$
2. $\lambda_j < 0 \to w_{ij}\lambda_j = -(-A_{ij})_+ \underset{\sum_i}{\Rrightarrow} \lambda_j = -\sum_i (A_{ij})_+$
3. $\lambda_j = 0$

$\square$

Данная лемма применяется для построения алогоритма нахождения общего решения задачи тематического моделирования:

$$\sum_{d\in D}\sum_{w\in d} n_{dw} \ln\left(\sum_{t\in T}\phi_{wt}\theta_{td}\right) + R(\Phi,\Theta) \to \max_{\Phi,\Theta} \tag{2}$$

$$\phi_{wt} \geq 0,\ \theta_{td} \geq 0,\ \sum_{w\in W}\phi_{wt} = 1,\ \sum_{t\in T}\theta_{td} = 1$$

Теорема 1. *Решение Φ, Θ задачи (2) при ограничениях неотрицательности и нормировки удовлетворяет следующей системе уравнений относительно переменных ϕ_{wt}, θ_{td} и вспомогательных переменных p_{tdw}:*

$$p_{tdw} = \underset{t\in T}{norm}(\phi_{wt}\theta_{td})$$

$$\phi_{wt} = \underset{w\in W}{norm}\left(n_{wt} + \phi_{wt}\frac{\partial R}{\phi_{wt}}\right),\ n_{wt} = \sum_{d\in D} n_{dw}p_{tdw}$$

$$\theta_{td} = \underset{t\in T}{norm}\left(n_{td} + \theta_{td}\frac{\partial R}{\theta_{td}}\right),\ n_{td} = \sum_{w\in d} n_{dw}p_{tdw}$$

Доказательство. Обозначим максимизируемую функцию в 2 за f. Применим лемму

$$\phi_{wt} = \underset{w \in W}{norm} \left(\phi_{wt} \frac{\partial f}{\partial \phi_{wt}} \right) = \underset{w \in W}{norm} \left(\phi_{wt} \sum_{d \in D} n_{dw} \frac{\theta_{td}}{p(w|d)} + \phi_{wt} \frac{\partial R}{\phi_{wt}} \right) =$$

$$= \underset{w \in W}{norm} \left(\sum_{d \in D} n_{dw} p_{tdw} + \phi_{wt} \frac{\partial R}{\phi_{wt}} \right)$$

$$\theta_{td} = \underset{t \in T}{norm} \left(\theta_{td} \frac{\partial f}{\partial \theta_{td}} \right) = \underset{t \in T}{norm} \left(\theta_{td} \sum_{w \in d} n_{dw} \frac{\theta_{td}}{p(w|d)} + \theta_{td} \frac{\partial R}{\theta_{td}} \right) =$$

$$= \underset{t \in T}{norm} \left(\sum_{w \in d} n_{dw} p_{tdw} + \theta_{td} \frac{\partial R}{\theta_{td}} \right)$$

□