Секция «Вычислительная математика, математическое моделирование и численные методы»

Выпуклая релаксация для выбора линейно-независимых столбцов в матрице

Научный руководитель – Оселедец Иван Валерьевич

Катруца Александр Михайлович

Аспирант

Сколковский институт науки и технологий, Информационные технологии, Москва, Россия

E-mail: aleksandr.katrutsa@phystech.edu

В работе исследуется проблема выбора линейно-независимых столбцов произвольной матрицы $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n] \in \mathbb{R}^{m \times n}$. Эта проблема является актуальной и возникает в различных приложениях. В частности, в задачах классификации и регрессии, где матрица $\mathbf{A}$ — это матрица плана с признаковым описанием объектов.

Предлагаемый подход основан на том, что матрица $\mathbf{A}$ неполного ранга тогда и только тогда, когда матрица $\mathbf{A}\mathbf{A}^\top$ вырождена. Так как произведение $\mathbf{A}\mathbf{A}^\top$ можно записать в виде

$$\mathbf{A}\mathbf{A}^\top = \sum_{i=1}^n \mathbf{a}_i \mathbf{a}_i^\top = \sum_{i=1}^n \mathbf{Y}_i,$$

то это позволяет переформулировать задачу определения линейно-независимых столбцов матрицы $\mathbf{A}$ как задачу определения такого бинарного вектора $\mathbf{x} \in \mathbb{B}^n$, что

$$\det\left(\sum_{i=1}^n x_i \mathbf{Y}_i\right) \to \max_{\mathbf{x} \in \mathbb{B}^n}. \tag{1}$$

Задача (1) не является выпуклой, так как целевая функция невогнута, и допустимое множество невыпукло, и поэтому не может быть эффективно решена. Для преобразования задачи (1) к задаче выпуклой оптимизации воспользуемся монотонным преобразованием и получим выпуклую целевую функцию, а также заменим допустимую область на пересечение её выпуклой оболочки с другим выпуклым множеством.

Известно, что функция $-\log\det(\mathbf{X})$ является выпуклой на множестве положительно определённых матриц $\mathbf{X} \in \mathbb{S}^n_{++}$ и функция $-\log x,\ x > 0$ монотонна. Также заметим, что выпуклая оболочка $\mathrm{Conv}(\mathbb{B}^n) = [0,1]^n$, и для получения более разреженного решения добавим ограничение на ℓ_1-норму решения, которое задаёт выпуклое множество по свойству нормы. Таким образом, получили следующую *выпуклую* задачу минимизации:

$$\begin{aligned} & -\log\det\left(\alpha\mathbf{I} + \sum_{i=1}^n z_i \mathbf{Y}_i\right) \to \min_{\mathbf{z} \in \mathbb{R}^n} \\ & \text{s.t. } \|\mathbf{z}\|_1 \le 1 \\ & z_i \ge 0, \end{aligned} \tag{2}$$

где $\alpha \simeq 10^{-10}$ — параметр регуляризации, обеспечивающий положительную определённость матрицы $\alpha\mathbf{I} + \sum\limits_{i=1}^n z_i \mathbf{Y}_i$.

Допустимым множеством задачи (2) является внутренность вероятностного симплекса с границей. Поскольку существуют быстрые алгоритмы поиска проекции точки на это множество, для решения задачи (2) предлагается использовать метод проекции градиента [1].

Источники и литература

1) Calamai, P. H., Moré, J. J. Projected gradient methods for linearly constrained problems // Mathematical programming. 1987, № 39(1). p. 93-116.