Анализ реализаций симплекс-метода в решателях с открытым кодом

Денис Лейбман
Научный руководитель: Роланд Хильдебранд
Команда: Екатерина Федоренко, Денис Лейбман

Московский физико-технический институт

leibman.da@phystech.edu

21 мая 2024 г.

Содержание

Напоминание

Стандартная форма линейной программы

$$\min_{x \geq 0} \; c^T x : Ax = b, \tag{1}$$

где $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n, 0 < \operatorname{rk} A = m < n$.

В процессе работы симплекс-метода поддерживаются два дизъюнктных множества индексов столбцов $B \subset [n], |B| = m$ и $N = [n] \setminus B$ – базисные и небазисные индексы, а также невырожденная подматрица A_{*B} - матрица базисных столбцов.

Примеры использования

- Производственное планирование
  Компания может определить оптимальные объемы производства различных продуктов с целью максимизации прибыли.
- Сети передачи данных
  Построение сети передачи данных так, чтобы обеспечить предопределённые требования за минимальную цену.
- Маршрутизация транспортных средств
  Определение оптимальных маршрутов для транспортных средств для минимизации общего расстояния или времени в пути

Итерация прямого симплекс-метода

1. CHUZC: выбрать некоторым способом индекс q в строке $\hat{c}_N$ текущих стоимостей – входящий элемент.
2. FTRAN: получить опорный q-ый столбец $\hat{A}_{*q} = A_{*B}^{-1} A_{*q}$.
3. CHUZR: выбрать некоторым образом индекс строки p – выходящий элемент, положить $\alpha = \hat{b}_p / \hat{A}_{pq}$. Обновляем столбец $\hat{b}$: $\hat{b} \leftarrow \hat{b} - \alpha \hat{A}_{*q}$.
4. BTRAN: получить строку $\pi_p^T = e_p^T A_{*B}^{-1}$.
5. PRICE: получить опорную p-ю строку $\hat{A}_{p*} = \pi_p^T A_{*N}$. Обновить $\hat{c}^T$: $\hat{c}^T \leftarrow \hat{c}^T - \hat{c}_q \hat{A}_{p*} / \hat{A}_{pq}$.
6. Либо заново провести LU-разложение для A_{*B}^{-1}, либо обновить текущее разложение, дописав новую матрицу в список факторов.

Наша работа и ее мотивация

Цели:

- Разобраться с работой симплекс метода.
- Собрать список проблем, с которыми приходится сталкиваться во время реализации.
- Проанализировать решатели с открытым кодом и собрать способы решения возникающих проблем.

Проделанная работа

- Разобрались с работой симплекс метода.
- Разобрали вместе с Екатериной решатель HiGHS, познакомились со способами решения вычислительных проблем на его примере.
- Разобрали 2 решателя и составили список решений выделенных проблем в них.
- Проанализировали получившиеся результаты и собрали их воедино.

Решатели с открытым кодом

GLPK (GNU Linear Programming Kit)

lpsolve
Mixed Integer Linear Programming (MILP) solver

HiGHS

SoPlex · SCIP · PaPILO · UG · GCG · ZIMPL
THE SCIP OPTIMIZATION SUITE

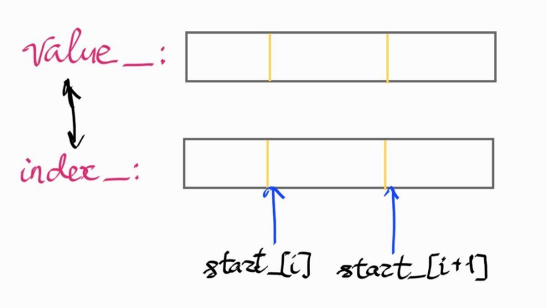
COIN|OR
COMPUTATIONAL INFRASTRUCTURE FOR OPERATIONS RESEARCH

Проблемы, которые были разобраны

- Множество проблем численного характера.
- Не систематизированный подход к решению этих проблем.
- Нулевые pivot элементы в LU разложении.
- Корректный выбор входящих и выходящих элементов базиса.
- Потеря прямой или двойственной допустимости при пересчете разложения базисной матрицы.
- Большие затраты памяти на хранение разреженных матриц при использовании стандартных структур данных.

Изучаемый решатель

- Метод Марковица для нахождения pivot элементов в LU разложении.
- Специальный собственный тип данных для повышения точности вычислений (четырехкратной точности).
- Хранение матрицы в виде:



- Специальный CHUZC для разреженной строки $e_p^T B^{-1} N$.
- Пертурбация стоимостей для избежания зацикливания и остановки.

Изучаемый решатель

GLPK (GNU Linear Programming Kit)

Борьба с зацикливанием

- Хранится массив c, в котором $c[i] = -1$, если базисная переменная $(x_B)_i$ незначительно меньше своей нижней границы, $c[i] = 1$ в соответствующем случае превышения верхней границы.
- Когда ищется допустимое для прямого симплекса решение, для некоторого i возможно $c[i] \neq 0$, хотя на самом деле выход за границу близок к нулю.
- Во избежание в таком случае зацикливаний такие индексы i зануляются искусственно.
- После каждого перехода к новому базису считается, что на предыдущем шаге все значения были строго допустимыми и на текущем могли несущественно выйти на границу.
- Если такие выходы происходят, выполняется пертурбация границ переменных, чтобы решение снова стало строго допустимым.

Выбор выходящего элемента

- Применяется метод, основанный на методе Харрис [Harris, 1973].
- Вместо того, чтобы просто сравнивать отношения $\hat{b}_p/\hat{A}_{pq}$, используется факт, что значения $\hat{b}_p$ хранятся не точно, для повышения численной устойчивости.
- Вычисляется минимальное допустимое изменение $\tilde{\theta}$, не выводящее за границы на переменную, причем границы сдвигаются на малое ε.
- Далее выбирается та строка p, для которой θ_p, посчитанное без сдвига границ на ε, не превышает $\tilde{\theta}$, а среди таких берется та, у которой больше $|\hat{A}_{pq}|$ для повышения численной устойчивости.

Выбор входящего элемента

Реализовано 2 метода:

- Стандартный метод Данцига, выбирающий такой столбец q, который максимизирует $|\hat{c}_q|$. Геометрический смысл состоит в выборе направления, которое имеет наиболее острый угол с антиградиентом целевой функции. Однако пространство, в котором считается угол, каждый раз меняется.
- Метод спроецированного наиболее острого ребра [Goldfarb, Reid, 1977] использует ту же идею, но уже в одном и том же пространстве $\mathbb{R}^n$. Сравниваются косинусы углов между вектором направления изменения и градиентом целевой функции.

Структуры для матриц и векторов

- Для разреженных матриц используется тот же способ с 3 массивами, что и в HiGHS.
- Для векторов используются 2 массива: индексы ненулевых элементов и массив всех элементов (в случае разреженного вектора только ненулевые элементы).
- Для перестановочных матриц P используется массив P_{per} длины $1 + 2n$, в котором $P_{ij} = 1 \Leftrightarrow P_{per}[i] = j, P_{per}[j] = i$.

LU разложение

- Как и в HiGHS, здесь используется выбор опорного элемента по стоимости Марковица в разреженном случае.
- Используется специальная структура Sparse vector area, которая хранит матрицу как список разреженных векторов, каждый из которых хранится как двусвязное кольцо (для сборки мусора).
- Для блочно диагональных разреженных матриц осуществляется особое разложение, являющееся LU разложением для каждого блока.
- Реализован метод Форреста-Томлина [Forrest, Tomlin, 1972] для обновления разложения матрицы, которое оставляет разреженными треугольные факторы.
- Поддерживается разложение, использующее рациональную арифметику.

Изучаемый решатель

Ремарки

- На самом деле для решения линейных программ SCIP использует другой пакет – SoPlex.
- Многие методы повторяют те, что уже были упомянуты выше.

Повторяющиеся решения

- Для обновления разложения матрицы используются либо домножение на eta-матрицу (как в HiGHS), так и метод Форреста-Томлина (как в GLPK).
- Для выбора опорного элемента в LU разложении используется метод Марковица.
- Также, как и в GLPK, есть альтернативные методы, использующие рациональную арифметику.
- Также, как и в GLPK, используются 2 способа хранения разреженных векторов через 2 массива: индексы ненулевых элементов и все элементы/ненулевые элементы. Есть и структура, хранящая множество разреженных векторов, соединенных двусвязным списком.

Выбор входящего элемента

- Стандартный метод Данцига.
- Метод Девекса, представляющий собой апроксимированный метод наиболее острого ребра, уменьшающий количество решаемых линейных систем и вычислений скалярных произведений.
- Стандартный метод самого острого ребра.

Выбор выходящего элемента

- Стандартный метод сравнения отношений (используется как шаблон для добавления новых методов, сам по себе не используется).
- Метод, являющийся модификацией метода Харрис, который пропускает вторую фазу, если уже после первой фазы видно, что есть надежный кандидат.
- Метод Харрис, использующий дополнительно пертурбацию границ переменных.

LU разложение

- Во всех решателях используется метод Марковица.
- Есть только некоторые отличия в реализации.

Хранение матрицы

- Через 3 массива значений, индексов и границ блоков.
- Через COO.
- Через двусвязный список разреженных векторов.

Борьба с зацикливанием

- Пертурбация стоимостей.
- Искусственное зануление ячеек в массиве штрафов за нарушение границ в случае незначительных нарушений.
- Пертурбация границ переменных.
- Правило Бланда выбора входящего и выходящего элемента.

Выбор входящего и выходящего элементов

Входящий элемент:

- Поддержка для разреженных матриц быстрого получения входящего элемента в CHUZC в HiGHS.
- Метод Данцига.
- Метод наиболее острого ребра.
- Метод Девекса.

Выходящий элемент:

- Стандартный 'книжный' метод сравнения отношений.
- Метод Харрис с различными вариациями:
  - Пропуск второй фазы, если после первой есть подходящий кандидат.
  - Пертурбация границ переменных.

Литература

D. Goldfarb, J. K. Reid (1977)
A practicable steepest-edge simplex algorithm
Mathematical Programming 12, 361-71.

P. M. J. Harris (1973)
Pivot selection methods of the Devex LP code
Mathematical Programming 5, 1-28.

J.J.H.Forrest and J.A.Tomlin (1972)
Updated triangular factors of the basis to maintain sparsity in the product form simplex method
Mathematical Programming 2, 263-78.