3 Симплекс-метод

Поиск оптимального решения ЗЛП путем простого перебора крайних точек допустимого множества возможен, но совершенно непрактичен с вычислительной точки зрения. Неэффективность такого подхода обусловлена астрономическим числом крайних точек у множеств, описываемых системами линейных неравенств. Например, такое простое множество, как n-мерный куб

$$B = \{x : 0 \leq x_j \leq 1, \ j = 1, 2, \ldots, n\}$$

имеет 2^n крайних точек. При достаточно скромной для практических приложений размерности $n = 100$ это число превосходит количество секунд, прошедших со времен Большого Взрыва – момента возникновения нашей Вселенной[1].

Ясно, что даже самый мощный суперкомпьютер не справится с задачей прямого перебора всех крайних точек и для более быстрого поиска оптимального решения ЗЛП необходим метод их направленного просмотра с максимально эффективным отсевом неперспективных вариантов. Одной из наиболее успешных идей в этом направлении был симплекс-метод, предложенный Дж. Данцигом и усовершенствованный затем многими вычислителями и программистами. Симплекс-метод – один из наиболее известных и распространенных алгоритмов решения ЗЛП, хотя в последнее время у него появились серьезные конкуренты в виде методов внутренних точек.

3.1 Базисные решения ЗЛП

Рассмотрим каноническую ЗЛП в матричной форме

$$cx \to \min, \quad x \in X = \{x \in \mathbb{R}^n : Ax = b, \quad x \geq 0\}. \tag{3.1}$$

Как было показано в предыдущем разделе, крайнюю точку допустимого множества X можно получить путем присвоения некоторым $(n - m)$ переменным нулевого значения, подстановка которых в ограничения $Ax = b$ приведет к невырожденной системе уравнений (2.7), решение которой даст значения оставшихся m неизвестных.

Обозначим через $N = \{i_1, i_2, \ldots, i_{n-m}\}$ множество индексов тех переменных, значения которых были положены равными нулю, а через $B = \{j_1, j_2, \ldots, j_m\}$ множество индексов тех переменных, значения которых были получены в результате решения системы уравнений (2.7). Объединение N и B дает совокупность всех индексов переменных задачи. Множество N будем называть *небазисным множеством* индексов, а B – *базисным множеством* индексов.

Во введенных обозначениях любую крайнюю точку x допустимого множества X задачи (3.1) можно представить в виде

$$x = (x_B, \ x_N), \quad x_B = \{x_i : i \in B\} \geq 0, \quad x_N = \{x_i : i \in N\} = 0. \tag{3.2}$$

[1]По современным представлениям, возраст Вселенной насчитывает 13.72 ± 0.12 млрд лет.

Компоненты вектора x_N будем называть *небазисными переменными*, а компоненты вектора x_B – *базисными переменными.* Если все базисные переменные положительны ($x_B > 0$), то такой базис называется *невырожденным.* Если вектор x_B имеет хоть одну нулевую компоненту, то базис называется *вырожденным.*

Определение 3.1 *Точки множества X вида* (3.2) *называются допустимыми базисными решениями канонической ЗЛП* (3.1).

Таким образом решение задачи (3.1) будем искать среди допустимых базисных решений.

Определение 3.2 *Допустимое базисное решение называется оптимальным, если в нем достигается экстремум целевой функции.*

Если ЗЛП имеет только невырожденные допустимые базисные решения, то и сама задача называется *невырожденной.* В противном случае ЗЛП называется *вырожденной.*

Множества базисных B и небазисных N индексов делит матрицу A ограничений общего вида на две подматрицы

$$A_B = \begin{pmatrix} a_{1j_1} & a_{1j_2} & \dots & a_{1j_m} \\ a_{2j_1} & a_{2j_2} & \dots & a_{2j_m} \\ \vdots & \vdots & \dots & \vdots \\ a_{mj_1} & a_{mj_2} & \dots & a_{mj_m} \end{pmatrix}, \quad A_N = \begin{pmatrix} a_{1i_1} & a_{1i_2} & \dots & a_{1i_{n-m}} \\ a_{2i_1} & a_{2i_2} & \dots & a_{2i_{n-m}} \\ \vdots & \vdots & \dots & \vdots \\ a_{mi_1} & a_{mi_2} & \dots & a_{mi_{n-m}} \end{pmatrix}$$

Матрицу A_B будем называть *базисной*, матрицу A_N – *небазисной.* При таком разбиении ограничения общего вида ЗЛП (3.1) перепишутся в виде

$$Ax = A_B x_B + A_N x_N = b. \tag{3.3}$$

Везде далее будем предполагать, что $\operatorname{rank} A = m$ и матрица A_B – невырожденная.

Аналогичным образом на базисные и небазисные коэффициенты можно разделить компоненты вектора целевой функции $c = (c_B,\ c_N)$, где $c_B = (c_i : i \in B),\ c_N = (c_i : i \in N)$. Тогда целевая функция задачи (3.1) перепишется в виде

$$cx = c_B x_B + c_N x_N. \tag{3.4}$$

В результате задачу (3.1) можно переписать в виде

$$\begin{gathered} c_B x_B + c_N x_N \to \min, \\ A_B x_B + A_N x_N = b, \\ x_B \geq 0, \quad x_N \geq 0. \end{gathered} \tag{3.5}$$

3.2 Основная идея симплекс-метода

Основная идея симплекс-метода состоит в последовательном улучшении значения целевой функции путем направленного перебора допустимых базисных решений. Геометрически процесс решения можно представить в виде последовательного перемещения из одной крайней точки допустимого множества в другую смежную такую, что значение целевой функции в новой точке будет не хуже, чем в предыдущей.

Предположим, что дано некоторое допустимое базисное решение $x^0 = (x_{B_0};\ x_{N_0})$ задачи (3.5). Компоненты x^0 разделены на базисные x_{B_0} (в соответствии с множеством B_0) и небазисные x_{N_0} (в соответствии с множеством N_0) переменные. Так как $x_{N_0} = 0$, то значение целевой функции (3.4) в точке x^0 равно $c_{B_0}x_{B_0}$. Возникает вопрос, достигается ли в точке x^0 минимальное значение целевой функции или можно найти другую крайнюю точку x^1 множества X такую, что $cx^1 < cx^0$. Очевидно, что точка x^1 будет определяться уже другими набором базисных B_1 и небазисных N_1 переменных.

Определим, как влияет изменение базиса на значение целевой функции. Из ограничений общего вида (3.3) выразим базисные переменные через небазисные:

$$x_B = A_B^{-1}b - A_B^{-1}A_N x_N = \beta - A_B^{-1}A_N x_N, \tag{3.6}$$

где $\beta = A_B^{-1}b$. Поскольку $x_N = 0$, то для текущего базиса B выполнено $x_B = \beta$.

Условие (3.6) дает возможность выразить целевую функцию (3.4) только через небазисные переменные x_N:

$$cx = c_B x_B + c_N x_N = c_B\beta + (c_N - c_B A_B^{-1}A_N)x_N = c_B\beta + (c_N - uA_N)x_N = c_B\beta + dx_N,$$

где $u = c_B A_B^{-1}$, $d = c_N - uA_N$. Величина $c_B\beta$ задает текущее значение целевой функции, а каждая из компонент вектора d показывает, как изменится это значение при изменении значения соответствующей небазисной компоненты из x_N на единицу. Вектор d называется *вектором модифицированных стоимостей*, а его компоненты — *модифицированными стоимостями.*

Нетрудно видеть, что значение целевой функции уменьшится, если для новых значений $x_N \geq 0$ величина $dx_N < 0$. Это возможно только в том случае, когда какая-либо компонента d отрицательна. В такой ситуации следует увеличить значения компонент x_N, соответствующих отрицательным элементам d. Если все модифицированные стоимости неотрицательны $d \geq 0$, то значение целевой функции улучшить невозможно, то есть текущая крайняя точка является оптимальным базисным решением. Таким образом, можно сформулировать критерий оптимальности решения ЗЛП.

Критерий оптимальности. *Если для текущего базиса B все компоненты вектора модифицированных стоимостей d неотрицательны, то точка $x^* = (x_B^*,\ \mathbf{0}_{n-m})$ является оптимальным базисным решением.*

Будем увеличивать значение только одной небазисной переменной, соответствующей отрицательной модифицированной стоимости. Выберем $d_{i_N} < 0$, тогда с целью уменьшения значения целевой функции присвоим переменной x_{i_N} некоторое положительное значение λ, при этом вектор x_N заменится на

$$\bar{x}_N = (0,\ 0,\ \ldots,\ 0,\ \underset{\substack{\uparrow \\ i_N}}{\lambda},\ 0\ \ldots,\ 0,\ 0)^T,$$

все компоненты $\bar{x}_N$ равны нулю, за исключением i_N-ой, которая приобретает значение $\lambda > 0$. Чем больше λ, тем меньше значение целевой функции.

Согласно соотношению (3.6) изменение значений компонент вектора x_N влияет на величину вектора x_B. Подставим $\bar{x}_N$ в равенство (3.6), получим вектор

$$\bar{x}_B = \beta - A_B^{-1} A_N \bar{x}_N = \beta - \lambda\alpha,$$

где α – это столбец матрицы $A_B^{-1}A_N$, соответствующий переменной x_{i_N}. Для того, чтобы новая точка $\bar{x} = (\bar{x}_B;\ \bar{x}_N)$ оставалась допустимым базисным решением задачи (3.5), значение λ можно увеличивать до тех пор, пока сохраняется условие неотрицательности на вектор $\bar{x}_B$.

Наибольшее $\lambda = \lambda_B$, при котором вектор $\bar{x}_B$ останется неотрицательным, можно определить по формуле:

$$\lambda_B = \min_{\alpha_i > 0} \frac{\beta_i}{\alpha_i} = \frac{\beta_{j_B}}{\alpha_{j_B}}, \tag{3.7}$$

где β_i и α_i – компоненты векторов β и α соответственно.

Если все компоненты вектора α неположительны, то λ можно выбрать сколь угодно большим и, следовательно, значение целевой функции сделать меньше любого наперед заданного числа, не нарушая условия неотрицательности для $\bar{x}_B$. Отсюда следует критерий неразрешимости ЗЛП.

Критерий неразрешимости. *Если в канонической ЗЛП найдется допустимое базисное решение $x = (x_B;\ x_N)$ такое, что существует индекс $i_N \in N$, при котором $d_{i_N} < 0$ и $\alpha \leq 0$, где α – столбец матрица $A_B^{-1}A_N$, соответствующий переменной x_{i_N}, то целевая функция на допустимом множестве неограничена снизу, следовательно, задача* (3.5) *не имеет решения.*

Чаще всего неограниченность целевой функции вызвана дефектами постановки задачи или ошибками при подготовке данных.

Отметим, что при $\lambda = \lambda_B$ компонента вектора $\bar{x}_B$ с номером j_B станет равной нулю

$$\bar{x}_{j_B} = \beta_{j_B} - \lambda_B \alpha_{j_B} = \beta_{j_B} - \frac{\beta_{j_B}}{\alpha_{j_B}} \alpha_{j_B} = 0.$$

Таким образом, одна из базисных переменных (с номером j_B) обращается в 0, а одна из небазисных переменных (с номером i_N) становится равной λ_B, следовательно, происходит смена базиса – переменная x_{j_B} выходит из базиса, а переменная x_{i_N} входит в базис. При этом получено новое допустимое базисное решение $\bar{x} = (\bar{x}_B;\ \bar{x}_N)$. Процедура изменения базиса повторяется, если в точке $\bar{x}$ не выполнен критерий оптимальности.

Идея направленного перебора допустимых базисных решений привела Дж. Данцига к алгоритму, известному как симплекс-метод.

3.3 Алгоритм симплекс-метода

Пусть задано начальное допустимое базисное решение $x^0 = (x_{B_0};\ x_{N_0})$, где B_0 и N_0 множества базисных и небазисных индексов. Положим $k = 0$.

Шаг 1. Проверка на оптимальность.

Для текущего базиса B_k вычислить модифицированные стоимости

$$d = c_{N_k} - c_{B_k} A_{B_k}^{-1} A_{N_k}. \tag{3.8}$$

Если все компоненты вектора d неотрицательны, то текущая точка x^k оптимальна и алгоритм заканчивает свою работу.

В противном случае, определить индекс $i_N \in N_k$ такой, что $d_{i_N} = \min\limits_{i \in N_k} d_i < 0$. Компонента x_{i_N} на следующей итерации войдет в базис. Перейти на шаг 2.

Шаг 2. Проверка на разрешимость.

Просмотреть столбец α матрицы $A_{B_k}^{-1} A_{N_k}$, соответствующий переменной x_{i_N}. Если все компоненты вектора α неположительны, то целевая функция исходной задачи неограничена снизу на допустимом множестве, следовательно, оптимального решения нет и алгоритм заканчивает свою работу.

В противном случае, определим индекс $j_B \in B_k$ такой, что

$$\frac{\beta_{j_B}}{\alpha_{j_B}} = \min_{\alpha_i > 0} \frac{\beta_i}{\alpha_i} = \lambda_{B_k}. \tag{3.9}$$

Компонента x_{j_B} на следующей итерации выйдет из базиса. Перейти на шаг 3.

Шаг 3. Пересчет базиса.

Построить новый базис $B_{k+1} = (B_k \setminus \{j_B\}) \cup \{i_N\}$, $\quad N_{k+1} = (N_k \setminus \{i_N\}) \cup \{j_B\}$. По формуле (3.6) вычислить новые значения переменных $x_{B_{k+1}}$. Перейти на шаг 1.

Заметим, что на шаге 1 для ввода в базис выбирается переменная с индексом i_N, соответствующая наименьшей отрицательной модифицированной стоимости d_{i_N}. С одной стороны, такой выбор разумен, поскольку минимальное d_{i_N} дает максимальное уменьшение значения целевой функции на единицу приращения небазисной переменной x_{i_N}. С другой стороны, итоговое уменьшение значения целевой функции (в данном случае это величина $d_{i_N} x_{i_N}$) зависит также и от того, на сколько допустимо увеличить значение x_{i_N}. В результате может оказаться, что максимальное уменьшение целевой функции на текущей итерации достигается путем введения в базис переменной, не соответствующей минимальному отрицательному значению модифицированной стоимости. Более того, строгое соблюдение правила $d_{i_N} = \min\limits_{i \in N_k} d_i < 0$ может оказаться неоптимальным с точки зрения общего количества итераций, необходимых для получения экстремума. Максимальный выигрыш на одной итерации может привести к медленному убыванию значения целевой функции на последующих шагах и, следовательно, замедлить процесс решения задачи.

Существует множество модификаций алгоритма симплекс-метода, отличающихся друг от друга правилами перехода от одного базиса к другому.

Как уже было отмечено, из теоремы 2.4 следует, что допустимая область ЗЛП имеет конечное число крайних точек. В канонической задаче допустимых базисных решений не больше чем C_n^m. Поэтому конечность симплекс-метода можно гарантировать только, если в процессе перебора крайних точек исключено повторение базисов. В противном случае процесс перебора может зациклиться.

Теорема 3.1 *Для невырожденной канонической ЗЛП симплекс-метод за конечное число шагов либо определит оптимальное базисное решение, либо установит, что задача не разрешима.*

`Доказательство.` Пусть $x^k = (x_{B_k};\ x_{N_k})$ — невырожденное допустимое базисное решение ЗЛП (3.5), то есть $x_{B_k} = \beta > 0$. Если для x^k выполнен критерий оптимальности, то алгоритм заканчивает свою работу. В противном случае выбираем небазисную переменную $x_{i_{N_k}}$ такую, что $d_{i_{N_k}} < 0$, и изменяем ее значение на λ_{B_k}, вычисляемое по формуле (3.7).

Если $\alpha \leq 0$, то алгоритм заканчивает свою работу, целевая функция неограничена снизу, следовательно, задача не имеет решений. В противном случае новое значение $x_{i_{N_{k+1}}} = \lambda_{B_k} = \dfrac{\beta_{j_B} > 0}{\alpha_{j_B} > 0} > 0$, следовательно, значение целевой функции в новой точке x^{k+1} будет строго меньше, чем значение в точке x^k

$$cx^{k+1} = c_{B_k}\beta + \lambda_{B_k} d_{i_{N_k}} < c_{B_k}\beta = cx^k.$$

Таким образом, на каждой итерации происходит строгое убывание значения целевой функции, что исключает повторение базиса, следовательно, процесс перебора рано или поздно закончится. □