Оптимизационные модели

Яцулевич Владимир Владимирович

1. БИНАРНЫЕ КВАДРАТИЧНЫЕ МОДЕЛИ (BQM)

Бинарные квадратичные модели (Binary Quadratic Model) — класс моделей основанный на двух фундаментальных классах: модель Изинга и модель QUBO (Quadratic Unconstrained Binary Optimization). Начнём с описания каждой из этих моделей.

Модель Изинга используется в статистической механике. Рассмотрим два состояния $|\uparrow\rangle$ и $|\downarrow\rangle$, которым будут соответствовать собственные значения $+1$ и -1. Будем рассматривать связи спинов, которые описываются либо корреляцией, либо антикорреляцией. Гамильтониан такой системы можно описать с помощью равенства:

$$H(\mathbf{s}) = \sum_{i=1}^{N} h_i s_i + \sum_{i=1}^{N} \sum_{j=i+1}^{N} J_{ij} s_i s_j, \tag{1}$$

где N —кол-во частиц, $\mathbf{s} = \{s_i\}_{i=1}^{N}$, $s_i \in \{-1, +1\}$, J_{ij} — коэффициент взаимодействия i-ой и j-ой частицы, h_i — коэффициент внешнего поля на i-ую частицу. Задача заключается в отыскивании минимального значения энергии.

QUBO модель очень похожа на модель Изинга. Только модель Изинга появляется в статистической механике, а модель QUBO изначально чисто математическая задача. Пусть задана верхнетреугольная матрица Q размерности $N \times N$ с вещественными коэффициентами и бинарный вектор $\mathbf{x} = \{x_i\}_{i=1}^{N}$, $x_i \in \{0, 1\}$. Рассмотрим целевую функцию

$$f(\mathbf{x}) = \sum_{i} Q_{ii} x_i + \sum_{i<j} Q_{ij} x_i x_j. \tag{2}$$

Задача заключается в минимизации целевой функции $f(\mathbf{x})$. Эту задачу можно переписать в более простой форме

$$\min_{\mathbf{x} \in \{0,1\}^n} \mathbf{x}^T Q \mathbf{x}. \tag{3}$$

Матрица Q в модели QUBO может быть двух типов: симметричная или верхнетреугольная. Для перехода к симметричной форме все элементы q_{ij} с индексами i и j нужно заменить на $(q_{ij} + q_{ji})/2$. Для перехода к верхнетреугольной матрице нужно заменить все элементы q_{ij} с индексами $j > i$ на сумму $q_{ij} + q_{ji}$. А элементы q_{ij} с индексами $j < i$ заменить нулями.

2. СВЕДЕНИЕ МОДЕЛИ ИЗИНГА К МОДЕЛИ QUBO

Пусть имеется целевой функционал модели Изинга.

$$H(\mathbf{s}) = \sum_{i=1}^{N} h_i s_i + \sum_{i=1}^{N} \sum_{j=i+1}^{N} J_{ij} s_i s_j. \tag{4}$$

Сделаем замену переменных $s = 2x - 1$. Если $x = 0$, то $s = -1$. Если $x = 1$, то $s = 1$.

$$\begin{aligned} H(\mathbf{x}) &= \sum_{i=1}^{N} \sum_{j>i} J_{ij}(2x_i - 1)(2x_j - 1) + \sum_{i=1}^{N} h_i(2x_i - 1) = \\ &= \sum_{i=1}^{N} \sum_{j>i} J_{ij}(4x_i x_i j - 2x_i - 2x_j + 1) + \sum_{i=1}^{N}(2h_i x_i - h_i) = \\ &= \sum_{i=1}^{N} \sum_{j>i} 4J_{ij} x_i x_j - \sum_{i=1}^{N} x_i \sum_{j>i} 2J_{ij} - \sum_{i=1}^{N} \sum_{j>i} 2J_{ij} x_j + \sum_{i=1}^{N} 2h_i x_i + k_0, \end{aligned} \tag{5}$$

где $k_0 = -\sum_{i=1}^{N} h_i$. Также введём ещё одно обозначение

$$a_i = \sum_{j>i} 2J_{ij} + 2h_i. \tag{6}$$

Тогда

$$H(\mathbf{s}) = \sum_{i=1}^{N} \sum_{j>i} 4J_{ij} x_i x_j + \sum_{i=1}^{N} a_i x_i - \sum_{i=1}^{N} \sum_{j>i} 2J_{ij} x_j + k_0. \tag{7}$$

Рассмотрим третье слагаемое

$$\begin{aligned} \sum_{i=1}^{N} \sum_{j>i} J_{ij} x_j = J_{12}x_2 + J_{13}x_3 + \dots + J_{1,N-1}x_{N-1} + J_{1N}x_N + \\ + J_{23}x_3 + J_{24}x_4 + \dots + J_{2,N-1}x_{N-1} + J_{2N}x_N + \\ \dots \\ + J_{N-2,N-1}x_{N-1} + J_{N-2,N}x_N + \\ + J_{N-1,N}x_N. \end{aligned} \tag{8}$$

Перегруппируем слагаемые, вынеся x_j за скобки.

$$\begin{aligned} \sum_{i=1}^{N} \sum_{j>i} J_{ij} x_j = (J_{1N} + J_{2N} + \dots + J_{N-2,N} + J_{N-1,N})x_N + \\ + (J_{1,N-1} + J_{2,N-1} + \dots + J_{N-2,N-1})x_{N-1} + \\ \dots \\ + (J_{13} + J23)x_3 + J_{12}x_2 = \sum_{i=1}^{N} b_i x_i, \end{aligned} \tag{9}$$

где $b_i = \sum_{j=1}^{i-1} J_{ij}$. Учитывая это разложение, получим

$$H(\mathbf{x}) = \sum_{i=1}^{N} \sum_{j>i} 4J_{ij}x_ix_j + \sum_{i=1}^{N} a_ix_i - 2\sum_{i=1}^{N} b_ix_i + k_0. \tag{10}$$

Введём новые обозначения $Q_{ij} = 4J_{ij}$, $Q_{ii} = a_i - 2b_i$. Тогда получим целевой функционал в форме QUBO.

$$H(\mathbf{x}) = \sum_{i=1}^{N} \sum_{j>i} Q_{ij}x_ix_j + \sum_{i=1}^{N} Q_{ii}x_i + k_0. \tag{11}$$

3. КВАДРАТИЧНЫЕ МОДЕЛИ С ОГРАНИЧЕНИЯМИ (CQM)

Квадратичные модели с ограничениями (Constrained Quadratic Model). Необходимо минимизировать функционал

$$\sum_i a_ix_i + \sum_{i\leqslant j} b_{ij}x_ix_j + c, \tag{12}$$

с заданными ограничениями

$$\sum_i a_i^{(m)}x_i + \sum_{i\leqslant j} b_{ij}^{(m)}x_ix_j + c^{(m)} \circ 0, \quad m = 1, \ldots, M, \tag{13}$$

где $\{x_i\}_{i=1}^N$ могут быть бинарными, целыми или вещественными переменными, a_i, b_{ij}, c — вещественные значения, $\circ \in \{\geqslant, \leqslant, =\}$, M — кол-во ограничений.

Рассмотрим более детально каким именно образом можно свести модель с линейными ограничениями к бинарной модели. Для начала определим условия для равенства. Пусть задано ограничение вида

$$x_1 + x_2 + x_3 = 1. \tag{14}$$

Для добавления этого условия к целевой функции необходимо представить ограничение в квадратичной форме.

$$\lambda(x_1 + x_2 + x_3 - 1)^2 = 0. \tag{15}$$

Если дано линейное ограничение общего вида

$$P(\mathbf{x}) = \sum_i a_ix_i - b = 0, \tag{16}$$

то для добавления его к целевой функции нужно сделать следующее

$$\lambda\left(\sum_i a_ix_i - b\right)^2 \equiv \lambda P(\mathbf{x})^2. \tag{17}$$

Далее рассмотрим пример с неравенством. Пусть задано неравенство с тремя двоичными переменными

$$x_1 + x_2 + x_3 \leqslant 2. \tag{18}$$

Поскольку мы сейчас рассматриваем только двоичные переменные, то мы можем от неравенства перейти к равенству. А затем использовать метод, описанный ранее. Для перехода к равенству нужно добавить две бинарные переменные y_1 и y_2.

$$x_1 + x_2 + x_3 = 2 - y_1 - y_2. \tag{19}$$

Здесь стоит добавить некоторое пояснение. Выражение «сумма меньше либо равна двух» на множестве двоичных цифр означает «сумма равна 2 или равна 1 или равна 0». Соответственно, первый исход достигается при $y_1 = y_2 = 1$, второй — при $y_1 = 0, y_2 = 1$ или $y_1 = 1, y_2 = 0$, а третий при $y_1 = y_2 = 0$.

После этого достаточно добавить ограничение к целевой функции, как это было показано ранее.

$$\lambda(x_1 + x_2 + x_3 + y_1 + y_2 - 2)^2. \tag{20}$$

В более общем случае линейное ограничение в виде неравенства

$$P(\mathbf{x}) = \sum_i a_i x_i - b \leqslant 0 \tag{21}$$

с положительными целыми коэффициентами a_i, b может быть добавлено к целевой функции через добавление дополнительного слагаемого вида

$$\lambda \left(\sum_i a_i x_i + \sum_{j=1}^{W} w_j y_j - b \right)^2 \equiv \lambda P(\mathbf{x})^2. \tag{22}$$

Число переменных W и их коэффициенты w_i могут быть определены через коэффициенты a_i и b. Чаще всего $W = b$.

В описанном решении нужно было, чтобы коэффициенты a_i, b были положительными и целыми. Если снять одно из ограничений, и допустить, что они могут быть отрицательными, то количество переменных возрастает до

$$W = b + \sum_{a_i < 0} |a_i| \tag{23}$$

в худшем случае.

По итогу полная целевая функция модели QUBO с ограничениями будет иметь вид

$$C(\mathbf{x}) = f(\mathbf{x}) + \sum_k \lambda_k P_k(\mathbf{x})^2. \tag{24}$$

4. ДИСКРЕТНЫЕ КВАДРАТИЧНЫЕ МОДЕЛИ (DQM)

Дискретные квадратичные модели (DQM) — полином некоторой дискретной переменной степени не превышающей двух. Дискретная переменная описывается некоторой коллекцией. Например $\{x, y, z\}$ или $\{1, 9, 0.5\}$, которые можно назвать классами. Целевой функционал определяется следующим образом:

$$H(\mathbf{d}) = \sum_i a_i(d_i) + \sum_{i,j} b_{ij}(d_i, d_j) + c, \tag{25}$$

где d_i — дискретная переменная, a_i и b_{ij} — вещественнозначные функции, c — смещение.

Любую дискретную модель можно свести к бинарной модели. Пусть D — это множество всевозможных классов. То есть $d_i \in D$. Тогда мы можем произвести кодирование категориального признака с помощью One-hot Encoding. А именно, каждому категориальному значению d поставим в соответствие бинарный вектор $\mathbf{x}$, который содержит только одно истинное значение, то есть

$$\sum_a x_{ia} = 1 \quad \forall i. \tag{26}$$

Пусть дискретная модель содержит N дискретных переменных d_i, каждая из которых содержит n_i классов. Тогда закодируем переменную d_i бинарной переменной x_{iu}, которая показывает принадлежит ли переменная d_i классу u. В таком случае целевой функционал можно переписать в виде

$$E(\mathbf{x}) = \sum_{i=1}^{N} \sum_{u=1}^{n_i} a_{iu} x_{iu} + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \sum_{u=1}^{n_i} \sum_{v=1}^{n_j} b_{ijuv} x_{iu} x_{jv}. \tag{27}$$

5. НЕПРЕРЫВНЫЕ МОДЕЛИ

С помощью квантового отжига также можно решать непрерывные оптимизационные задачи. Но в данном случае решение можно будет получить только с некоторой точностью. Пусть задана вещественная переменная $\tilde{x}$ объёмом n бит. Тогда эту переменную можно закодировать следующим образом

$$\tilde{x} = c \sum_{i=0}^{N-1} b^i x_i, \tag{28}$$

где $x_i \in \{0, 1\}$ — двоичные переменные, $b > 1$ определяет логарифмическое разрешение, $c \in \mathbb{R}$ определяет масштаб вещественной переменной, то есть её диапазон. Основная проблема такого подхода заключается в том, что с увеличением точности представления вещественного числа, увеличивается число кубитов, из-за чего увеличивается влияние шумов, в следствие чего увеличивается ошибка. То есть при небольшой точности представления числа получается небольшая физическая ошибка, но большая математическая ошибка. А при высокой точности представления числа получается большая физическая ошибка, но небольшая математическая ошибка.