Formal verification of Quadratic Unconstrained Binary Optimisation (QUBO) matrix construction for Quantum Annealer based optimal feature selection

Pauli Taipale

January 15, 2025

Contents

1 QUBO matrix form verification

We want to verify that when evaluated, the proposed QUBO matrix produces the following sums

$$\sum_{j=1}^{n} x_j |\rho_{V_j}| \tag{1}$$

$$\sum_{j=1}^{n} \sum_{\substack{k=1,\\ k\neq j}}^{n} x_j x_k |\rho_{jk}| \tag{2}$$

presented in [1]. If the evaluated QUBO matrix produces these sums, it is correctly constructed.

Correlation coefficients of model variables with respect to predicted outcome are

$$\rho_{Vj} = \begin{bmatrix} \rho_{V_1} \\ \rho_{V_2} \\ \vdots \\ \rho_{V_n} \end{bmatrix}$$

Correlation coefficients between n model variables can be represented as

$$\rho_{ij} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ \rho_{n1} & \cdots & \rho_{nn\text{-}1} & \rho_{nn} \end{bmatrix}$$

In this case the sign of correlation does not matter, so lets also specify that all the correlation coefficients are absolute values: $\rho_{Vj} := |\rho_{Vj}|$ and $\rho_{ij} := |\rho_{ij}|$. With only non-negative values in the matrix, we can later set signs on the matrix elements so that we end up with the correct energy function.

Assuming that a correct QUBO matrix can be constructed by replacing the diagonal of ρ_{ij} with ρ_{Vj}, we verify this by direct calculation of vector product form of the objective function we want to minimize for optimal feature selection:

$$E(\mathbf{x}) = \mathbf{x}^T Q \mathbf{x}, \tag{3}$$

which is equivalent with the sums (1) and (2) if the QUBO matrix Q is correctly constructed.

Starting with the simplest case of two model variables. x_1 and x_2 are binary elements of vector $\mathbf{x}$ that indicate whether or not a model variable associated with them has been chosen.

$$\begin{aligned}
\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \rho_{V_1} & \rho_{12} \\ \rho_{21} & \rho_{V_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} x_1\rho_{V_1} + x_2\rho_{21} & x_1\rho_{12} + x_2\rho_{V_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= x_1 \left(x_1\rho_{V_1} + x_2\rho_{21} \right) + x_2 \left(x_1\rho_{12} + x_2\rho_{V_2} \right) \\
&= x_1^2\rho_{V_1} + x_2^2\rho_{V_2} + x_1x_2\rho_{12} + x_2x_1\rho_{21} \\
&= \underbrace{\sum_{j=1}^{2} x_j |\rho_{V_j}|}_{\text{for binary values } x^2 = x.} + \sum_{j=1}^{2} \sum_{\substack{k=1, \\ k\neq j}}^{2} x_j x_k |\rho_{jk}|
\end{aligned}$$

When using three model variables we get

$$\begin{aligned}
\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} \rho_{V_1} & \rho_{12} & \rho_{13} \\ \rho_{21} & \rho_{V_2} & \rho_{23} \\ \rho_{31} & \rho_{32} & \rho_{V_3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} x_1\rho V_1 + x_2\rho_{21} + x_3\rho_{31} & x_1\rho_{12} + x_2\rho_{V_2} + x_3\rho_{32} & x_1\rho_{13} + x_2\rho_{23} + x_3\rho_{V_3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\
&= \left(x_1\rho_{V_1} + x_2\rho_{21} + x_3\rho_{31} \right) x_1 \\
&\quad + \left(x_1\rho_{12} + x_2\rho_{V_2} + x_3\rho_{32} \right) x_2 \\
&\quad + \left(x_1\rho_{13} + x_2\rho_{23} + x_3\rho_{V_3} \right) x_3 \\
&= \sum_{j=1}^{3} x_j |\rho_{V_j}| + \sum_{j=1}^{3} \sum_{\substack{k=1, \\ k\neq j}}^{3} x_j x_k |\rho_{jk}|
\end{aligned}$$

Similarly for n model variables, the result is

$$\begin{aligned}
\begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \end{bmatrix} \begin{bmatrix} \rho_{V1} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{V2} & \cdots & \rho_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ \rho_{n1} & \cdots & \rho_{nn\text{-}1} & \rho_{Vn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} &= \left(x_1\rho_{V_1} + x_2\rho_{21} + x_3\rho_{31} + \cdots + x_n\rho_{n1} \right) x_1 \\
&\quad + \left(x_1\rho_{12} + x_2\rho_{V_2} + x_3\rho_{32} + \cdots + x_n\rho_{n2} \right) x_2 \\
&\quad + \left(x_1\rho_{13} + x_2\rho_{23} + x_3\rho_{V_3} + \cdots + x_n\rho_{n3} \right) x_3 \\
&\quad \vdots \\
&\quad + \left(x_1\rho_{1n} + x_2\rho_{2n} + x_3\rho_{3n} + \cdots + x_n\rho_{V_n} \right) x_n \\
&= \sum_{j=1}^{n} x_j |\rho_{V_j}| + \sum_{j=1}^{n} \sum_{\substack{k=1, \\ k\neq j}}^{n} x_j x_k |\rho_{jk}|. \qquad \blacksquare
\end{aligned}$$

2 QUBO element signs

To obtain an objective function that is minimized at the optimum, we need to set the sign of the sum corresponding to feature indepence as negative and the sign of the sum corresponding to influence between the features as positive. We combine the terms using a parameter $\alpha(0 \leq \alpha \leq 1$) to represent the relative weighting of independence (greatest at $\alpha = 0$) and influence (greatest at $\alpha = 1$). The objective function which we want to minimize in order to select optimal set of features for an predictive model is thus

$$E(\mathbf{x}) = -\alpha \sum_{j=1}^{n} x_j \left|\rho_{Vj}\right| + (1-\alpha) \sum_{j=1}^{n} \sum_{\substack{k=1 \\ k \neq j}}^{n} x_j x_k \left|\rho_{jk}\right| \tag{4}$$

We can see that in order to get the vector product form of the objective function to match the summation form, we need to set the signs of diagonal terms ρ_{Vj} of the QUBO matrix as negative. Thus the final form of the QUBO matrix Q for n variables is

$$Q = \begin{bmatrix} -\alpha\rho_{V1} & (1-\alpha)\rho_{12} & \dots & (1-\alpha)\rho_{1n} \\ (1-\alpha)\rho_{21} & -\alpha\rho_{V2} & \dots & (1-\alpha)\rho_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ (1-\alpha)\rho_{n1} & \dots & (1-\alpha)\rho_{nn\text{-}1} & -\alpha\rho_{Vn} \end{bmatrix} \tag{5}$$

References

[1] Andrew Milne Maxwell Rounds and Phil Goddard. Optimal feature selection in credit scoring and classification using a quantum annealer, 2017.