**Note2: Theory, results and discussion on the 2D Ising model classification**

**Classification and logistic regression**

The classification problem resembles the regression problem, except that the target values we want to predict in classification are discrete variables (categories), which could take a binary or multiclass form. In binary classification the target value takes only two values, typically 0 and 1, or. Our goal in the classification problem is to predict the target value by some input via some non-linear function , which leads to solving the problem using logistic regression. The logistic regression function takes the form

Where is defined as

where are the parameters we want to estimate (estimators) and are the biases. The function as defined above, is commonly known as the sigmoid function. Because of the minimization of the cost function in logistic regression leads to non-linear equation for the estimators (Jensen, 2018), it need to be treated by gradient methods for minimizing the cost function. For a given training example the cost function for logistic regression including regularization can be defined as (Ng, 20)

where the second term is the penalty. We can rewrite it in vectorized form by including all training examples

We take the partial derivatives wrt the estimators

Finally, we update the estimators by gradient methods, such as gradient descent. Gradient descent, including penalty, can be defined as

**2D Ising model: results and discussion**

For this experiment we turn our focus at the 2D Ising model. In this section we will classify the magnetic phase of the 2D Ising model by using logistic regression. The Hamiltonian for the classical Ising model in this case reads

where expresses a summation of all adjacent neighbors in a 2D square lattice. The data was downloaded from XXXXXXXXXXXXXXXXXXXXXXXXXXXXX. We are not going to classify the disordered phase at (or close to) the critical temperature . In this case, we will predict the magnetic phases by binary classification using as the disordered phase and as the ordered phase. The input data, or design matrix, consists of the 2D square lattice structured as row-vectors, and the target phases as a column vector.

van Wieringen W. N. (2015). *Lecture notes on ridge regression*, arXiv:1509.09169.