Genetic alghoritm for Quantum Support Vector Machine

Lorenzo Tasca

October 2024

Contents

1 Introduction

2 Theoretical Background

2.1 Classical machine learning

2.1.1 Support vector machine

The Support Vector Machine (SVM) is a binary classification alghoritm, whose goal is to build the maximum margin separator between the two classes, that is the separator that maximizes the distance of the closest point from each class. The standard SVM alghoritm is a linear alghoritm, so in particular it will try and build the separating margin as a hyperplane in a d-dimensional space (so a $(d-1)$-dimensional plane), where d is the number of features. The points that touch the margin, or that are on the wrong side of it, are called support vectors. The distance between the decision boundary and the support vectors is called margin. The alghoritm will find the biggest possible margin.

Let's start assuming that the classes are linearly separable. We are provided with a dataset with N d-dimensional istances $\{\underline{x}_i\}_{i=0,\cdots,N-1}$. The two classes will be labelled with

$$y \in \{-1, 1\}.$$

The margin will be the the set of points

$$\{\underline{x} \in \mathbb{R}^d : \beta_0 + \underline{\beta}^T \cdot \underline{x} = 0\}, \tag{1}$$

for appropriate parameters $\beta_0 \in \mathbb{R}$ and $\underline{\beta} \in \mathbb{R}^d$, which define the hyperplane and must be found by the alghoritm. Now we have to find

$$\max_{\beta_0, \underline{\beta}}(m),$$

with the constraint

$$\frac{1}{||\underline{\beta}||} y_i(\beta_0 + \underline{\beta}^T \cdot \underline{x_i}) \geq m, \forall i = 0, \cdots, N-1, \tag{2}$$

that can be rewritten as

$$y_i(\beta_0 + \underline{\beta}^T \cdot \underline{x_i}) \geq m||\underline{\beta}||, \forall i = 0, \cdots, N-1. \tag{3}$$

The constraint prevents data points from falling into the margin. Rescaling $\underline{\beta}$ up to a multiplicative factor does not change the hyperplane it defines, so for convenience we can choose its norm such that

$$||\underline{\beta}|| = \frac{1}{m}. \tag{4}$$

Therefore the problem becomes minimizing

$$\frac{1}{2}||\underline{\beta}||,$$

with the constraint

$$y_i(\beta_0 + \underline{\beta}^T \cdot \underline{x_i}) \geq 1, \forall i = 0, \cdots, N-1. \tag{5}$$