Estimation of the neuromodulation parameters from the planned volume of tissue activated in deep brain stimulation

Thesis presented as partial requirement for the degree of:
Master in Electrical Engineering

Viviana Gómez Orozco

Director: Mauricio Alexander Álvarez López - MEng, PhD
Co-director: Óscar Alberto Henao Gallo - MSc, PhD

Universidad Tecnológica de Pereira
Facultad de Ingenierías Eléctrica, Electrónica, Física y Ciencias de la Computación
Maestría en Ingeniería Eléctrica - Línea de automática
Grupo de Investigación Automática
Pereira-Risaralda
2015
# Contents

Acknowledgements ............................................. ii
Abstract ....................................................... iii
1 Introduction .................................................. 1
2 Problem Statement ........................................... 2
3 Justification .................................................. 4
4 Aims ............................................................ 5
  4.1 General Aim ............................................... 5
  4.2 Specific Aims ............................................. 5
5 Background ................................................... 6
  5.1 The finite element method and deep brain stimulation ......... 6
  5.2 Volume of tissue activated and multicompartment neuron models ... 7
  5.3 Gaussian process .......................................... 7
  5.4 Support vector machine ................................... 8
6 Materials and Methods ....................................... 10
  6.1 Direct problem ............................................ 10
    6.1.1 Extracellular potential model ....................... 10
    6.1.2 Volume of tissue activated ......................... 12
    6.1.3 Computer emulator of the simulation model ......... 13
  6.2 Inverse problem ......................................... 14
    6.2.1 Procedure with SVMs ................................ 16
7 Results ....................................................... 18
  7.1 Direct problem ............................................ 18
  7.2 Generated dataset ........................................ 19
  7.3 Inverse problem ......................................... 19
    7.3.1 Evaluation of performance of the support vector machines . . . 19
    7.3.2 Evaluation of performance of the proposed methodology to reconstruct the volume of tissue activated . .... 23
8 Discussion and conclusion .................................... 26
I want to thank my family for their continuous and unconditional support during my studies, and my advisors, PhD Mauricio Alexander Álvarez López and PhD Óscar Alberto Henao Gallo, for their help and indispensable guidance during the development of this project. I am also thankful to the MSc in Electrical Engineering program at UTP and to the Automática research group, especially to PhD Genaro Daza Santacoloma and PhD Álvaro Ángel Orozco Gutiérrez for their contributions on several aspects of this work.

This research is developed under the project “Estimación de los parámetros de neuromodulación con terapia de estimulación cerebral profunda en pacientes con enfermedad de Parkinson a partir del volumen de tejido activo planeado”, financed by Colciencias with code 111065740687. The author, Viviana Gómez Orozco was supported by the 645 agreement, “Jóvenes Investigadores e Innovadores”.
Deep brain stimulation (DBS) is a therapy with promissory results for the treatment of movement disorders. It delivers electric stimulation via an electrode to a specific target brain region. The spatial extent of the neural response to this stimulation is known as volume of tissue activated (VTA). Changes in stimulation parameters that control VTA, such as amplitude, pulse width and electrode configuration can affect the effectiveness of the DBS therapy. In this study, we develop a novel methodology for estimating suitable DBS neuromodulation parameters, from a planned VTA, that attempts to maximize the therapeutic effects and to minimize the adverse effects for a patient in treatment. For estimating the continuous outputs (amplitude and pulse width), we use multi-output support vector regression, taking as the input space, the geometry of the VTA. For estimating the electrode polarity configuration, we perform several classification problems, also using support vector machines from the same input space. Our methodology attains satisfactory results for both the regression setting, and for predicting active contacts and their polarity. Combining biological neural modeling techniques together with machine learning, we introduce a promising area of research where parameters of neuromodulation in DBS can be tuned by manually specifying a desired geometric volume.
1. Introduction

Deep brain stimulation (DBS) is a therapy used for the treatment of neurological conditions, such as Parkinson’s Disease, essential tremor, and dystonia, among others. It is considered when drug therapy cannot suitably control the movement disorder symptoms. DBS involves the placing of an electrode within a target brain region (the basal ganglia, the thalamus, or other subcortical structures). Although DBS is an effective therapy, understanding the effects of neuronal response to electrical stimulation (its action mechanisms) remains unclear [1] [2] [3]. The fundamental purpose of DBS is to modulate neural activity with applied electric fields [4]. In this regard, a measure of the effects of deep brain stimulation is to estimate the volume of tissue activated (VTA), namely, the spatial spread of direct neural response to external electrical stimulation via a deep brain stimulation (DBS) electrode [5] [6].

Once the DBS electrode is implanted, one essential step is the configuration of the neurostimulation parameters. The neurostimulation is carried out by different active contacts in the electrode. Each active contact generates a rectangular pulse function, that shares the same amplitude, pulse width, and frequency. Furthermore, a contact could be active or inactive. If a contact is active, it can behave as a cathode or anode. Therefore, the neurostimulation parameters are the amplitude, pulse width, and frequency of the rectangular pulse function, and each lead contact configuration (cathode, anode or no stimulation). Although the frequency is an important parameter in DBS, it is considered that it does not have an influence in VTA estimation [7]. Accordingly, in this project, we consider as neurostimulation parameters the amplitude, and pulse width of the rectangular pulse function, and the configuration of each lead contact.

Variations in the electric stimulation parameters affect the spread of the activation, which can have serious consequences on therapeutic effects, or induce side effects when such parameters are not carefully adjusted [8]. The process for manually getting the right parameters can be quite time consuming since it usually relies on the experience of the medical specialist, commonly demanding several clinical sessions for a successful result. Therefore, in order to achieve the desired therapeutic benefits, and to reduce the amount of time spent by the patient and the specialist in clinical sessions, it is important to find a procedure for optimally adjusting the DBS device parameters.
2. Problem Statement

Once the DBS electrode is implanted in the patient’s brain via surgery, one essential step is the configuration of the neurostimulation parameters. The stimulation parameters are the amplitude, the pulse width, the frequency and each lead contact configuration (cathode, anode or no stimulation). Then, the stimulation is addressed by a rectangular pulse function, which can vary the amplitude and pulse width modulating the VTA for each electrode configuration [9]. Variations in the electric stimulation parameters affect the spread of activation, and can have serious consequences on therapeutic effects or induce side effects when such parameters are not carefully adjusted [8].

The VTA, and its visualization, jointly with reconstructions of the brain structures surrounding the implanted electrode [6], has been proposed as an alternative to accelerate the process of stimulation parameters adjustment, also minimizing the adverse side effects that can occur if such stimulation parameters are not carefully adjusted [8]. As a measure representing the extent of direct neural activation in response to the electrical stimulation, the VTA, as part of a visualization system, allows the medical specialist to observe the brain structures that are responding directly to the electrical stimulation. Based on this information the clinician determines the possible clinical effects a given stimulation configuration can have on the patient. This approach, although considerably facilitates the adjustment of stimulation parameters, still involves its search with the method of trial and error, this time using computer simulation. Thus, the idea to develop a novel methodology for estimating suitable DBS neuromodulation parameters arises. The aim is to allow the specialist to define an objective VTA, that attempts to maximize the therapeutic effects and to minimize the adverse effects for a patient in treatment.

Although there is a huge amount of literature on the subject of computing the VTA given a specific set of neuromodulation parameters, that is, the direct problem [3] [5] [7] [10], the inverse problem, this is, the problem of computing a set of specific neuromodulation parameters given a desired VTA, has received less attention. A previous study about the estimation of VTA uses an approach based on artificial neural networks to characterize the effects the adjustment of stimulation parameters over DBS [11] [12]. A system is provided in [12], which seeks the correlation between the calculated, and the desired VTA, that allows them to determine the appropriate stimulation parameters for the desired volume. The drawbacks of this approach are that it can not appropriately represent high stimulation parameter values and/or complex electrode configurations (more than two active contacts), and the assumption of an isotropic tissue medium.

In this project, we propose a novel methodology which allows us to find a suitable configuration of the neurostimulation parameters that attempts to maximize the therapeutic benefits, and to minimize the side effects of the VTA for DBS treatment. We first employ a computer simulator of the VTA generation process, where the input data corresponds to different configurations of the neuromodulation parameters, and the output data corresponds to the graphical representation of the VTA. The simulator uses a model to compute the extracellular potential generated by the DBS, plus a model for estimating the neural activation that takes place due to the electrical stimulation. We refer to this step as the direct problem. We then solve an inverse
problem that takes as input space the graphical representation of the VTA, and the output space corresponds to the values of the parameters that we want to adjust. To solve the inverse problem, we employ a framework based on support vector machines using them either for multi-output regression (for simultaneously tuning the amplitude and the pulse width), or for classification (for tuning the configuration of each contact). The relevance of the SVM underlies in its robustness for generalization, and its ability for easily dealing with high-dimensional input spaces.
The process for getting the right stimulation parameters can be quite time consuming since it usually relies on the experience of the specialist, including several clinical sessions for a successful result. Therefore, to achieve the desired therapeutic benefits, and reduce the amount of time spent by the patient and the specialist in clinical sessions, it becomes important to find a procedure for optimally adjusting the DBS device parameters. The development of a novel approach where the neuromodulation parameters are estimated from the graphical representation of the VTA, we refer to this as the inverse problem, may be promising alternative. However, this problem lacks, in the state of the art, an established methodology for its solution [13], [14], and presents challenges from the point of view of computational cost, due to the high-dimensional data involved.

The development of this proposal will combine biological neural modeling techniques together with machine learning, and introduce a promising area of research where parameters of neuromodulation in DBS can be tuned by manually specifying a desired geometric volume.
4. Aims

4.1 General Aim

To develop a methodology to estimate the neuromodulation parameters that generate the planned volume of tissue activated in deep brain stimulation employing machine learning algorithms.

4.2 Specific Aims

Aim 1. To build a dataset with different geometrical representations of volume of tissue activated and their corresponding stimulation parameters.

Aim 2. To employ machine learning algorithms that capture the characteristics of the dataset built in specific aim 1 to estimate the neuromodulation parameters that generate the planned volume of tissue activated.

Aim 3. To evaluate the performance of the proposed methodology in terms of its ability to reconstruct the volume of tissue activated through an error metric.
5. Background

This section introduce the theory of the finite element method (FEM), and describes the framework for the proposed methodology. The first subsection provides an explanation about the FEM, and its role in deep brain stimulation. Then, we describe the volume of tissue activated and how is estimated. Finally, we present the theory that support the use of support vector machine to solve the inverse problem.

5.1 The finite element method and deep brain stimulation

The finite element method (FEM) is a numerical method that allows us to find solutions to problems that may be represented by partial differential equations and a set of boundary conditions. The method involves discretizing the problem domain, in a process known as meshing, to create a series of smaller subdomains or finite elements. The numerical solution corresponds to the values of the quantity of interest in the nodes or edges of discretized domain. The values inside each element are calculated using a series of interpolation functions previously established. The solution is obtained by solving a linear system of equations constructed from the differential equation governing the problem and its associated boundary conditions. The main advantage of FEM over other numerical methods is that the subdomains can have different shapes and sizes, which allows the method to work with intricate geometries in 2D or 3D. This allows for a greater emphasis on the regions of interest by making the mesh denser in such places, in order to achieve a more accurate solution [15]

In deep brain stimulation (DBS) the distribution of the extracellular potential ($V_o$) is modeled by Poisson’s equation (Equation 5.1) through finite element method. This equation computes the spatial distribution of the electric potential. If there are not sources the problem is reduced to solve Laplace’s equation. Figure 5.1 shows an example of the spatial distribution of the extracellular potential generated by the DBS electrode.

![Figure 5.1: Electric potential obtained solving Equation 5.1](image)
5. Background

\[ -\nabla (\sigma \nabla V_o) = 0, \]  

(5.1)

where \(\nabla\) is the differential operator, and \(\sigma\) is the electric conductivity of the tissue of the medium.

5.2 Volume of tissue activated and multicompartment neuron models

The volume of tissue activated (VTA) is a measure of the spatial spread of direct neural response to external electrical stimulation via a deep brain stimulation electrode. The gold standard for VTA estimation is to couple the extracellular potential generated by the electrode with multicompartment neuron models [11], and to solve the so-called cable equation, for each compartment in each neuron, to compute the changes in the transmembrane potential induced by the stimulation. In this context the cable equation is given by

\[
\frac{dV^{(n)}}{dt} = \frac{1}{c_m} \left[ \frac{1}{r_a d\pi} (\phi_m + \phi_o) - i_{ion} \right]
\]

\[
\phi_m = \frac{V^{(n-1)}_m - 2V^{(n)}_m + V^{(n+1)}_m}{\Delta x^2}
\]

\[
\phi_o = \frac{V^{(n-1)}_o - 2V^{(n)}_o + V^{(n+1)}_o}{\Delta x^2},
\]

where \(V^{(n)}_m\) is the transmembrane potential at the \(n^{th}\) compartment, \(V^{(n)}_o\) is the extracellular potential interpolated onto the \(n^{th}\) compartment, \(c_m\) is the specific membrane capacitance, \(\Delta x\) and \(d\) are the compartment length and diameter, respectively, and \(i_{ion}\) is the total ionic current flowing through the membrane at a given moment in time. \(r_a\) is the axial resistance per unit length and it is obtained as \(\frac{4R_a}{\pi d^2}\), where \(R_a\) is the specific membrane resistance.

Furthermore, such multicompartment neuron models are usually restricted to axons, since they have been shown to drive the basic neural response to external electrical stimulation [4]. Studies that included structures other than axons, such as the soma and dendritic trees, have consistently found that DBS produces a decoupling between axonal and somatic firing. In normal conditions axonal firing is dictated by somatic activity and the axon acts as a transmission line that carries information from the soma to the neurons downstream. During DBS the soma is largely inactivated and the axonal firing rate is replaced by the stimulation frequency, masking any pathological conditions that can alter somatic firing patterns [1] [16] [17]. In addition to the former behavior, the fundamental biophysics of how axonal response to external electrical stimulation works is independent of neuron type [4]. Both factors have contributed to the widespread use in the DBS modeling literature of the mammalian myelinated axon model first proposed by McIntyre et al. in [18].

5.3 Gaussian process

A Gaussian process (GP) is a random process with the characteristic that a finite number of random variables taken of the process follow a multivariate Gaussian distribution. The GP is completely specified by a
mean function and covariance function. The covariance function encodes the degree of similarity between observations in terms of the input values. A Gaussian process is generally used as a prior distribution on a space of functions. Uncertainty over the space of functions is updated using a set of observations and Bayesian inference. The Bayesian inference is a method of statistical inference in which a hypothesis about the behavior of a system is introduced before observing any data, a prior distribution, then Bayes’ theorem is used to calculate the corresponding posterior distribution given the observed data, that is, the uncertainty in the hypothesis after the data are observed [19]. The Gaussian processes can be used to solve both regression and classification problems [20].

5.4 Support vector machine

A support vector machine (SVM) is a machine learning technique characterized by its robustness for generalization, and its ability for easily dealing with high-dimensional input spaces. SVM map the training data to a high-dimensional feature space through a mapping vector, and it creates a separation hyperplane that constitutes a nonlinear separation boundary in the input space. Using a kernel function it is possible to calculate the separation hyperplane without mapping to a feature space. Initially, SVMs were employed for classification problems, later they were extended for regression problems [21]. A detailed mathematical explanation about SVM for classification and for single output regression can be found in appendix A.

Support vector regression for multiple outputs

The SVM regression method for multiple outputs that we use in this work is an approach which defines a hyper-sphere insensitivity zone, that allows us to penalize only once the samples that are not placed inside the insensitivity zone for solving multiple outputs regression problems [22] [23] [24]. Let \( \mathbf{x}_i \in \mathbb{R}^m \) be the input vector, and \( \mathbf{y} \in \mathbb{R}^Q \), the output vector. The relationship between \( \mathbf{x}_i \) and \( \mathbf{y} \) is assumed to follow

\[
\mathbf{y} = \mathbf{W}^\top \phi(\mathbf{x}) + \mathbf{b},
\]

where \( \mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_Q] \), \( \mathbf{b} = [b_1, \ldots, b_Q]^\top \), with a vector \( \mathbf{w}_j \) and a constant \( b_j \) for every output. The function \( \phi(\mathbf{x}) \) refers to a non-linear transformation to a higher-dimensional space \( d \), where \( d \gg m \).

Given a dataset \( \{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^n \), the purpose is to find the set of parameters \( \{\mathbf{W}, \mathbf{b}\} \) that minimizes the objective function

\[
L_P(\mathbf{W}, \mathbf{b}) = \frac{1}{2} \sum_{j=1}^{Q} \|\mathbf{w}_j\|^2 + C \sum_{i=1}^{n} L(u_i),
\]

where \( C \) is a regularization constant, \( L(u_i) \) is a loss-function, with \( u_i = \sqrt{\mathbf{e}_i^\top \mathbf{e}_i} \), and \( \mathbf{e}_i = \mathbf{y}_i - \mathbf{W}^\top \phi(\mathbf{x}_i) - \mathbf{b} \). As a loss function \( L(u_i) \), the authors of [23] use a quadratic loss with respect to an user defined constant \( \epsilon \)

\[
L(u) = \begin{cases} 
0, & u < \epsilon \\
(u - \epsilon)^2, & u \geq \epsilon.
\end{cases}
\]
An iteratively reweighted least squares (IRLS) procedure is used to estimate the parameters $W$, and $b$.

The solution of the problem above can also be expressed in terms of the vector of coefficients $\beta^j$ for each output, which relates to the original vectors $w^j$ through $w^j = \Phi^\top \beta^j$, where $\Phi = [\phi(x_1), \ldots, \phi(x_n)]^\top \in \mathbb{R}^{n \times d}$. The prediction $y_*$ for a new input vector $x_*$ can be computed as $y_* = K_* \beta + b$, where $\beta = [\beta^1, \ldots, \beta^Q]$, and $K_*$ is the kernel between $x_*$ and the training set.
6. Materials and Methods

This section presents all the methods and materials used in the development of the project. First, we provide a detailed explanation about how to compute volumes of tissue activated for a set of realistic stimulation parameters (direct problem), using the gold standard for VTA estimation, and a Gaussian process classifier in order to emulate the action of the gold standard for VTA estimation. After that, we present the steps to follow for building the dataset needed to solve the problem of computing a set of specific neuromodulation parameters given a desired VTA (inverse problem) via support vector machine, specified the Toolbox. Finally, we describe the error metric employed to evaluate the performance of the proposed methodology.

6.1 Direct problem

The key neurostimulation parameters in VTA estimation are amplitude, pulse width and electrode contacts configuration. In this work, we use a clinical DBS electrode (Medtronic DBS 3389 electrode, ACTIVA-RC stimulator [25]) which offers wide neurostimulation parameter ranges. In clinical studies, the DBS amplitude should not exceed a certain value that depends on the neurostimulator, e.g 3.6 V for the Soletra, and 5.5 V for the Kinetra neurostimulator. To observe changes in patient response, 0.5 V steps in amplitude levels are also needed [26].

The solution of the direct problem in VTA uses two models, the first one to compute the extracellular potential generated by the DBS electrode, and the second one to estimate the neural activation in response to the electrical stimulation [27]. The model that computes the extracellular potential was implemented in COMSOL Multiphysics 4.2. (Comsol Inc., Stockholm Sweden) [28], and it is used to solve Poisson’s equation by means of the finite element method (FEM). The model that estimates the amount of neural activation was computed using Neuron 7.3, configured as a Python module [29] [30]. Figure 6.1 is a schematic representation of the methodology used to obtain the solution of the direct problem.

In what follows, we briefly explain the extracellular potential model, and the computation of the volume of tissue activated.

6.1.1 Extracellular potential model

A simplified 3D model of a clinical DBS electrode positioned in the middle of a conductive extracellular medium with an isotropic conductivity [31] was built in COMSOL Multiphysics 4.2., and save as a model M-file. The electrode model consisted of four conductive contacts (with a conductivity of \( \sigma = 4 \times 10^6 \text{Sm}^{-1} \) 1.27 mm in diameter and 1.5 mm in height separated by insulating bands (\( \sigma = 1 \times 10^{-10} \text{Sm}^{-1} \) 0.5 mm in height, and of an insulating semicircular tip with radius 0.635 mm (Figure 6.2). The conductor extracellular medium consisted of a sphere of diameter 10 cm. The electric conductivity of the brain tissue can represented of two different ways, first it was assumed to be homogeneous and isotropic with conductivity \( \sigma = 0.3 \text{Sm}^{-1} \). Then, it was represented by anisotropic conductivities. Such anisotropic conductivities can
6. Materials and Methods

Figure 6.1: Graphical representation of the methodology used for solving the direct problem. First, the neurostimulation parameters (amplitude, pulse width, and conductive contacts configuration) are selected. Then, an extracellular potential model is executed, in the finite element method (FEM) software COMSOL Multiphysics 4.2, to compute the spatial and temporal distribution of the extracellular potential generated by the neurostimulation parameters. After that, a model of multicompartment myelinated axons is implemented, in NEURON 7.3 configured as a Python module, to determine axonal response to the electric stimulation. Finally, the volume of tissue activated is computed as the volume generated by the active axons.

be obtained from magnetic resonance imaging by mean of diffusion tensors. They were estimated from the DTI30 dataset, with the RESTORE (Robust Estimation of Tensor by Outlier Rejection) algorithm, and then linearly transformed to conductivity tensor. In both cases a relative permittivity $\varepsilon_r = 80$, and an encapsulation tissue layer of 0.5 mm around the electrode with conductivity of 0.18 Sm$^{-1}$ were employed. The DBS pulse was modeled as a rectangular pulse by imposing time dependent Dirichlet boundary conditions on the contacts of the DBS electrode considered as actives. The remaining contacts were left inactive. Dirichlet boundary conditions ($V_o = 0$ V) were also imposed on the boundaries of the extracellular medium. Finally, zero current flow conditions were imposed on the surfaces of the non active contacts, and insulating components of the electrode. The Poisson’s equation was used to compute the spatial and temporal distribution of the extracellular potential.

Once the extracellular potential model was built, it was run into Matlab 2012a using COMSOL with Matlab (COMSOL-Matlab LiveLink). We automated the solution of the Poisson’s equation for several possible neurostimulation parameter combinations (Table 6.1) by using an M-file in Matlab. A random sample of 500 neurostimulation configurations was taken for the available parameter values under study (Table 6.1), where each configuration described amplitude, pulse width, and the active contacts during the electrode stimulation.

$^1$Available at www.cabiatl.com/CABI/resources/dti-analysis/
6. Materials and Methods

Figure 6.2: A simplified 3D model of the DBS 3389 electrode. Four conductive contacts (numbered from 0 to 3 and separated by insulating bands), and an insulating semicircular tip comprise the simplified model.

Table 6.1: Set of possible neurostimulation parameters and electrode configuration used in this work

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplitude $(A)$</td>
<td>0.5-5.5 V in 0.5 V steps</td>
</tr>
<tr>
<td>Width pulse $(W)$</td>
<td>60-450 $\mu$s in 30 $\mu$s steps</td>
</tr>
<tr>
<td>Contact 0 $(c_0)$</td>
<td>0, -1, 1</td>
</tr>
<tr>
<td>Contact 1 $(c_1)$</td>
<td>0, -1, 1</td>
</tr>
<tr>
<td>Contact 2 $(c_2)$</td>
<td>0, -1, 1</td>
</tr>
<tr>
<td>Contact 3 $(c_3)$</td>
<td>0, -1, 1</td>
</tr>
</tbody>
</table>

* No stimulation (0), cathodal stimulation (1), anodal stimulation (-1)

6.1.2 Volume of tissue activated

The gold standard for VTA estimation consists in coupling the electric potential due to DBS with a model of multicomartment myelinated axons distributed around the electrode shaft. The volume of tissue activated is computed as the volume generated by the active axons. To compute the change in the transmembrane potential induced by the stimulation, the multicomartment myelinated axonal model [13] [27] was implemented in Python 2.7 with Neuron 7.3 configured as a Python module. Each axon includes 21 nodes of Ranvier, and 2 myelin attachment segments (MYSA), 2 paranode main segments (FLUT), and 3 internode segments (STIN) between each node. For more detailed information, see [18]. The axonal field was modeled as fibers of 5.7 $\mu m$ diameter. The straight axons were oriented in four different directions, perpendicular to the axis of the electrode, and positioned at a distance between axons of 0.5 mm in both the vertical and horizontal
The electrical potential was interpolated from the elements of the FEM mesh onto each of the sections that integrated the axonal model.

### 6.1.3 Computer emulator of the simulation model

Solving exactly the model for the multicompartiment myelinated axons described above is computationally expensive. For each neurostimulation parameter configuration, running the full simulation model takes around one hour. In order to reduce the computational complexity of computing 500 times the direct problem, we used a computer emulator for the simulation model. The computer emulator was based on a Gaussian process (GP) classifier. The GP classifier was trained to emulate the action of the gold standard for VTA estimation, that is, it was trained to determine the spatial extent of neuronal activation, predicting which axons were activated by the applied extracellular stimulus [36]. We assume a logistic likelihood function \( f \), prescribe a general purpose kernel (squared exponential covariance function with automatic relevance determination), and use Laplace’s approximation to the posterior. The hyperparameters of the kernel were selected minimizing the Laplace approximation of the negative log marginal likelihood. The inference method was chosen due to computational cost considerations [37]. The GP emulator reduces the computation time to about four mins [36], allowing the complete computation of 500 samples in a reasonable time. The methodology used for emulating the multicompartiment myelinated axonal model is summarized in Figure 6.3. Our final dataset is made of 500 values for the neurostimulation parameters together with the 500 geometrical VTAs generated by each configuration.

It is important to clarify that each VTA was estimated for the same elements of the FEM mesh with labels \( \{0, 1\} \) that determine which of these elements were not activated (label 0) or were activated (label 1) (see Figure 6.4). The applied stimulus was different for each sample.
6. Materials and Methods

Figure 6.4: The sphere represents the region of interest (subthalamic nucleus), where the direct problem was solved. Elements of the FEM mesh are closer together around the electrode than in other places. This is to improve the precision of the solution around the electrode. Green dots are elements labeled active (‘1’), and red dots are elements labeled not active (‘0’). We want to clarify that green and red dots are not superposed, it is a 3D image.

6.2 Inverse problem

Each geometrical representation of the VTA is described using an $m$-dimensional vector made of zeros, and ones. A value of zero in this vector indicates that the mesh element of the finite element solution (from the direct problem) was inactive due to the particular configuration of the neurostimulation parameters. A value of one in this vector indicates that the mesh element was active. The value of $m$ is determined by the elements of the FEM mesh (in our case, $m = 177,924$), that is, the elements necessary to cover an spherical volume of 10 cm centered around the subthalamic nucleus, our target brain region for DBS.

As explained in the introduction, we want to map the geometrical representation of the VTA to the neurostimulation parameters that were used for computing that VTA. In short, we want to build a vector-valued function $f$ that maps $m$-dimensional binary vectors $x$, to a six-dimensional vector $y = [A\ W\ c_0\ c_1\ c_2\ c_3]^\top$, where $A$ refers to the amplitude of the pulse, $W$ refers to the pulse width, and $c_k$ refers to the condition of contact $k$, with $k = 0, 1, 2, 3$.

The space of the brain region that we want to cover with the VTA leads to vectors $x$ with a dimensionality $m$ greater than a thousand, depending of the volume resolution that we aim for the VTA. For building functions in such high dimensional input spaces, we use a framework based on Support Vector Machines, where the computations involved are dependent on the number of samples (in our case, less than 500 for the training...
6. Materials and Methods

For constructing the function \( y = f(x) \), we use two types of support vector learners. We jointly model the amplitude \( A(x) \), and the pulse width \( W(x) \) using multiple-output support vector regression through the scheme proposed in [22] [23]. Although, we could use support vector regression for modeling \( A(x) \), and \( W(x) \) independently, we experimentally found that modeling them jointly offered better results. This result is expected since in generating the VTA, both \( A(x) \), and \( W(x) \) are correlated. We also modeled the contacts \( c_0, c_1, c_2, \) and \( c_3 \) as four different classification tasks, using support vector classifiers for each of them. The methodology employed for solving the inverse problem is depicted in Figure 6.5.

![Figure 6.5: Graphical representation of the methodology used to solve the inverse problem. First, the direct problem is executed several times for different configuration parameters. Then, a dataset is built for the neurostimulation parameters together with the geometrical VTAs. This information is used to train a support vector machine to obtain the neurostimulation parameters of a specific geometry of the VTA.](image)

The subject of multiple-output regression in the context of SVMs has been less studied in the literature, and in what follows, we briefly summarize the theory behind [22] [23] (section 5.4). The theory behind support vector classifiers is well known, and it can be found in different machine learning textbooks [38] [19]. For a detailed mathematical explanation about SVM for classification and for single output regression can be found in appendix A.

For all the SVM configurations (multiple output regression, and the independent classification tasks), we use a kernel function \( k(x, x') \) based on a Hamming distance \( D_H(x, x') \) between the binary vectors \( x \), and \( x' \), vectors obtained from the mesh given by the FEM model. The Hamming distance measures the number of positions at which the corresponding vectors, \( x \) and \( x' \), have different symbols. For the SVMs, we also tried the classical radial basis function (RBF) kernel, but the results were extremely poor. The RBF kernel might be useful when the input space is continuous, which is not our case.
6.2.1 Procedure with SVMs

For all the experiments with SVMs (multi-output regression, and classification), we performed training with seventy percent observations in the dataset, this is, 350 sample points; and testing, with thirty percent of the data, this is, 150 sample points. In the four classification problems \((c_0, c_1, c_2, \text{ and } c_3)\) we employed the PRTools Matlab toolbox [39], with our own kernel function. For each of these classification problems, the classes considered were three, namely, active contact behaving as anode (label 1), active contact behaving as cathode (label -1), and inactive contact (label 0). The performance measure that we report for classification is the accuracy. For the multi-output regression problem, we use the code provided in [40]. The performance measure that we report for regression is the relative difference between the test value and the predicted value. In both cases, the experiments were run 20 times with different training and testing sets to assess the performance.

We also evaluate the performance of the different SVM in terms of the dimensionality \(m\) of the input vectors \(x\), in other words, in terms of the spatial resolution of the VTA. We start with a dimensionality of \(m = 177,924\), and then we uniformly subsample each vector by factors of 10, 50, 100 and 200 obtaining input spaces of dimensionalities \(m = 17,793\), \(m = 3,559\), \(m = 1,780\) and \(m = 890\), respectively. Figure 6.6 shows an example of the geometries of the VTA for the different input data resolutions considered.

![Figure 6.6: Graphic representation of the VTA geometry taking different resolution values to assess behavior of SVM under lower dimensional input data \((A = 1.0 \ V, W = 390 \ \mu s, c_0 = 1, c_1 = -1, c_2 = -1, c_3 = 0)\), where (a) \(m=177,924\). (b) \(m= 17,793\). (c) \(m=3,559\). (d) \(m=1,780\). (e) \(m=890\).]

We use different statistical tests to study if there are differences that are significant among the performances obtained in terms of the dimensionality \(m\), for a particular classification task, or for the multiple-output

\[^{2}\text{The relative difference between two values } x \text{ and } y \text{ is defined as } |x - y|/\max(x, y), \text{ where } |x| \text{ is the absolute value of } x, \text{ and } \max(x, y) \text{ is the maximum value between } x \text{ and } y\]
regression task. First, we apply a Lilliefors test for normality over the 20 repetitions of each value of \( m \). If the null hypothesis for normality is rejected, we perform a Kruskal-Wallis test to compare median performances among the values of \( m \), otherwise we use ANOVA. If the null hypothesis for equal medians is rejected, we perform a multiple comparison test using Tukey-Kramer to study further which performances in terms of \( m \) are different. All the significance levels are measured at 5%. Details for this procedure can be found in \([41]\).

The error metric used in function of the spatial resolution of the VTA was the Positive Matching Index (PMI). The PMI is a similarity measure between objects based on lists of their attributes \([42]\). It ranges from 0 to 1, where a value of zero indicates that the compared samples are dissimilar, while a value of one indicates that the samples are similar. This measure is expressed in terms of the true positives (TP), false positives (FP) and false negatives (FN) with respect to the reference data set

\[
PMI_{FP=FN} = \frac{TP}{|TP - FP|} = \frac{TP}{|TP - FN|}
\]

\[
PMI_{FP\neq FN} = \frac{TP}{|FP - FN|} \ln \left( \frac{TP + \max(FP,FN)}{TP + \min(FP,FN)} \right).
\]

We also tried a metric based on the predictor error,\(^3\) but the results gave a comparable information, for this reason it was not included in the document. In addition, this measure is unbounded.

\(^3\)The predictor error between two values is defined as \( \frac{FP+FN}{FP} \).
7. Results

This section shows the results for the different phases of the proposed methodology. First, we simulate the direct problem for several neuromodulation parameter configurations. Then, we compare the data obtained with direct problem, under both isotropic and anisotropic conditions. Finally, we evaluate the performance of the support vector machines, in terms of its ability to reconstruct the volume of tissue activated.

7.1 Direct problem

The main objective of the direct problem is to estimate the VTA given a specific set of neuromodulation parameters, using a DBS electrode model inside the brain tissue, onto a region of interest, under both ideal (isotropic tissue conductivities) or realistic (anisotropic tissue conductivities) assumptions, to modulate the neural activity. The idea is to deliver an electrical pulse to the target region, and then study the neural response to this stimulation, the VTA. Figure 7.1 shows the potential distribution for an specific neurostimulation parameter setting and the corresponding estimated VTA.

![Diagram of model solving the direct problem](image)

Figure 7.1: Graphic representation of the model that solves the direct problem. (a) Extracellular potential generated by the DBS electrode in Comsol Multiphysics \((A = 4.5 \text{ V}, W = 90 \mu s, c_0 = -1, c_1 = 0, c_2 = 1, c_3 = -1)\), with a schematic representation of the stimulation delivered through each electrode contact. (b) Estimated VTA induced by the electrical stimulation.
7. Results

7.2 Generated dataset

The purpose of the dataset generated was to capture different parameter configurations and their corresponding VTAs. Figure 7.2 shows six samples of possible parameter combinations (see Table 6.1), under isotropic conditions, and how they affect the volume that represents the spatial spread of neural activation due to the electric stimulation. A monopolar stimulation, with a low amplitude and a high pulse width value, is represented in Figure 7.2(a), it shows that high values of pulse width induce a small neural response. A bipolar stimulation, anode-cathode, is described by Figure 7.2(b), while a cathode-cathode is described by Figure 7.2(c). The relevance in this configuration is to show how the polarity affects the VTA. The other Figures (Figure 7.2(d), Figure 7.2(e) and Figure 7.2(f)) allow us to see the influence of the parameter variations onto the estimated volume.

Figure 7.3 shows the same samples of figure 7.2 considering anisotropic tissue conductivity conditions. In this case, figures show that the volumes of tissue activated estimated do not have the same symmetry with respect to the axis of the electrode, as exhibited in the isotropic case. These changes in the volume are the result of considering the conductivity tensors in the electric potential, which affect the direction of propagation.

7.3 Inverse problem

7.3.1 Evaluation of performance of the support vector machines

Table 7.1 summarizes the accuracy obtained for each classification task in terms of the different resolutions of the input data, considering isotropic conductivity conditions. We obtain a slightly better accuracy for contacts $c_0$, $c_1$ and $c_3$ when the dimensionality of the input data is $m = 17,793$, under isotropic conductivities. The best result for contact $c_2$ was obtained when the dimensionality of the input space was equal to $m = 177,924$ (see Table 7.1). We apply the statistical significance test described in section 6.2.1 for each classification task, obtaining as a result that the performances for all the contacts are not statistically different in terms of the value of $m$. This result can be explained by the fact that we are assuming isotropic conductivities for the extracellular potential model, and the shapes for the different VTA generated have regular forms. Based on the results of Table 7.1 we conclude that the classification accuracy is in the range of 80% (an error rate of 20%) for contacts $c_1$ and $c_2$, and 86% (an error rate of 14%) for contacts $c_0$ and $c_3$.

Table 7.1: Accuracy in each electrode contact configuration (anodal, cathodal or no stimulation) with different values of dimensionality $m$ of the input space $x$, under isotropic conditions

<table>
<thead>
<tr>
<th>Dimensionality $m$ of $x$</th>
<th>Contact 0 ($c_0$)</th>
<th>Contact 1 ($c_1$)</th>
<th>Contact 2 ($c_2$)</th>
<th>Contact 3 ($c_3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>890</td>
<td>0.8455±0.0277</td>
<td>0.7741±0.0257</td>
<td>0.7929±0.0333</td>
<td>0.8562±0.0209</td>
</tr>
<tr>
<td>1,780</td>
<td>0.8528±0.0267</td>
<td>0.7953±0.0225</td>
<td>0.8030±0.0282</td>
<td>0.8594±0.0227</td>
</tr>
<tr>
<td>3,559</td>
<td>0.8524±0.0292</td>
<td>0.8002±0.0269</td>
<td>0.8116±0.0238</td>
<td>0.8613±0.0252</td>
</tr>
<tr>
<td>17,793</td>
<td>0.8612±0.0249</td>
<td>0.8105±0.0224</td>
<td>0.8173±0.0241</td>
<td>0.8607±0.0243</td>
</tr>
<tr>
<td>177,924</td>
<td>0.8607±0.0265</td>
<td>0.8061±0.0231</td>
<td>0.8192±0.0236</td>
<td>0.8569±0.0257</td>
</tr>
</tbody>
</table>

For the anisotropic case (see Table 7.2) when the dimensionality of the input data is $m = 17,793$ the best results in terms of accuracy were for the contacts $c_1$ and $c_2$, while the others two contacts ($c_0$ and $c_3$) achieved better accuracy when the dimensionality was $m = 17,793$. A statistical test was made as in
Figure 7.2: Graphic representation of different samples of the dataset generated, under isotropic conductivity conditions, and their respective amplitude (V), pulse width (µs) values, and electrode configuration.
Figure 7.3: Graphic representation of different samples of the dataset generated, under anisotropic conductivity conditions, and their respective amplitude (V), pulse width ($\mu$s) values, and electrode configuration.
isotropic case, obtaining similar results. Although, the shapes generated for the VTA have irregular forms. These irregular volumes can be justified because the tissue conductivity affects the VTA via changes in the propagation of the stimulus waveform and consequently on axonal activation. We can conclude that the classification accuracy is around 70% (an error rate of 30%) for contacts $c_1$ and $c_2$, and 77% (an error rate of 23%) for contacts $c_0$ and $c_3$. At this point, we may note that the accuracy is lower when we include the anisotropy, it can be explained by the fact that anisotropic conductivities generate complex shapes of the VTA.

Table 7.2: Accuracy in each electrode contact configuration (anodal, cathodal or no stimulation) with different values of dimensionality $m$ of the input space $x$, under anisotropic conditions.

<table>
<thead>
<tr>
<th>Dimensionality $m$ of $x$</th>
<th>Contact 0 ($c_0$)</th>
<th>Contact 1 ($c_1$)</th>
<th>Contact 2 ($c_2$)</th>
<th>Contact 3 ($c_3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>890</td>
<td>0.7690±0.0311</td>
<td>0.6863±0.0324</td>
<td>0.6767±0.0294</td>
<td>0.7471±0.0206</td>
</tr>
<tr>
<td>1,780</td>
<td>0.7761±0.0285</td>
<td>0.7027±0.0349</td>
<td>0.6842±0.0311</td>
<td>0.7469±0.0202</td>
</tr>
<tr>
<td>3,559</td>
<td>0.7779±0.0281</td>
<td>0.7152±0.0343</td>
<td>0.6885±0.0260</td>
<td>0.7560±0.0215</td>
</tr>
<tr>
<td>17,793</td>
<td>0.7795±0.0256</td>
<td>0.7113±0.0333</td>
<td>0.6920±0.0287</td>
<td>0.7613±0.0213</td>
</tr>
<tr>
<td>177,924</td>
<td>0.7797±0.0252</td>
<td>0.7102±0.0355</td>
<td>0.6891±0.0276</td>
<td>0.7680±0.0231</td>
</tr>
</tbody>
</table>

Table 7.3 shows the relative difference for the amplitude and the pulse width, in terms of the dimensionality $m$ of the input space considering isotropic conditions. Applying the statistical tests described before, we find that the discrepancies among the relative differences are not significant for the amplitude, nor for the pulse width. Again, this could be explained due to the regular forms of the VTA as a consequence of the isotropic conductivities assumed in the extracellular potential model. The relative difference for the amplitude is about 27%, whereas the relative difference for the pulse width is about 28%.

Table 7.3: Relative differences for the amplitude and pulse width with several values of the dimensionality $m$ for the input space $x$, under isotropic conditions.

<table>
<thead>
<tr>
<th>Dimensionality $m$ of $x$</th>
<th>Amplitude ($V$)</th>
<th>Pulse width ($\mu$s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>890</td>
<td>0.2825±0.1795</td>
<td>0.2868±0.2042</td>
</tr>
<tr>
<td>1,780</td>
<td>0.2737±0.1773</td>
<td>0.2898±0.2002</td>
</tr>
<tr>
<td>3,559</td>
<td>0.2727±0.1743</td>
<td>0.2871±0.2013</td>
</tr>
<tr>
<td>17,793</td>
<td>0.2743±0.1769</td>
<td>0.2888±0.2031</td>
</tr>
<tr>
<td>177,924</td>
<td>0.2750±0.1775</td>
<td>0.2888±0.2043</td>
</tr>
</tbody>
</table>

For the anisotropic case, Table 7.4 summarizes the relative differences for the amplitude and pulse width. Also, a statistical test was applied getting the same results as in the case isotropic, i.e., for both amplitude and pulse width they are not statistically different in terms of the value of $m$, respectively. Again, this can explained by the fact that these volumes have more complex shapes. We attain relative differences around 30% for the amplitude, and around 32% for the pulse width.
Table 7.4: Relative differences for the amplitude and pulse width with several values of the dimensionality $m$ for the input space $x$, under anisotropic conditions

<table>
<thead>
<tr>
<th>Dimensionality $m$ of $x$</th>
<th>Amplitude (V)</th>
<th>Pulse width ($\mu$s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>890</td>
<td>$0.3086 \pm 0.1973$</td>
<td>$0.3184 \pm 0.2036$</td>
</tr>
<tr>
<td>1,780</td>
<td>$0.3035 \pm 0.2022$</td>
<td>$0.3202 \pm 0.2096$</td>
</tr>
<tr>
<td>3,559</td>
<td>$0.3036 \pm 0.2029$</td>
<td>$0.3203 \pm 0.2078$</td>
</tr>
<tr>
<td>17,793</td>
<td>$0.2991 \pm 0.2202$</td>
<td>$0.3150 \pm 0.2055$</td>
</tr>
<tr>
<td>177,924</td>
<td>$0.2985 \pm 0.2033$</td>
<td>$0.3143 \pm 0.2045$</td>
</tr>
</tbody>
</table>

7.3.2 Evaluation of performance of the proposed methodology to reconstruct the volume of tissue activated

As mentioned in section 6.2.1, the test set had 150 samples, but to reconstruct the VTA only 116 samples were used. This reduction can be justified because the estimation of the parameters is being done independently, that is, the test set creates 34 configurations where all contacts are inactive. During the DBS at least one contact must be active. Then, the neuromodulation parameters estimated were used to compute new VTAs. In order to evaluate the performance of the proposed methodology in terms of its ability to reconstruct the volume tissue activated, we employ an error metric that allows the comparison of the new and objective VTAs.

Figures 7.4 and 7.5 show the results obtained with the error metric used to assess the performance of the proposed methodology in terms of its ability to reconstruct the volume of tissue activated, either under isotropic and anisotropic conductivity conditions. We find that the proposed methodology has a high ability to reconstruct the volume tissue activated. Box-whisker plots present information independent about the different values of dimensionality of the input space used. In general terms the median of the data is similar in all considered spatial resolutions which agrees with the results obtained by the performance measures of SVMs. We can note slight differences between the results provide by the Figures 7.4 and 7.5, where the PMI gives better results in the case isotropic. The methodology achieves successful results in our first approach to estimate the neurostimulation parameters during DBS.

Figure 7.6 shows the different volumes of tissue activated for the dimensionalities studied, considering both conductivity conditions. It displays the methodology ability not only to reconstruct the volume of tissue activated, but also to represent the effects that tissue conductivity has on the VTA. The tissue conductivity affects the VTA via changes in the propagation of the stimulus waveform and consequently on axonal activation. The volumes of tissue activated when isotropic conductivity is assumed exhibit symmetry with respect to the electrode axis. Such symmetry disappears for the anisotropic case.
Figure 7.4: Box-whisker plots of the proposed methodology to reconstruct the volume of tissue activated, for the different values of dimensionality \( m \) of the input space \( x \), under isotropic conductivity conditions.

Figure 7.5: Box-whisker plots of the proposed methodology to reconstruct the volume of tissue activated, for the different values of dimensionality \( m \) of the input space \( x \), under anisotropic conductivity conditions.
Figure 7.6: Volumes of tissue activated obtained with the neuromodulation parameters estimated by the proposed methodology, for the different values of dimensionality \((A = 5.0 \, V, \, W = 60 \, \mu s, \, c_0 = -1, \, c_1 = -1, \, c_2 = -1, \, c_3 = -1)\), where (a) Test sample. (b) Estimated sample for \(m=177,924\). (c) Estimated sample for \(m=17,793\). (d) Estimated sample for \(m=3,559\). (e) Estimated sample for \(m=1,780\). (f) Estimated sample for \(m=890\), under isotropic conductivity conditions. The other volumes from (g) to (l) exhibit the same information, but under anisotropic conductivity conditions.
In this study, we attempt to solve the inverse problem of VTA, previously described in section 2, which has not been extensively studied in the literature. This is the reason why we employed a framework based on support vector machines. Based on this premise, we developed a novel methodology to address the problem of finding optimal neurostimulation parameters that increase the effectiveness of the DBS therapy, independently of whether isotropic or anisotropic conductivities are assumed. Our method, first builds a dataset of volumes of tissue activated with their corresponding neurostimulation parameters. This dataset is built with an accurate physiological model that combines an extracellular potential model, and a multicompartment myelinated axonal model. We then use this dataset to design a machine learning algorithm that maps from the space of volumes of tissue activated to the values of the neurostimulation parameters. Since the representation of each VTA is given as a long vector of zeros and ones, in this work, we use a kernel machine, a SVM, that allows us to handle spaces of high dimensionality. We obtain classification accuracies over 80% for predicting the states of the electrode contacts, and relative differences equal or lower than 30% for the amplitude and the pulse width of the rectangular pulse function. Additionally, the Figure 7.6 confirms the results obtained with error metric (PMI) used, in terms of the methodology ability to reconstruct the volume of tissue activated.

To the best of our knowledge, this is the first attempt in the literature for DBS, that looks to automatically tune the neurostimulation parameters from a previously specified VTA. Although the results shown in this study may be considered as preliminary, we think these results are promising, and leave plenty of room for further improvement. First, in this work we solved the classification problems and the regression problems independently. We hope that by solving simultaneously the classification and regression problems, we may improve the performance metrics. This is a foundational idea in successful frameworks like transfer learning or multi-task learning. Second, the choice of the kernel may also help to improve the results. We would like to experimentally test the performance under different distance or similarity measures (other than the Hamming distance), and under different and more sophisticated kernels that exploit the geometric structure of the VTA [43] [44]. Third, the experiments of this study show that it is possible to reduce the dimensionality of the input space without affecting the performance metrics. Reducing the dimensionality of the input space may ease the design of the machine learning model, improving the performance for classification and regression.
This section presents the additional mathematical details that allow us a better understanding of the algorithm, support vector machine, used to solve the inverse problem. We provide a mathematical explanation about SVM for classification and for single output regression.

9.1 SVM for classification

In support vector machines to classification problems can be presented two scenarios regarding the training data set, the first one assumes that the training data set is linearly separable in feature space, and second one assumes that the training data set is not linearly separable in feature space. In practice, conditional class distributions can overlap, in this case exact separation of training data can lead to a poor generalization. This difficulty is overcome allowing some data to be on the wrong side of margin boundary, but applying a penalty, which increases with distance from this boundary. Slack variables \( \xi_i \geq 0 \) \((i = 1, \ldots, m)\) are introduced for each training data, with \( \xi_i = 0 \) for data that are on or into the correct margin boundary and \( \xi_i = |y_i - f(x_i)| \) for other data [19][21].

The purpose is to maximize the margin while data misplaced are penalized, that is to say minimize

\[
\frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} \xi_i,
\]

where \( w \in \mathbb{R}^m \), \( b \in \mathbb{R} \), the constant \( C > 0 \) is the parameter that express the balance between penalization by slack variables and the margin. The problem constraints are given as

\[ y_if(x_i) \geq 1 - \xi_i, \quad i = 1, \ldots, N \quad \text{with} \quad \xi_i \geq 0. \]

Lagrange multipliers \( a_n \geq 0 \) and \( \mu_n \geq 0 \) are introduced to solve the optimization problem

\[
L(w, b, a) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} \xi_i - \sum_{i=1}^{N} a_i [y_if(x_i) - 1 + \xi_i] - \sum_{i=1}^{N} \mu_i \xi_i,
\]

where \( a_i = [a_1, \ldots, a_i]^{T} \) and calculating the derivatives \( L(w, b, a) \) with respect to \( w \), \( b \) and \( \xi_i \) equal to zero.
\[ \frac{\partial L}{\partial w} = 0 \quad \Rightarrow \quad w = \sum_{i=1}^{N} a_i y_i \phi(x), \]
\[ \frac{\partial L}{\partial b} = 0 \quad \Rightarrow \quad 0 = \sum_{i=1}^{N} a_i y_i, \]
\[ \frac{\partial L}{\partial \xi_i} = 0 \quad \Rightarrow \quad a_i = C - \mu_i. \]

The dual representation is

\[ \tilde{L}(a) = \sum_{i=1}^{N} a_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j y_i y_j k(x_i, x_j), \]

and the constraints are

\[ 0 \leq a_i \leq C; \quad \sum_{i=1}^{j} a_i y_i = 0. \]

The Karush Kuhn Tucker conditions that are satisfied by the solution are

\[ a_i \geq 0; \quad y_i f(x_i) - 1 + \xi_i \geq 0; \quad a_i (y_i f(x_i) - 1 + \xi_i) = 0, \]
\[ \mu_i \geq 0; \quad \xi_i \geq 0; \quad \mu_i \xi_i = 0 \quad \forall \quad i = 1, \ldots, m. \]

The classification for a new input data from training model is made by

\[ f(x) = \sum_{i=1}^{N} a_i y_i k(x, x_i') + b. \]

### 9.2 SVM for regression

SVM can be applied to regression problems by introducing a different loss function, which is more sensitive to outliers. For this reason, we using an \( \epsilon \)-insensitive loss function \[19\] \[45\] defined a

\[ L_\epsilon(y, \hat{y}) = \begin{cases} 
0, & \text{if } |y - \hat{y}| \leq \epsilon \\
|y - \hat{y}| - \epsilon, & \text{otherwise}
\end{cases} \]
Then, any datapoint that is inside $\epsilon$-tube around the prediction is not penalized. The objective function is

$$J = C \sum_{i=1}^{N} L_{\epsilon}(y_i, \hat{y}_i) + \frac{1}{2} \|w\|^2,$$

where $\hat{y}_i = f(x_i) = w^T \phi(x) + b$ and $C = \frac{1}{\lambda}$ is a regularization constant.

Applying a procedure similar to SVM based on classification, two slack variables $\xi_i \geq 0$ and $\hat{\xi}_i \geq 0$ are introduced, where $\xi_i \geq 0$ corresponds a data for which $y_i > f(x_i) + \epsilon$ and $\hat{\xi}_i \geq 0$ corresponds a data for which $y_i > f(x_i) - \epsilon$.

The condition for data to locate inside $\epsilon$-tube is

$$f(x_i) - \epsilon \leq y_i \leq f(x_i) + \epsilon.$$

Introducing the slack variables allows points to lie outside the $\epsilon$-tube provided the slack variables are nonzero, and the corresponding conditions are

$$y_i \leq f(x_i) + \epsilon + \xi_i; \quad y_i \geq f(x_i) - \epsilon - \hat{\xi}_i.$$

The error function for support vector regression is

$$J = C \sum_{i=1}^{N} (\xi_i + \hat{\xi}_i) + \frac{1}{2} \|w\|^2,$$

where it is minimized with constraints $\xi_i \geq 0; \hat{\xi}_i \geq 0; y_i \leq f(x_i) + \epsilon + \xi_i; y_i \geq f(x_i) - \epsilon - \hat{\xi}_i$.

To solve the optimization problem Lagrange multipliers $a_n \geq 0, \hat{a}_n \geq 0, \mu_n \geq 0$ and $\hat{\mu}_n \geq 0$ are introduced.

$$L = C \sum_{i=1}^{N} (\xi_i + \hat{\xi}_i) + \frac{1}{2} \|w\|^2 - \sum_{i=1}^{N} (\mu_i \xi_i + \hat{\mu}_i \hat{\xi}_i) - \sum_{i=1}^{N} a_i (\epsilon + \xi_i + f(x_i) - y_i) - \ldots$$

$$\sum_{i=1}^{N} \hat{a}_i (\epsilon - \hat{\xi}_i - f(x_i) + y_i).$$

Replacing $f(x_i)$ for the model previously described and calculating the derivatives $L(w, b, a)$ with respect to $w, b, \xi_i$ and $\hat{\xi}_i$ equal to zero

$$\frac{\partial L}{\partial w} = 0 \quad \Rightarrow \quad w = \sum_{i=1}^{N} (a_i - \hat{a}_i) \phi(x),$$

$$\frac{\partial L}{\partial b} = 0 \quad \Rightarrow \quad 0 = \sum_{i=1}^{N} (a_i - \hat{a}_i),$$

$$\frac{\partial L}{\partial \xi_i} = 0 \quad \Rightarrow \quad a_i + \mu_i = C,$$

$$\frac{\partial L}{\partial \hat{\xi}_i} = 0 \quad \Rightarrow \quad \hat{a}_i + \hat{\mu}_i = C.$$
The dual representation is

\[
\tilde{L}(a_i - \hat{a}_i) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (a_i - \hat{a}_i)(a_j - \hat{a}_j)k(x_i, x_j) - \epsilon \sum_{i=1}^{N} (a_i + \hat{a}_i) + \sum_{i=1}^{N} (a_i - \hat{a}_i)y_i,
\]

and the constraints are

\[
a_i + \mu_i = C; \quad \hat{a}_i + \hat{\mu}_i = C; \quad 0 \leq a_i \leq C \quad 0 \leq \hat{a}_i \leq C \quad \sum_{i=1}^{N} (a_i + \hat{a}_i) = 0.
\]

The Karush-Kuhn-Tucker conditions that are satisfied by the solution are

\[
\begin{align*}
    a_i(\epsilon + \xi_i + f(x_i) - y_i) & = 0; \quad (C - a_i)\xi_i = 0, \\
    \hat{a}_i(\epsilon + \hat{\xi}_i - f(x_i) + y_i) & = 0; \quad (C - \hat{a}_i)\hat{\xi}_i = 0.
\end{align*}
\]

The estimation for a new input data is made by

\[
f(x) = \sum_{i=1}^{N} (a_i, \hat{a}_i)k(x, x'_i) + b.
\]
Publications

Refereed journal papers


Refereed conference papers


[40] Fernando Pérez Cruz. Multioutput SVR [msvr].


