Classification with Support Vector Machines, New Quadratic Programming Algorithm

[1]A.Chikhaoui, [2]A. Mokhtari,
[1]University of Tiaret Algeria, [2]University of Laghouat Algeria,
ah_chikhaoui@yahoo.fr, Aek_Mokhtari@yahoo.fr

Abstract-*Support vector machines (SVM) are excellent tools for classification and regression. They seek the optimal separating hyperplan and maximal margin. The modeling results often lead to solving a quadratic programming problem. In this paper, we present a simple method to determine the hyperplan H that separates two classes of examples so that the distance between these two classes is maximal. This method is based on the geometric interpretation of the norm of a linear mapping. The result model of our algorithm modeling is a maximization of a concave quadratic program. This quadratic program is resolved by projection method. Example illustrates the method.*

Keywords
Support vector machines, separating hyperplan, maximizing concave function, cosine, projection method.

1 Introduction

Learning to rank is an important problem in web page ranking information retrieval and other applications. Support Vector Machines (SVMs) are a powerful machine learning technique. Vapnik [7] showed how training a support vector machine for the pattern recognition problem leads to quadratic optimization problem (QP). The size of the optimization problem depends on the number of training examples. With 10000 training examples and more it becomes impossible to keep matrix data in memory. SVM$^{\text{Light}}$ uses the decomposition idea of Osuna and al. ([7]) and decompose the problem into a series of smaller tasks. This decomposition splits the initial problem in an inactive and en active part. These algorithms may need a long training time. To tackle this problem, T. Joachims [5], uses a method for selecting the working set, successive “shrinking” of the optimization problem and incremental updates of the gradient (Joachims [6]). Burges form AT&T [1], has even developed a QP solver specifically for training SVM.

In this paper we introduce new support vector machines method in order to define a decision surface separating two opposing classes of a training set of vectors.

This method associates a distance parameter with each vector of the SVM’s training set. The distance parameter is calculating as the shortest of distances from each vector of one class to the opposite class. The method determines initial separating hyperplan and its maximum margin, where the margin is defined as the shortest distances of the hyperplan from the closest points of the two classes. The optimal vectors to preselect as potential support vectors are those closest to the decision hyperplan. The vectors with the smallest distance are then selected as pivots.

To determine the optimal hyperplan, we will use the well-known result:

if f is a linear map from R^n into R defined by $f(x)=\langle a,x\rangle,\ a\in R^n$.

Then $\|a\|=d(0,H)$ where H is the hyperplan defined by $H=\{x\in R^n : \langle a,x\rangle=1\}$.

The optimal hyperplan will be a boundary point of the set of feasible solutions which can be an extreme point.

2 Partition of examples $\tilde{X}_+$ and $\tilde{X}_-$

Suppose that separating hyperplan with maximum margin be written as $ax+b=0$.

2.1 Formulation of the optimization problem

The inequalities $ax+b\geq 1$ and $ax+b\leq 1$ become $\frac{a}{2}x_+ + \frac{b}{2}\geq\frac{1}{2}$ and $\frac{a}{2}x_- + \frac{b}{2}\leq -\frac{1}{2}$, and the hyperplan is $\frac{a}{2}x+\frac{b}{2}=\frac{1}{2}$; i.e. $ax+b=0$. As the couple (a,b) is set to a multiplicative coefficient, the separating problem becomes then

$$\begin{cases} \inf \|a\|^2 \\ ax_+ + b \geq \frac{1}{2} \\ ax_- + b \leq -\frac{1}{2} \end{cases}$$

Suppose that $\bar{x}_+$ is support vector, $a\bar{x}_+ + b = \frac{1}{2}$

$\Rightarrow\ b = \frac{1}{2} - a\bar{x}_+$.

Then $ax_+ + b \geq \frac{1}{2} \Leftrightarrow$

$ax_+ + \frac{1}{2} - a\bar{x}_+ \geq \frac{1}{2} \ \Leftrightarrow\ a(x_+ - \bar{x}_+) \geq 0$

$\Leftrightarrow\ a(\bar{x}_+ - x_+) \leq 0$ and

$ax_- + b \leq -\frac{1}{2} \Leftrightarrow ax_- + \frac{1}{2} - a\bar{x}_+ \leq -\frac{1}{2} \Leftrightarrow a(x_- - \bar{x}_+) \leq -1$.

Then
$$\begin{cases} \inf \|a\|^2 \\ a(\bar{x}_+ - x_+) \leq 0 \\ a(x_- - \bar{x}_+) \leq -1 \end{cases} = \begin{cases} -Max\{-\|a\|^2\} \\ a(\bar{x}_+ - x_+) \leq 0 \\ a(x_- - \bar{x}_+) \leq -1 \end{cases}$$

and consequently, the separating problem becomes

$$(P) = \begin{cases} Max\{-\|a\|^2\} \\ \Omega = \begin{Bmatrix} a(\bar{x}_+ - x_+) \leq 0, & x_+ \in X_+ \\ a(x_- - \bar{x}_+) \leq -1, & x_- \in X_- \end{Bmatrix} \end{cases},$$

$f(a) = \sum_{i=1}^{n} -a_i^2 = -\|a\|^2$ is concave, defined on closed bounded convex of R^n, then the local maximum is global, but $\frac{\partial f}{\partial a_i}(a) = 0$ for all i,

$\Rightarrow 2a_i = 0, \quad a_i = a_i^* = 0$.

The critical point $a^* = 0 \in R^n$ is not feasible solution, then the solution of the problem is the projection of $a^* = 0$ on Ω . This is a particular case of general optimization problem of concave quadratic programming, where $\alpha_i = 0, \beta_i = 1, a^* = -\frac{\alpha_i}{2\beta_i} = 0$.

This problem of maximizing concave quadratic function under linear constraints has solved by Chikhaoui and all. [3]. It is noted that $P_{a(\bar{x}_+ - x_+)}(0) = 0.$

This was made possible through the form $\begin{cases} \frac{a}{2}x_+ + \frac{b}{2} \geq \frac{1}{2} \\ \frac{a}{2}x_- + \frac{b}{2} \leq -\frac{1}{2} \end{cases}$, this minimize the computing time.

Increase in a margin.

Let (H) the separating hyperplan of wide margin of equation $ax + b = 0$.

We know that for all $x_+ \in X_+$, $x_- \in X_-$, we have
$$\begin{cases} \frac{a}{2}x_+ + \frac{b}{2} \geq \frac{1}{2} \\ \frac{a}{2}x_- + \frac{b}{2} \leq -\frac{1}{2} \end{cases}$$

$$\Rightarrow \begin{cases} ax_+ + b \geq \frac{1}{2} \\ -ax_- - b \geq +\frac{1}{2} \end{cases} \Rightarrow (ax_+ + b) + (-ax_- - b) \geq \frac{1}{2} + \frac{1}{2}$$

; from where

$a(x_+ - x_-) \geq 1, \ \forall x_+ \in X_+, \ \forall x_- \in X_-$.

By the inequality of Cauchy Schwartz, $\frac{1}{\|a\|} \leq \|x_+ - x_-\|$, $\quad \forall x_+ \in X_+, \forall x_- \in X_-$,

by passing to the lower bound, we obtain $\frac{1}{\|a\|} \leq \inf_{x_+ \in X_+, x_- \in X_-} \|x_+ - x_-\|$.

Whence important proposal:

Let $\tilde{H}_+ = \{x \in R^n : ax - a\bar{x}_+ = 0\}$, $\tilde{H}_- = \{x \in R^n : ax - a\bar{x}_- = 0\}$.

Proposal:

The width of the strip is increased by the constant $K = \inf_{x_+ \in X_+, x_- \in X_-} \|x_+ - x_-\|$, and this is best constant.

Proof: Indeed, suppose x_+, x_- are two support vectors, i.e. $x_+ \in \tilde{H}_+, x_- \in \tilde{H}_-$;

$d(x_+, \tilde{H}_+) = \frac{ax_+ + b}{\|a\|} = \frac{1}{2}.\frac{1}{\|a\|}$,

$d(x_-, \tilde{H}_-) = \frac{|ax_- + b|}{\|a\|} = \frac{1}{2}.\frac{1}{\|a\|}$

$\Rightarrow d\left(x_+, \tilde{H}_+\right) + d\left(x_-, \tilde{H}_-\right) = \frac{1}{\|a\|}$, and $\frac{1}{\|a\|} \le \inf_{x_+ \in X_+, x_- \in X_-} \|x_+ - x_-\| = K$.

This is the best ever because in cases where $\tilde{X}_+ = \{x_+\}$ and $\tilde{X}_- = \{x_-\}$, then $\frac{1}{\|a\|} = \inf_{x_+ \in X_+, x_- \in X_-} \|x_+ - x_-\| = K$; this completes the proof.

We see that the margin width does not exceed $\|x_+ - x_-\|$.

This leads us to consider the separating hyperplan with the widest possible margin $\tilde{H}$.

2.2 Partition of X_+ and X_-

Let $\inf_{x_+ \in X_+, x_- \in X_-} \|x_+ - x_-\| = \|\bar{x}_+ - \bar{x}_-\|$,

$\tilde{a} = \bar{x}_+ - \bar{x}_-$,

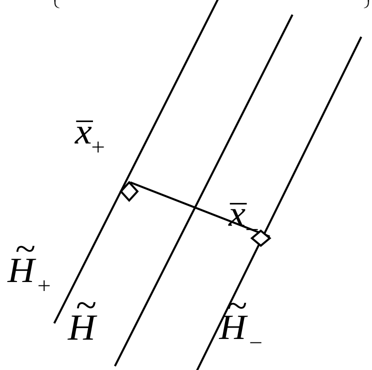
$\tilde{H}_+ = \left\{x \in R^n : \tilde{a}x - \tilde{a}\bar{x}_+ = 0\right\}$,

$\tilde{H}_- = \left\{x \in R^n : \tilde{a}x - \tilde{a}\bar{x}_- = 0\right\}$,

$$\tilde{H} = \left\{x \in R^n : \tilde{a}x - \left(\frac{\tilde{a}\bar{x}_+ + \tilde{a}\bar{x}_-}{2}\right) = 0\right\}$$

$$= \left\{x \in R^n : (\bar{x}_+ - \bar{x}_-)x - \frac{(\bar{x}_+ - \bar{x}_-)(\bar{x}_+ + \bar{x}_-)}{2} = 0\right\}$$

$$= \left\{x \in R^n : (\bar{x}_+ - \bar{x}_-)x + \frac{\|\bar{x}_-\|^2 - \|\bar{x}_+\|^2}{2} = 0\right\}$$



The existence of optimal separating hyperplan H , and construction of $\tilde{H}$ define a partition of X_+ and a partition of X_- :

$$\tilde{X}_+ = \left\{x \in X_+ : ax - a\bar{x}_+ < \frac{1}{2}\right\},$$

$X_+ = \tilde{X}_+ \cup \left(X_+ / \tilde{X}_+\right)$

$$\tilde{X}_- = \left\{x \in X_- : ax - a\bar{x}_- > -\frac{1}{2}\right\},$$

$X_- = \tilde{X}_- \cup \left(X_- / \tilde{X}_-\right)$

If the hyperplans H and $\tilde{H}$ separate the sets $X_+ / \tilde{X}_+$ and $X_- / \tilde{X}_-$ and $\tilde{H}$ is optimal. $\tilde{X}_+ = \tilde{X}_- = \phi$ then $H = \tilde{H}$. Stop.

3 Finding Optimal separating hyperplan H: case $H \neq \tilde{H}$

The maximum margin separating $\tilde{X}_+$ and $\tilde{X}_-$ is greater than or equal to the maximum margin separating X_+ and X_- because the optimal separating hyperplan H separates $\tilde{X}_+$ and $\tilde{X}_-$.

Suppose that the maximum margin between $\tilde{X}_+$ and $\tilde{X}_-$ is strictly greater than that between X_+ and X_-, then this separating hyperplan is between H and $\tilde{H}$ and hence it separates $X_+ / \tilde{X}_+$ and $X_- / \tilde{X}_-$. So this separating hyperplan separates X_+ and X_- with a wider margin than strictly greater than that to H . Contradiction, because H is optimal.

Proposition:

The optimal separating hyperplan of sets $\tilde{X}_+$ and $\tilde{X}_-$ is optimal separating hyperplan for sets X_+ and X_- .

Proof: Denote by H^* optimal separating hyperplan of $\tilde{X}_+$ and $\tilde{X}_-$ whose normal is a^* . There positive λ_1! and λ_2! such that $a^* = \lambda_1 a + \lambda_2 \tilde{a}$.

In fact, $\begin{cases} \langle a, a^* \rangle = \|a\| . \|a^*\| \cos\alpha \\ \langle \tilde{a}, a^* \rangle = \|\tilde{a}\| . \|a^*\| \cos\beta \end{cases}$

$\langle a, \lambda_1 a + \lambda_2 \tilde{a} \rangle = \|a\| . \|a^*\| \cos\alpha$

$\Rightarrow \lambda_1 \langle a, a \rangle + \lambda_2 \langle a, \tilde{a} \rangle = \|a\| . \|a^*\| \cos\alpha$

$\Rightarrow \lambda_1 \|a\|^2 + \lambda_2 \langle a, \tilde{a} \rangle = \|a\| . \|a^*\| \cos\alpha$.

As well

$\langle \tilde{a}, \lambda_1 a + \lambda_2 \tilde{a} \rangle = \|\tilde{a}\| . \|a^*\| \cos\beta$

$\Rightarrow \lambda_1 \langle \tilde{a}, a \rangle + \lambda_2 \langle \tilde{a}, \tilde{a} \rangle = \|\tilde{a}\| . \|a^*\| \cos\beta$

$\Rightarrow \lambda_1\langle\tilde{a},a\rangle + \lambda_2\|\tilde{a}\|^2 = \|\tilde{a}\|\|a^*\|\cos\beta$. Then

$$\begin{cases} \lambda_1\|a\|^2 + \lambda_2\langle a,\tilde{a}\rangle = \|a\|.\|a^*\|\cos\alpha \\ \lambda_1\langle\tilde{a},a\rangle + \lambda_2\|\tilde{a}\|^2 = \|\tilde{a}\|.\|a^*\|\cos\beta \end{cases}$$

Where θ is the angle formed between hyperplans H and $\tilde{H}$, as $H \neq \tilde{H}$, $\theta \neq 0$;
then $\cos\theta \neq 1$, and $\Delta = \|a\|^2\|\tilde{a}\|^2(1-\cos^2\theta)$. $\Delta > 0$

The system has a unique solution λ_1 and λ_2.

Then suppose that the margin of H^* is strictly greater than that of H, as H and $\tilde{H}$ separate $X_+/\tilde{X}_+$ and $X_-/\tilde{X}_-$. i.e. a and $\tilde{a}$ are solution of problem

$$\begin{cases} \underset{\Omega'}{Max}\left\{-\|a\|^2\right\} \\ \Omega' = \begin{cases} a(-x_+ + \bar{x}_+) \leq 0 \\ a(x_- - \bar{x}_+) \leq -1 \\ \bar{x}_+ \in X_+/\tilde{X}_+ \\ \bar{x}_- \in X_-/\tilde{X}_- \end{cases} \end{cases}$$

Ω' is bounded below convex, then $a^* \in \Omega'$.

The separating hyperplan H^* separates then X_+ and X_- whose margin is strictly greater than that of H. Contradiction, H is optimal by hypothesis.

Consequence:

To separate X_+ and X_-, just separate the sample $\tilde{X}_+$ and $\tilde{X}_-$. We then have a smaller number of constraints.

Example1.

$$X_+ = \left\{(1,0),\left(2,\frac{3}{2}\right),(5,1),\right\}$$

$$X_- = \{(0,1),(-2,1),(-2,2),\}$$

Here,

$$\inf_{\substack{x_+\in X_+\\ x_-\in X_-}}\left(\|x_+ - x_-\|\right) = \|(1,0)-(0,1)\| = \sqrt{2};$$

$x_+ = (1,0)$, $x_- = (0,1)$

$\tilde{a} = (x_+ - x_-) = (1,-1)$

$\tilde{H}_+ : (1,-1)x - (1,-1)(0,0) = 0$

$\tilde{H}_+ : x_1 - x_2 - 1 = 0$

$\tilde{H}_- : x_1 - x_2 + 1 = 0$ $\tilde{H} : x_1 - x_2 = 0$.

$\left(2,\frac{3}{2}\right) \in \tilde{X}_+$ car $2 - \frac{3}{2} - 1 = -\frac{1}{2} < \frac{1}{2}$,

$(5,1) \notin \tilde{X}_+$ car $5 - 1 - 1 = 3 > \frac{1}{2}$

$$\Rightarrow \tilde{X}_+ = \left\{(1,0),\left(2,\frac{3}{2}\right)\right\}$$

$(-2,1) \notin \tilde{X}_-$ because $-2-1+1 = -2 < -\frac{1}{2}$,

$(-2,2) \notin \tilde{X}_-$ because $-2-2+1 = 3 < -\frac{1}{2}$

$\Rightarrow \tilde{X}_- = \{(0,1)\}$.

The constraint set Ω, of problem becomes

$$\begin{cases} a\left((1,0)-\left(2,\frac{3}{2}\right)\right) \leq 0 \\ a((0,1)-(1,0)) \leq -1 \end{cases}$$

the solution is

$$\begin{cases} 2a_{11} + 3a_{12} = 0 \\ a_{11} + a_{12} + 1 = 0 \end{cases} \Rightarrow \begin{cases} a_{11} = \frac{3}{5} \\ a_{12} = -\frac{2}{5} \end{cases}. \qquad (H)$$

has the equation

$$\left(\frac{3}{5},-\frac{2}{5}\right)(x_1,x_2) + \left(\frac{1}{2} - \left(\frac{3}{5},-\frac{2}{5}\right)\right)(1,0) = 0,$$

$$H: \quad \frac{3}{5}x_1 - \frac{2}{5}x_2 - \frac{1}{10} = 0.$$

4 Projection Method ([2])

Consider the problem quadratic result of our

$$\Delta = \begin{vmatrix} \|a\|^2 & \langle a,\tilde{a}\rangle \\ \langle\tilde{a},a\rangle & \|\tilde{a}\|^2 \end{vmatrix} = \|a\|^2.\|\tilde{a}\|^2 - \langle a,\tilde{a}\rangle\langle\tilde{a},a\rangle = \|a\|^2\|\tilde{a}\|^2 - \|a\|^2\|\tilde{a}\|^2\cos^2\theta$$

modeling:

$$(P^{'})=\begin{cases} \underset{\Omega}{Max}\left\{-\|a\|^2\right\} \\ \Omega=\left\{a\in R^n,\ \begin{pmatrix} a(-x_+ + \bar{x}_+)\le 0 \\ a(x_- - \bar{x}_+)\ \le -1\end{pmatrix},\ x_+\in \tilde{X}_+,\ x_-\in\tilde{X}_-\right\}\end{cases}$$

Since the function $f(x)=\sum_{i=1}^{n}-a_i^2=-\|a\|^2$ is concave defined on a closed convex of R^n, then the local maximum of f is global. But $\frac{\partial f}{\partial a_i}(a)=0$,

$\Leftrightarrow\ a_i=0\quad i=1,2,...,n$

Critical point $(a_i^*)_i=0$ *for all* $i=1,2,...,n$ is not feasible solution of problem $(P^{'})$.

Then the solution of problem $(P^{'})$ is the projection of point $0\in R^n$ on Ω.

This is a particular case of more general problem of quadratic optimization:

$f(a)=\sum_{i=1}^{n}\alpha_i a_i+\sum_{i=1}^{n}\beta_i a_i^2$ *with* $\alpha_i\in R^n, \beta_i<0$ *for all* i ,

under linear constraints.

In classification with SVM we have $a^*=\left(-\frac{\alpha_i}{2\beta_i}\right)_i$ with $\alpha_i=0,\ \forall i\ \ \beta_i=-1.$ For more details see [3]. We recall that if a concave function f defined on closed convex and that the critical point does not belong to convex, then the maximum of f is reached on a boundary point of closed convex. See [3].

The projection of point $0\in R^n$ on the hyperplan $a(x_- - \bar{x}_+)\ =-1$ is given by

$$P_{a(x_- - \bar{x}_+)=-1}(0)=0-\frac{1}{\|(x_- - \bar{x}_+)\|^2}(x_- - \bar{x}_+).$$

Example1

$X_+=\{(1,3),(1.5,4),(2,3),(3,3.5),(3,4)\}$

$X_-=\{(1,1.5),(1.5,1),(2,1),(2,2),(2.5,1.5)\}$.

$\inf_{\substack{x_+\in X_+\\ x_-\in X_-}}(\|x_+-x_-\|)=\|(2,3)-(2,2)\|=1$;

$x_+=(2,3),\quad x_-=(2,2)$

$$\tilde{a}=(x_+-x_-)=(2,3)-(2,2)=(0,1)$$

$$\tilde{H}_+:(0,1)(x_1,x_2)+\left(\frac{1}{2}-(0,1)(2,3)\right)=\frac{1}{2}$$

$$\Rightarrow \tilde{H}_+:\ x_2-3=0$$

$$\tilde{H}_-:(0,1)(x_1,x_2)+\left(-\frac{1}{2}-(0,1)(2,2)\right)=-\frac{1}{2}$$

$$\Rightarrow \tilde{H}_-:\ x_2-2=0$$

$$\tilde{H}:(0,1)(x_1,x_2)+\left(\frac{-3-2}{2}\right)=0$$

$$\Rightarrow \tilde{H}:\ x_2-\frac{5}{2}=0$$

Construction of $\tilde{X}_+$ and $\tilde{X}_-$

$(1,3)\in\tilde{X}_+$ because $3-3\le 0$

$(1.5,4)\notin\tilde{X}_+$ because $4-3\not\le 0$

$(2,3)\in\tilde{X}_+$ because $3-3\le 0$

$(3,3.5)\notin\tilde{X}_+$ because $3.5-3\not\le 0$

$(3,4)\notin\tilde{X}_+$ because $4-3\not\le 0$

$(1,1.5)\notin\tilde{X}_-$ because $1.5-2\not\ge 0$

$(1.5,4)\notin\tilde{X}_+$ because $4-3\not\ge 0$

$(2,3)\in\tilde{X}_+$ because $3-3\le 0$

$(3,3.5)\notin\tilde{X}_+$ because $3.5-3\not\ge 0$

$(3,4)\notin\tilde{X}_+$ because $4-3\not\ge 0$

Then $\tilde{X}_+=\{(2,3),(1,3)\}$ $\quad\tilde{X}_-=\{(2,2)\}$

Constraints are:

$$\begin{cases} a((2,2)-(2,3))\le -1 & \Leftrightarrow\ a_2\ge 1 \\ a((2,1)-(2,3))\le -1 & \Leftrightarrow\ a_2\ge \frac{1}{2}\end{cases}$$

$$P_{a_2=1}(0)=0-\frac{-1}{1}(0,1)=(0,1)\qquad\Rightarrow\|P_{a_2=1}(0)\|=1$$

$$P_{a_2=\frac{1}{2}}(0)=0-\frac{-\frac{1}{2}}{1}(0,1)=\left(0,\frac{1}{2}\right)\quad\Rightarrow\left\|P_{a_2=\frac{1}{2}}(0)\right\|=\frac{1}{4}$$

The solution is $a=(0,1)$, because $a=\left(0,\frac{1}{2}\right)$ is not feasible solution, and the optimal separating hyperplan is:

$$H:(0,1)(x_1,x_2)+\left(\frac{1}{2}-(0,1)(2,3)\right)=\frac{1}{2}$$

$$\Rightarrow x_2-\frac{5}{2}=0.$$

Remark: The feasible solution set of separating hyperplans is the half-space $a_2 \geq 1$ and the projection of 0 on this half-space is $(0,1)$. The set of feasible solutions do here no extreme point. It is interesting to study the nature of the set of separating hyperplans.

Example 2:

$$X_+=\left\{\left(2,2,\frac{1}{2}\right),(2,3,2),(3,3,1)\right\}$$

$$X_-=\left\{\left(1,0,\frac{1}{2}\right),(1,-1,3)\right\}$$

$$\inf_{\substack{x_+\in X_+\\ x_-\in X_-}}\|x_+-x_-\|=\|(2,2,0.5)-(1,0,0.5)\|=\sqrt{5}$$

$$x_+=(2,2,0.5),\quad x_-=(1,0,0.5)$$

$$\tilde{a}=(x_+-x_-)=(1,2,0)$$

$$\tilde{H}:(1,2,0)(x_1,x_2,x_3)-\frac{3}{2}=0$$

$\tilde{X}_+=\left\{\left(2,2,\frac{1}{2}\right)\right\}$ $\tilde{X}_-=\left\{\left(1,0,\frac{1}{2}\right)\right\}$. Here, inside of band is empty. So $(\tilde{H})=(H)$.

5 Conclusion

In this paper, we gave a geometric interpretation of the hyperplan that separates two classes linearly separable. In fact, the search algorithm to the optimum is nothing other than a particular case of general optimization problem:

$$\begin{cases}\sum_{i=1}^{n}\alpha_i x_i+\beta_i x_i^2, \quad with\ \ \alpha_i=0,\ \beta_i=-1\\ Ax\leq b\end{cases}$$

The nature of solution (extreme point or not) provides to better track the support vectors.

6 References

1. Burges C. J.C. 1998. A Tutorial on Support Vector machines for Pattern Recognition.
Data Mining and Knowledge Discovery, 2(2): 121-167.
2. Chikhaoui A, Djebbar B., Bellabacci A, Mokhtari A., 2009.Optimization of a quadratic
function under its Canonical form *Asian Journal of Applied Sciences* 2(6):499-510
3. Chikhaoui A, Djebbar B., and Mekki R., 2010. New Method for Finding an Optimal Solu
tion to Quadratic Programming Problem, *Journal of Applied Sciences* 10(15):1627- 1631-2010. ISSN 1812-5654.
4. Chikhaoui A, Bellabacci A, Mokhtari A.,2012. New Algorithm for maximizing of linearly
constrained convex quadratic programming. IRECOS Vol 7, ISSN 1828-6003.
5. Joachims T.,2003, Making Large-Scale SVM Learning practical, Advances in kernel
methods-support learning. Page 169-184. MIT Press.
6. Joachims, T., Granka, L., Pan, Bing, Hembrooke, H.,Radlinski, F., and Gay, G. Evaluating
the accuracy of implicit feedback from clicks and query reformulations in web search.
ACM Transactions on Information Systems (TOIS), 25(2), April 2007.
7. Osuna, E., Freund, R. and Girosi, F. An improved training algorithm for support vector machines. In Proceedings of the 1997 IEEE Workshop on Neural Networks for Signal Processing, Eds. J. Principe, L. Giles, N. Morgan,
8. Vapnik, V.N., 1998, Statistical learning theory. John Wiley & Sons.
9. N.M., Deris S. and Chin K.K., 2003, A comparaison of support vector machines training. Jurnal Teknologi, Vol 39, pp 45-56.
10. Wilson E., pages 276 – 285, Amelia Island, FL, 1997.