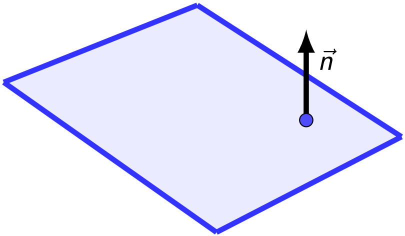
Support Vector Machine

using python.

Support vector Machine

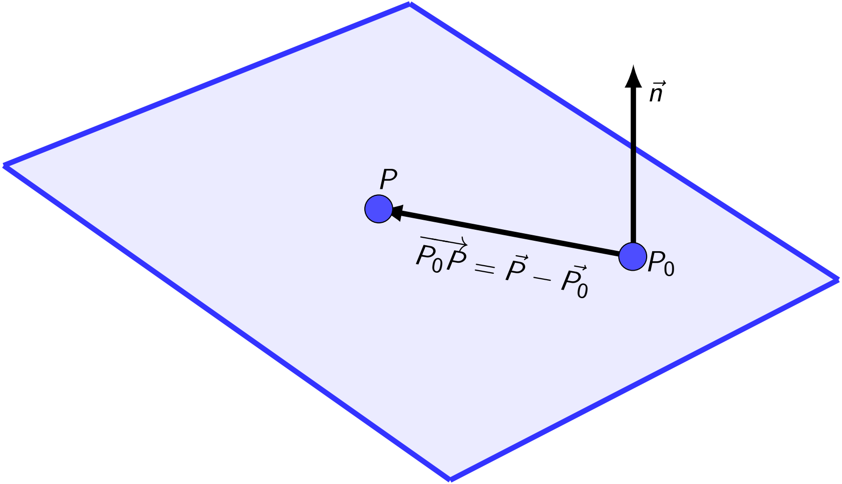
Support Vector Machine (SVM) is a supervised algorithm.

A plane



It is important concept, we are going to said that $\vec{n}$ is a normal(ortogonal) vector.

Plane



Vectorial equation of plane

$$\overrightarrow{P_0P}.\vec{n} = 0 \tag{1}$$

It is important notice, that for any vector (P) over the plane, the previous equation is satisfied.

Genaral equation of a plane

Assume that $\vec{P} = (x, y, z)$, $\vec{P_0} = (x_0, y_0, z_0)$ and $\vec{n} = (w_0, w_1, w_2)$

$$\begin{aligned} \overrightarrow{P_0P}.\vec{n} &= 0 \\ (x - x_0, y - y_0, z - z_0).\vec{n} &= 0 \\ w_0(x - x_0) + w_1(y - y_0) + w_2(z - z_0) &= 0 \end{aligned} \tag{2}$$

given that we known -$w_0x_0 - w_1y_0 - w_2z_0$, we can rewrite as:

$$\begin{aligned} w_0x + w_1y + w_2z &= w_0x_0 + w_1y_0 + w_2z_0 \\ w_0x + w_1y + w_2z &= b \end{aligned} \tag{3}$$

Genaral equation of a plane

If we assume that $\vec{w} = (w_0, w_1, w_2)$ and $\vec{x} = (x, y, z)$ we can rewrite as (in terms of the most papers):

$$\vec{w}.\vec{x} = b \tag{4}$$

or its equivalent:

$$\vec{w}.\vec{x} - b = 0 \tag{5}$$

From the last equation note that b is a scalar.

$\vec{w}$ perpendicular to plane π

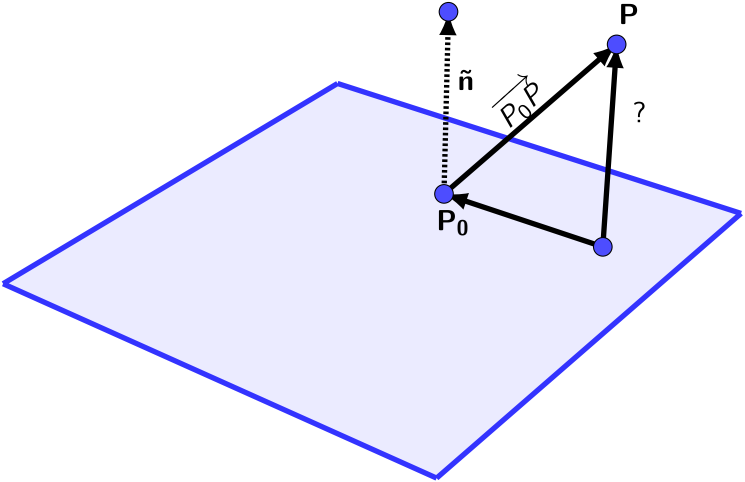
The plane π described by $\vec{w}.\vec{x} = b$, if $\vec{n}$ is a normal vector of π then its components are the same of $\vec{w}$.
Assume that if the point $P_i \in \pi$ and is described by (x_i, y_i, z_i) then:

$$w_0 x_i + w_1 y_i + w_2 z_i = b \tag{6}$$

for **any** two points P_a, P_b is equivalent to say that **any vector** $\overrightarrow{P_a P_b}$ is perpendicular to π if $\overrightarrow{P_a P_b}.\vec{n} = 0$, assuming that $\vec{n} = \vec{w}$ we have

$$(\vec{P_b} - \vec{P_a}).\vec{w} = \vec{P_b}.\vec{w} - \vec{P_a}.\vec{w} = b - b = 0. \tag{7}$$

Distance from point to plane

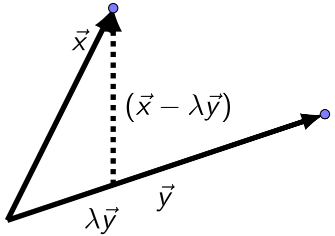


Distance from point to plane

The previous problem is restricted to find only the projection of the vector $\overrightarrow{P_0P}$ over the normal vector $\vec{n}$, remember that the last vector is perpendicualar to any point in the plane.

Projection

There is a λ scalar such that the vectors $\vec{x}, \vec{y}$ be orthogonal (in this case is the projection of $\vec{x}$ over $\vec{y}$).



Remember the orthogonality imply that:

$$(\vec{x} - \lambda\vec{y}).\vec{y} = 0 \tag{8}$$

$$\vec{x}.\vec{y} - \lambda\vec{y}.\vec{y} = 0 \tag{9}$$

$$\lambda = \frac{\vec{x}.\vec{y}}{\|\vec{y}\|^2} \tag{10}$$

This result is very handy, and we said that we poject the vector $\vec{x}$ over $\vec{y}$:

$$project_{\vec{y}}\vec{x} = \left(\frac{\vec{x}.\vec{y}}{\|\vec{y}\|^2}\right)\vec{y} \tag{11}$$

in PCA this operation is fundamental.

Distance from a point to hyperplane

Our problem is project $\overrightarrow{P_0P}$ over $\vec{n}$ and get projection magnitude:

$$project_{\vec{n}}\overrightarrow{P_0P} = \left(\frac{\overrightarrow{P_0P}.\vec{n}}{\|\vec{n}\|^2}\right)\vec{n} \tag{12}$$

Now the magintude of the vector will be:

$$\begin{aligned} \|project_{\vec{n}}\overrightarrow{P_0P}\| &= \left\| \left(\frac{\overrightarrow{P_0P}.\vec{n}}{\|\vec{n}\|^2}\right)\vec{n}\right\| \\ &= \frac{|\overrightarrow{P_0P}.\vec{n}|}{\|\vec{n}\|^2}\|\vec{n}\| = \frac{|\overrightarrow{P_0P}.\vec{n}|}{\|\vec{n}\|}. \end{aligned} \tag{13}$$

Distance from point to plane

Therefore the distance d from a point to plane will be:

$$d = \frac{|\overrightarrow{P_0 P}.\vec{n}|}{\|\vec{n}\|} \tag{14}$$

Assume that $P = (x, y, z)$ and $P_0 = (x_0, y_0, z_0)$ thus

$$d = \frac{|(P - P_0).\vec{n}|}{\|\vec{n}\|} \tag{15}$$

$$d = \frac{|w_0(x - x_0) + w_1(y - y_0) + w_2(z - z_0)|}{\sqrt{w_0^2 + w_1^2 + w_2^2}} \tag{16}$$

Distance from point to plane

$$d = \frac{|w_0x + w_1y + w_2z - w_0x_0 - w_1y_0 - w_2z_0|}{\sqrt{w_0^2 + w_1^2 + w_2^2}} \tag{17}$$

we know $-w_0x_0 - w_1y_0 - w_2z_0$

$$d = \frac{|w_0x + w_1y + w_2z - b|}{\sqrt{w_0^2 + w_1^2 + w_2^2}} \tag{18}$$

for any P point describe by $\vec{P}$ we can rewrite

$$d = \frac{|\vec{w}.\vec{P} - b|}{\|\vec{w}\|} \tag{19}$$

Distance from point to plane

From origin

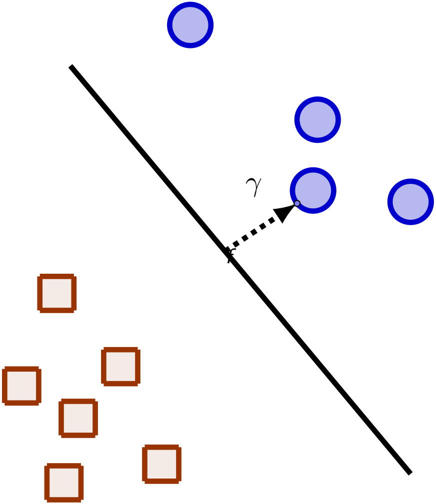
In **SVM** it is important the distance from the origin to the plane, $P = O = (0, 0, 0)$ therefore:

$$d = \frac{|\vec{w}.\vec{0} - b|}{\|\vec{w}\|} = \frac{|-b|}{\|\vec{w}\|} \tag{20}$$

Margin

Assume that margin(γ) will be the distance from the hyperplane to its closest point.

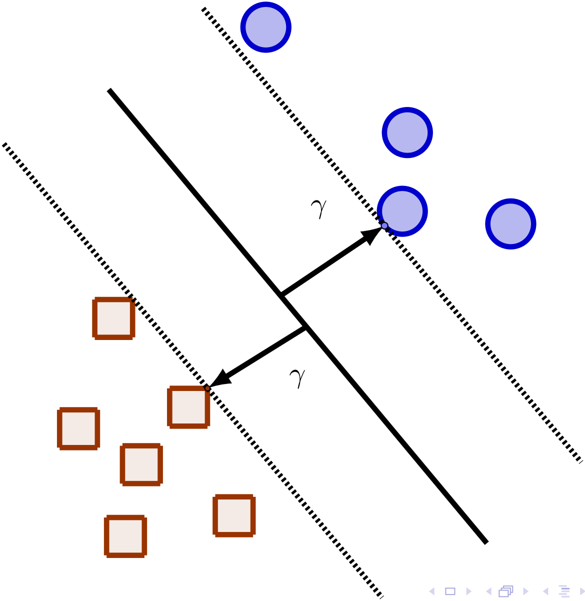
Margin



Margin

Now for **SVM** we are seeking get the maximun distance from the hyperplane to each nearest point in each class, namely we are seeking the maximun margin for the classes.

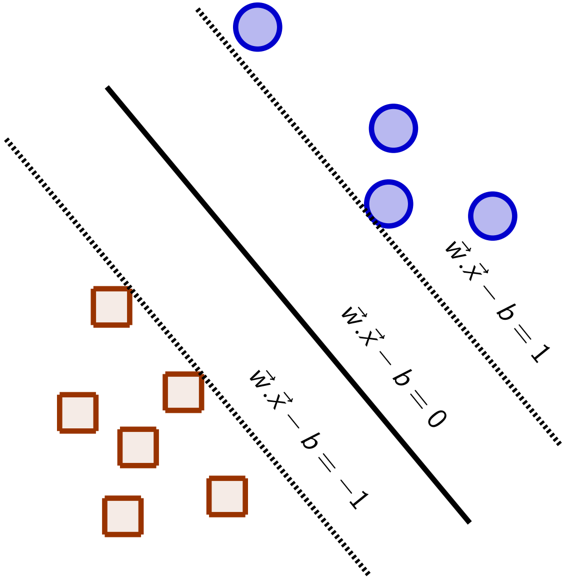
Maximun margin hyperplane



Codification

In this binary example, we are going to assume that the depenent variable y have two possible values $(-1, 1)$.

SVM



According to the previous figure, we need maximize the distance among the two planes that could be described as:

$$d(O, \vec{w}.\vec{x} - b - 1) - d(O, \vec{w}.\vec{x} - b + 1) \tag{21}$$

Now according to equation (15) this could be expressed as:

$$\frac{|-b-1|}{\|\vec{w}\|} - \frac{|-b+1|}{\|\vec{w}\|} = \frac{2}{\|\vec{w}\|} \tag{22}$$

Constrained optimization problem

The problem is

$$\underset{\vec{w}}{\text{maximize}} \quad \frac{2}{\|\vec{w}\|} \tag{23a}$$

$$\text{subject to}$$

$$\vec{w}.\vec{x} - b \geq 1 \text{ if } y_i = 1, \tag{23b}$$

$$\vec{w}.\vec{x} - b \leq -1 \text{ if } y_i = -1 \tag{23c}$$

Notice that $\max \frac{1}{f(x)} = \min f(x)$, and the two resctrictions could be described by the following:

$$y_i(\vec{w}.\vec{x} - b) - 1 \geq 0 \tag{24}$$

Constranied optimization problem

$$\underset{\vec{w}}{\text{minimize}} \quad \|\vec{w}\| \tag{25a}$$

$$\text{subject to}$$

$$y_i(\vec{w}.\vec{x} - b) - 1 \geq 0 \tag{25b}$$

The question is how solve this inequality constranied?

Important thigs to remark

The normal vector $\vec{n}$ of the hyperplane $\vec{w}.\vec{x} - b$ is $\nabla(\vec{w}.\vec{x} - b)$, is the gradient;
namely;

$$f(x, y, z) = w_0 x + w_1 y + w_2 z - b = 0 \tag{26}$$

$$\vec{n} = \vec{w} = \nabla f(x, y, z) \tag{27}$$

Material and references

- Francisco Calderon proffesor at PUJ "Máquinas de soporte Vectorial" Youtube Video
- Sergio monsalve Matemáticas básicas para economistas. Vol. 1. Álgebra lineal (Con notas históricas y contextos económicos)