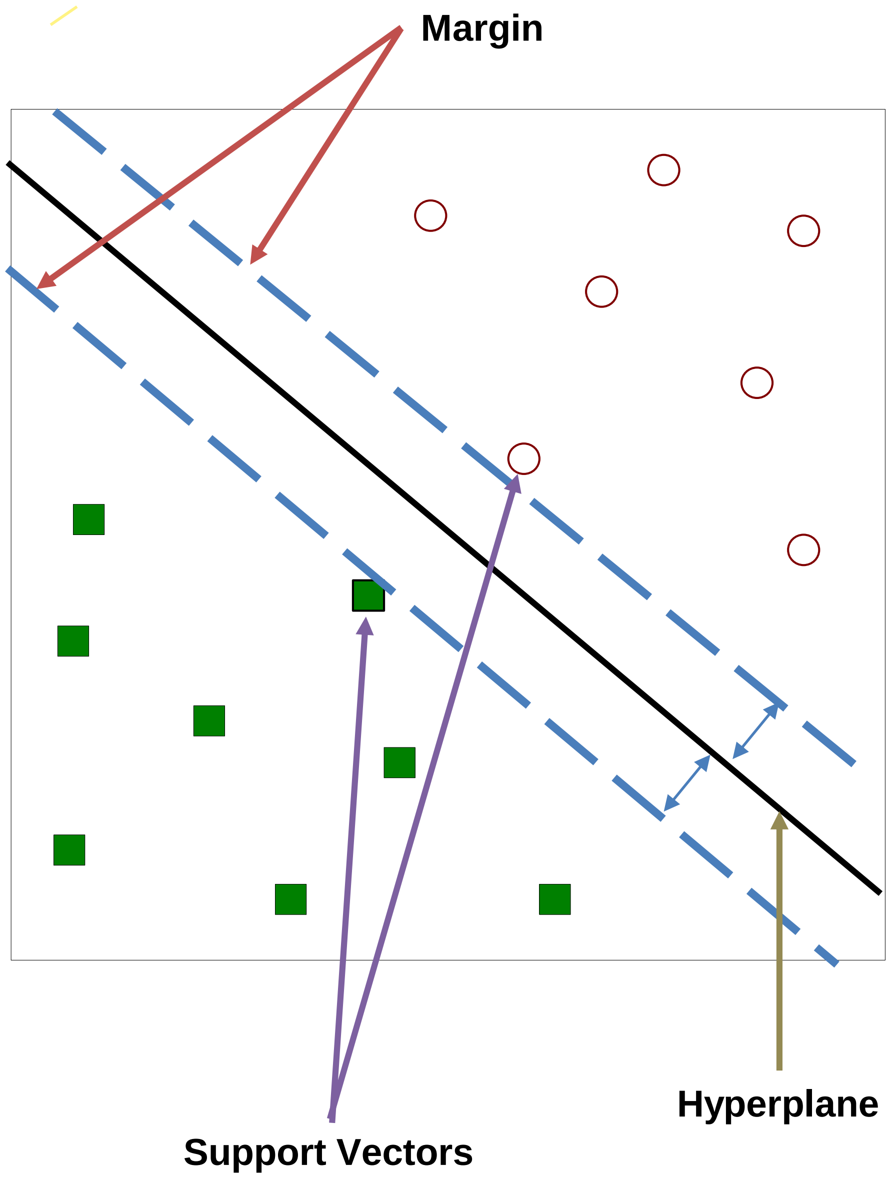
Support Vector Machine

Support Vector Machine(SVM)

- Decision Boundary
- Which Solution is Better?
- **Hyperplane**
  - Best divides the dataset into classes
  - Line for 2D space, Plane or a hyperplane for higher dimensions.
- **Margin**
  - The margin is the distance between the hyperplane and the nearest points from either class
  - A larger margin implies a better generalization for the classifier
  - **Hard Margin SVM**: Assumes data is perfectly separable by a hyperplane without any errors
  - **Soft Margin SVM**: Allows some misclassifications or errors
- **Support Vectors**
  - Points closest to the hyperplane
  - Decide the position and orientation of hyperplane
  - "Supporting" points that define the boundary between classes.



SVM

$y = \beta_0 + \beta_1 x$

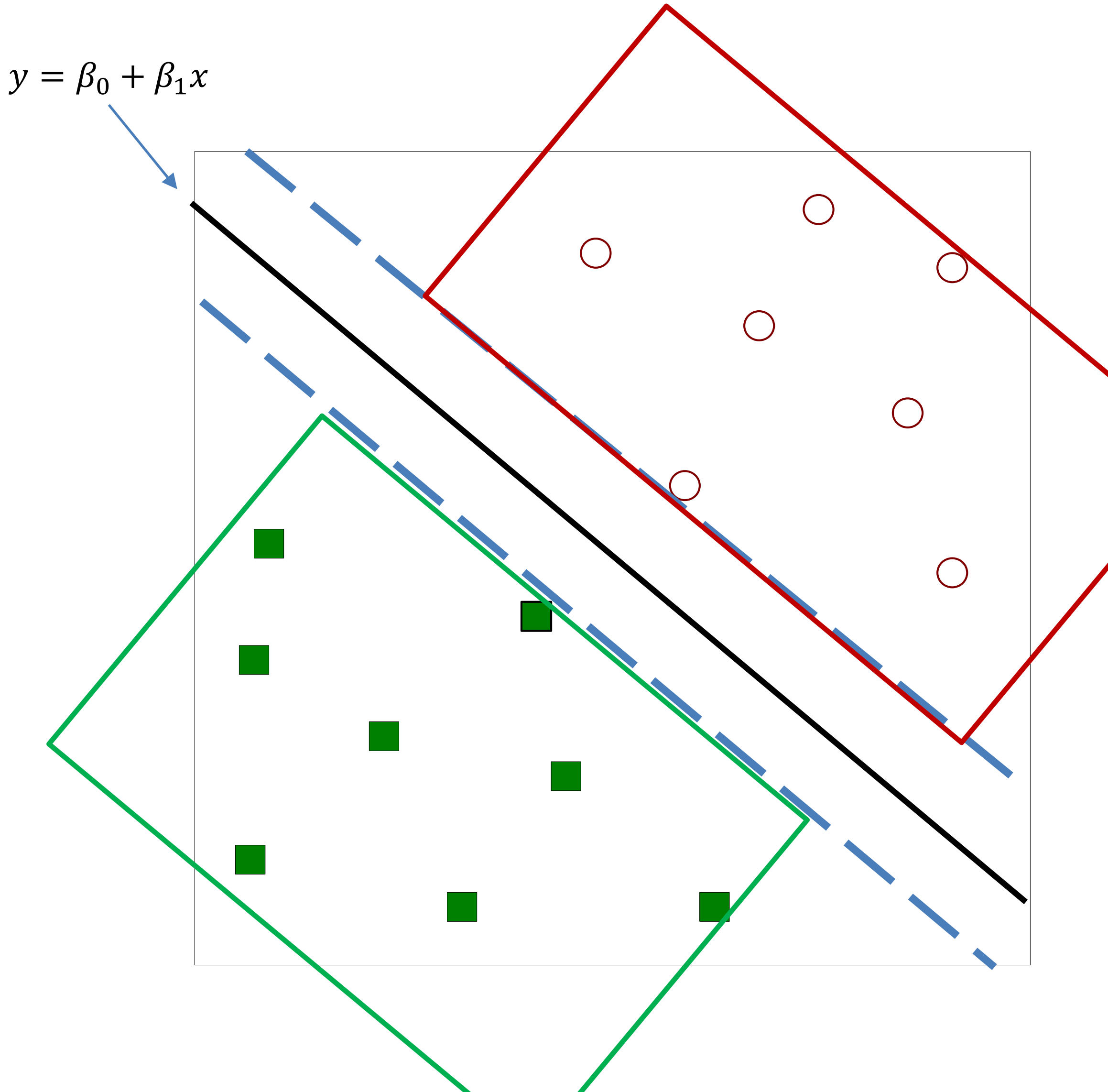
$\beta_0 + \beta_1 x - y = 0$

What if

$\beta_0 + \beta_1 x - y > 0$

Or

$\beta_0 + \beta_1 x - y < 0$



SVM Example 1

Sl. No.	Feature 1	Feature 2	Class/ Labels
1	3	1	Positive
2	3	-1	Positive
3	6	1	Positive
4	6	-1	Positive
5	1	0	Negative
6	0	1	Negative
7	0	-1	Negative
8	-1	0	Negative

Positive Labels (+1)

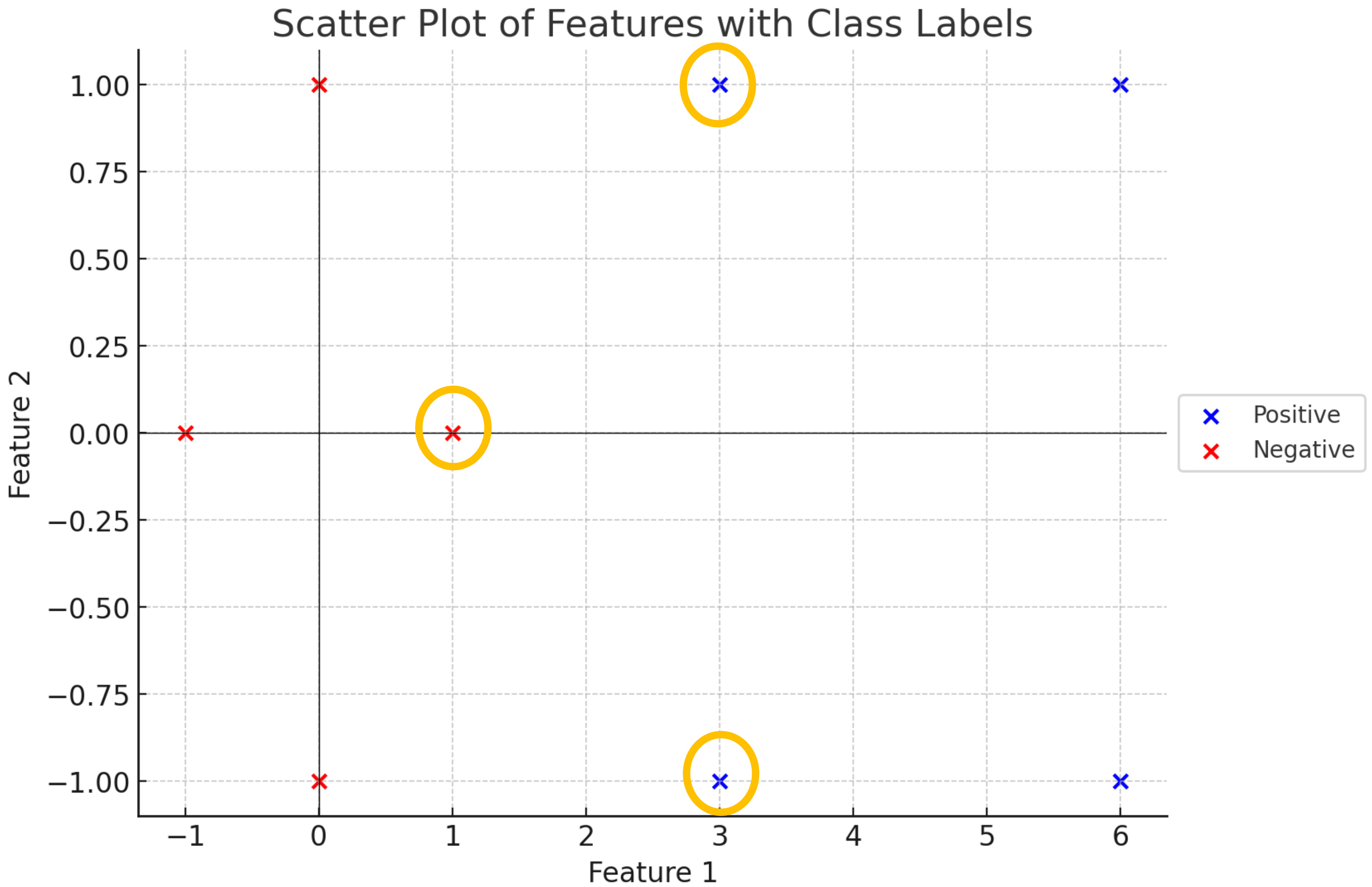
$$\left\{\begin{pmatrix}3\\1\end{pmatrix}\begin{pmatrix}3\\-1\end{pmatrix}\begin{pmatrix}6\\1\end{pmatrix}\begin{pmatrix}6\\-1\end{pmatrix}\right\}$$

Negative Labels (-1)

$$\left\{\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0\\-1\end{pmatrix}\begin{pmatrix}-1\\0\end{pmatrix}\right\}$$

Support Vectors

$$\left\{s_1 = \begin{pmatrix}1\\0\end{pmatrix}, s_2 = \begin{pmatrix}3\\1\end{pmatrix}, s_3 = \begin{pmatrix}3\\-1\end{pmatrix}\right\}$$



SVM Example 1

Support Vectors

$$\left\{s_1 = \begin{pmatrix}1\\0\end{pmatrix}, s_2 = \begin{pmatrix}3\\1\end{pmatrix}, s_3 = \begin{pmatrix}3\\-1\end{pmatrix}\right\}$$

Vector Augmentation

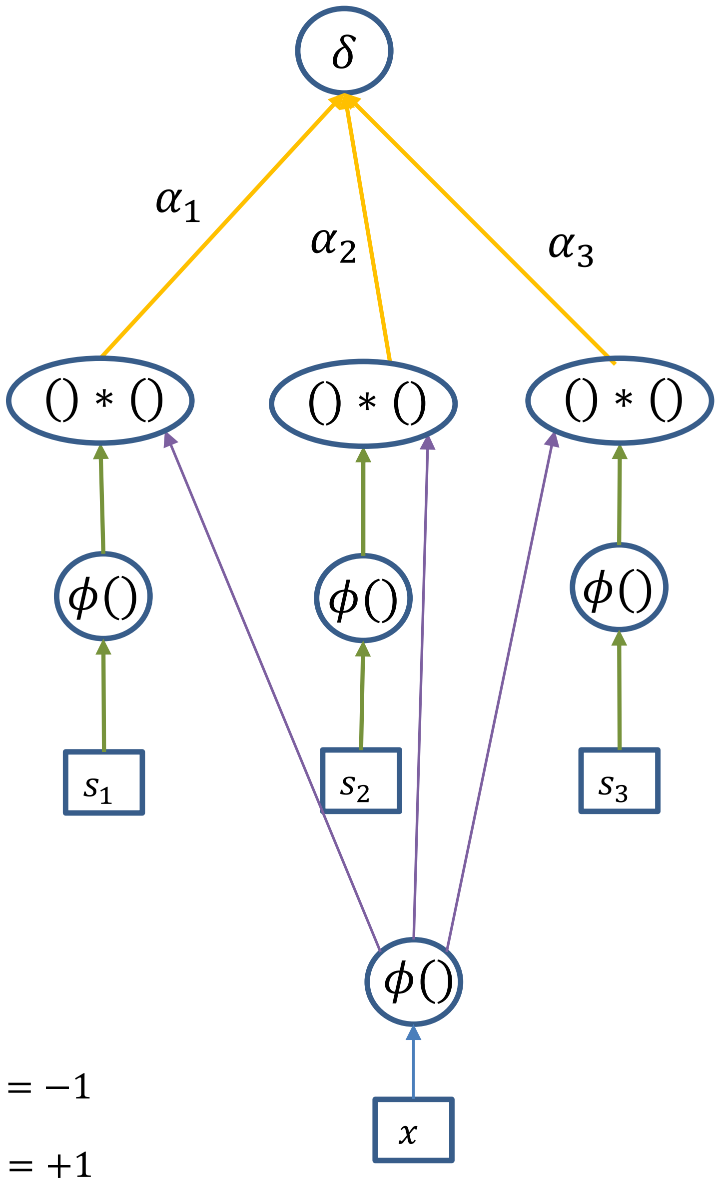
$s_1 = (10), \widetilde{s_1} = (101)$ *Negative*

$s_2 = (31), \widetilde{s_2} = (311)$ *Positive*

$s_3 = (3\,(-1)), \widetilde{s_3} = (3\,(-1)1)$ *Positive*

Augmented Support Vectors

$$\left\{\widetilde{s_1} = \begin{pmatrix}1\\0\\1\end{pmatrix}, \widetilde{s_2} = \begin{pmatrix}3\\1\\1\end{pmatrix}, \widetilde{s_3} = \begin{pmatrix}3\\-1\\1\end{pmatrix}\right\}$$



Find the values of α_i

$$\alpha_1\phi(s_1).\phi(s_1) + \alpha_2\phi(s_2).\phi(s_1) + \alpha_3\phi(s_3).\phi(s_1) = -1$$

$$\alpha_1\phi(s_1).\phi(s_2) + \alpha_2\phi(s_2).\phi(s_2) + \alpha_3\phi(s_3).\phi(s_2) = +1$$

$$\alpha_1\phi(s_1).\phi(s_3) + \alpha_2\phi(s_2).\phi(s_3) + \alpha_3\phi(s_3).\phi(s_3) = +1$$

$\Rightarrow$

$$\alpha_1\widetilde{s_1}.\widetilde{s_1} + \alpha_2\widetilde{s_2}.\widetilde{s_1} + \alpha_3\widetilde{s_3}.\widetilde{s_1} = -1$$

$$\alpha_1\widetilde{s_1}.\widetilde{s_2} + \alpha_2\widetilde{s_2}.\widetilde{s_2} + \alpha_3\widetilde{s_3}.\widetilde{s_2} = +1$$

$$\alpha_1\widetilde{s_1}.\widetilde{s_3} + \alpha_2\widetilde{s_2}.\widetilde{s_3} + \alpha_3\widetilde{s_3}.\widetilde{s_3} = +1$$

SVM Example 1

$$\alpha_1 \tilde{s}_1 . \tilde{s}_1 + \alpha_2 \tilde{s}_2 . \tilde{s}_1 + \alpha_3 \tilde{s}_3 . \tilde{s}_1 = -1$$

$$\alpha_1 \tilde{s}_1 . \tilde{s}_2 + \alpha_2 \tilde{s}_2 . \tilde{s}_2 + \alpha_3 \tilde{s}_3 . \tilde{s}_2 = +1$$

$$\alpha_1 \tilde{s}_1 . \tilde{s}_3 + \alpha_2 \tilde{s}_2 . \tilde{s}_3 + \alpha_3 \tilde{s}_3 . \tilde{s}_3 = +1$$

$$\alpha_1 (101).(101) + \alpha_2 (311).(101) + \alpha_3 (3\ (-1)1).(101) = -1$$

$$\alpha_1 (101).(311) + \alpha_2 (311).(311) + \alpha_3 (3\ (-1)1).(311) = +1$$

$$\alpha_1 \boxed{(101).(3\ (-1)1)} + \alpha_2 (311).(3\ (-1)1) + \alpha_3 (3\ (-1)1).(3\ (-1)1) = +1$$

Compute Dot Product of $(101).(3\ (-1)1)$

$$(101).(3\ (-1)1) = 1 * 3 + 0 * (-1) + 1 * 1 = 4$$

$$2\alpha_1 + 4\alpha_2 + 4\alpha_3 = -1 \quad \text{Eq. 1}$$

$$4\alpha_1 + 11\alpha_2 + 9\alpha_3 = +1 \quad \text{Eq. 2}$$

$$4\alpha_1 + 9\alpha_2 + 11\alpha_3 = +1 \quad \text{Eq. 3}$$

Divide Eq 1 by 2

$$2\alpha_1 + 4\alpha_2 + 4\alpha_3 = -1 \rightarrow \alpha_1 + 2\alpha_2 + 2\alpha_3 = -\frac{1}{2} \quad \text{Eq. 4}$$

Subtract (Eq 2 – Eq 4)

$$(4\alpha_1 + 11\alpha_2 + 9\alpha_3) - (\alpha_1 + 2\alpha_2 + 2\alpha_3) = +1 - \left(-\frac{1}{2}\right)$$

$$3\alpha_1 + 9\alpha_2 + 7\alpha_3 = \left(\frac{3}{2}\right) \quad \text{Eq. 5}$$

Subtract (Eq 3 – Eq 4)

$$(4\alpha_1 + 9\alpha_2 + 11\alpha_3) - (\alpha_1 + 2\alpha_2 + 2\alpha_3) = +1 - \left(-\frac{1}{2}\right)$$

$$3\alpha_1 + 7\alpha_2 + 9\alpha_3 = \left(\frac{3}{2}\right) \quad \text{Eq. 6}$$

Subtract (Eq 6 – Eq 5)

$$(3\alpha_1 + 7\alpha_2 + 9\alpha_3) - (3\alpha_1 + 9\alpha_2 + 7\alpha_3) = \left(\frac{3}{2}\right) - \left(\frac{3}{2}\right)$$

$$-2\alpha_2 + 2\alpha_3 = 0 \rightarrow \alpha_3 = \alpha_2 \quad \text{Eq. 7}$$

Substitute Eq 7 in Eq 4

$$\alpha_1 + 2\alpha_2 + 2\alpha_3 = -\frac{1}{2} \rightarrow \alpha_1 + 2\alpha_3 + 2\alpha_3 = -\frac{1}{2}$$

$$\alpha_1 = -\frac{1}{2} - 4\alpha_3 \quad \text{Eq. 8}$$

Substitute Eq 7 & Eq 8 in Eq 5

$$3\alpha_1 + 9\alpha_2 + 7\alpha_3 = \left(\frac{3}{2}\right) \rightarrow 3\left(-\frac{1}{2} - 4\alpha_3\right) + 9\alpha_3 + 7\alpha_3 = \left(\frac{3}{2}\right)$$

$$\alpha_3 = \frac{3}{4}$$

Eq. 7

$$\alpha_2 = \frac{3}{4}$$

Eq. 8

$$\alpha_1 = -\frac{7}{2}$$

SVM Example 1

$\alpha_1 = -\frac{7}{2} \qquad \alpha_2 = \frac{3}{4} \qquad \alpha_3 = \frac{3}{4}$

Augmented Support Vectors

$$\left\{ \tilde{s}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \tilde{s}_2 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \tilde{s}_3 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \right\}$$

How to get the hyperplane that
separates the positive and negative class

$\tilde{w} = \sum_{i=1}^{n} \alpha_i \tilde{s}_i$

$$\tilde{w} = -\frac{7}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \frac{3}{4}\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} + \frac{7}{4}\begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

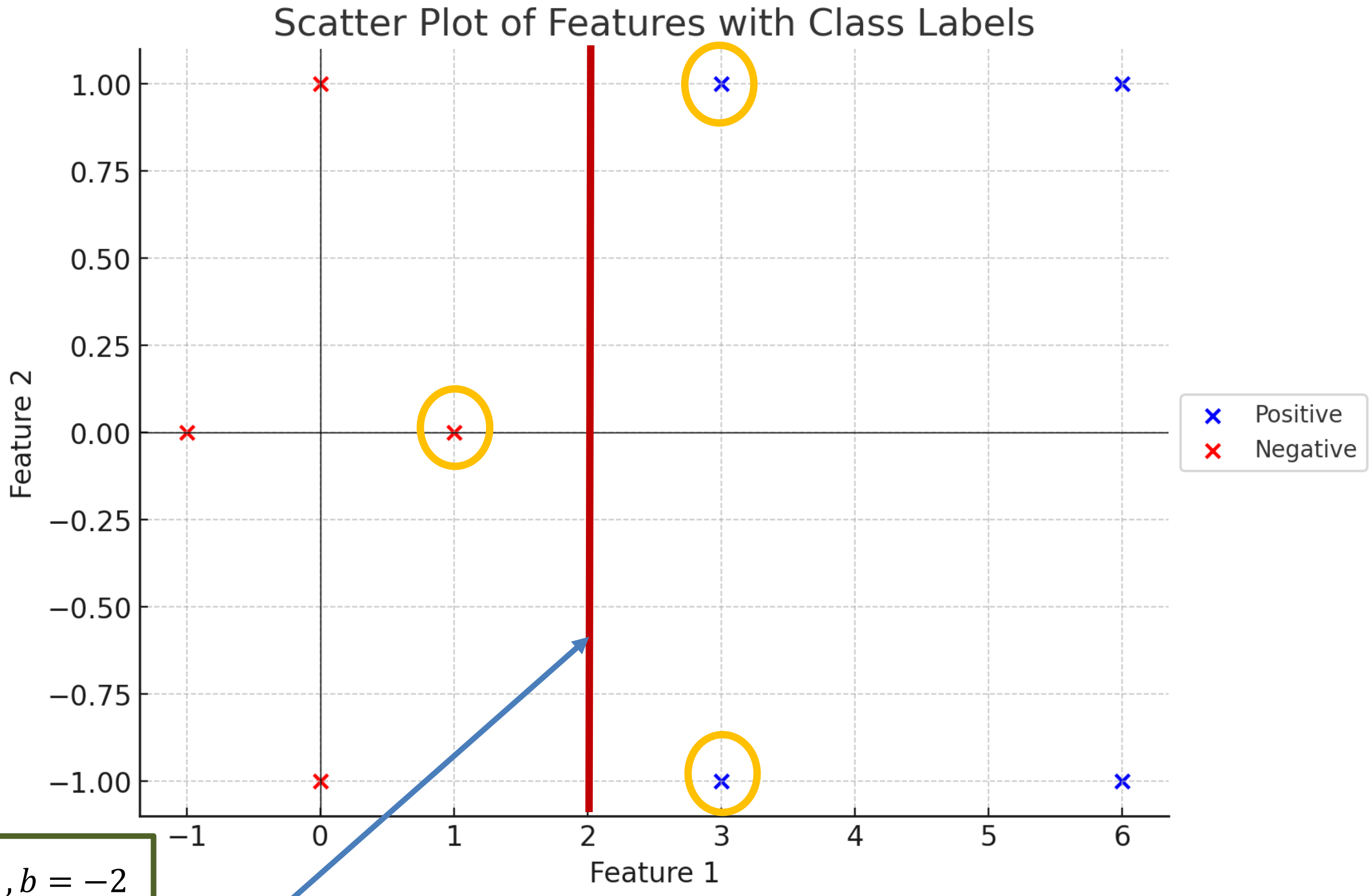
$\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$

x coefficient

Y Coefficient

Bias (Intercept)

$\boxed{(Weight\ Vector(w) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, b = -2}$



Hyperplane

What about Margin?

$\beta_0 + \beta_1 x - y = 0$

$-2 + 1x - 0y = 0 \rightarrow \boxed{x = 2}$

SVM Example 1

$$\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

x coefficient

Y Coefficient

Bias (Intercept)

$$(Weight\ Vector(w) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, b = -2$$

Margin Calculation

Step 1: Compute the Norm of (w)

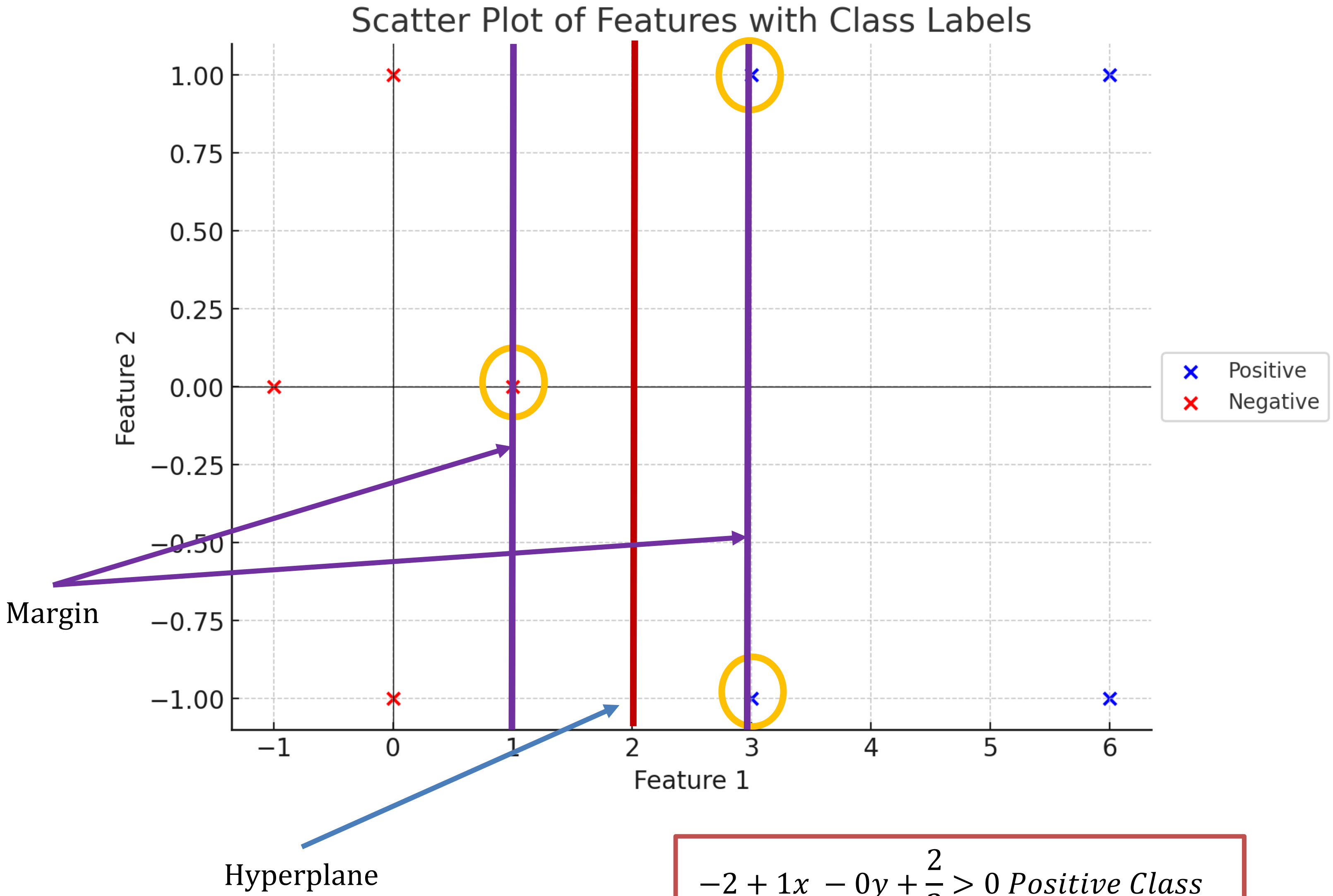
$$\|w\| = \sqrt{1^2 + 0^2} = 1$$

Step 2: Compute the Margin

$$M = \frac{2}{\|w\|}$$

$$M = \frac{2}{1} = 2$$

Hyperplane is exactly in the middle of Margin



$$-2 + 1x - 0y + \frac{2}{2} > 0 \ Positive\ Class$$

$$-2 + 1x - 0y - \frac{2}{2} < 0 \ Negative\ Class$$

SVM Characteristics

- The learning problem is formulated as a convex optimization problem
  - Efficient algorithms are available to find the global minima
  - Many of the other methods use greedy approaches and find locally optimal solutions
  - High computational complexity for building the model

- Robust to noise
- Overfitting is handled by maximizing the margin of the decision boundary,
- SVM can handle irrelevant and redundant attributes better than many other techniques
- The user needs to provide the type of kernel function and cost function
- Difficult to handle missing values

SVM – RBF Kernel

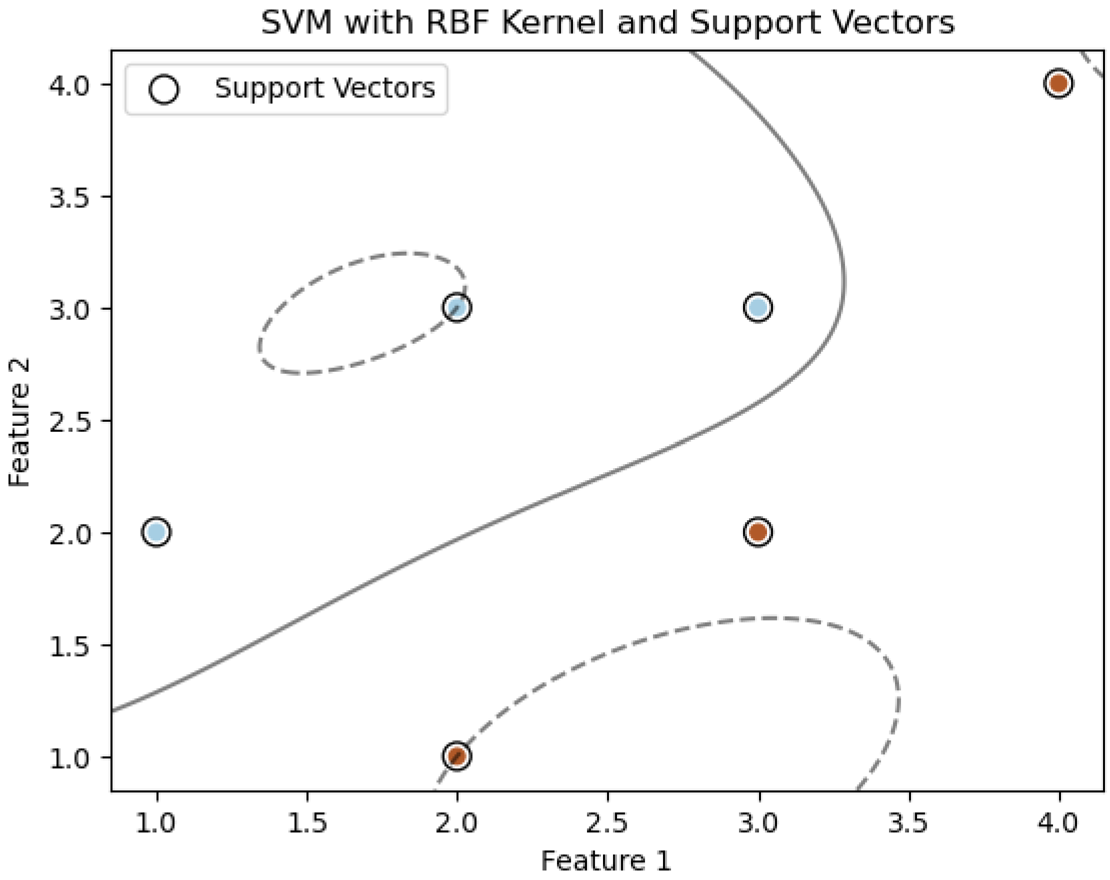
- Radial Basis Function(RBF) also known as **Gaussian Kernel**

$$K(x_1, x_2) = e^{(\gamma\, ||x_1 - x_2||^2)}$$

x_1, x_2 *are two data points*

$||x_1 - x_2||^2$ *is the squared euclidean distance between* x_1 & x_2

γ *defines the spread of kernel*



- Commonly used Kernels
  - Linear
  - Polynomial
  - RBF
  - Sigmoid
  - Laplacian
  - Rational Quadratic
  - Anova (Analysis of Variance)

SVM Example 2(Non-Linear)

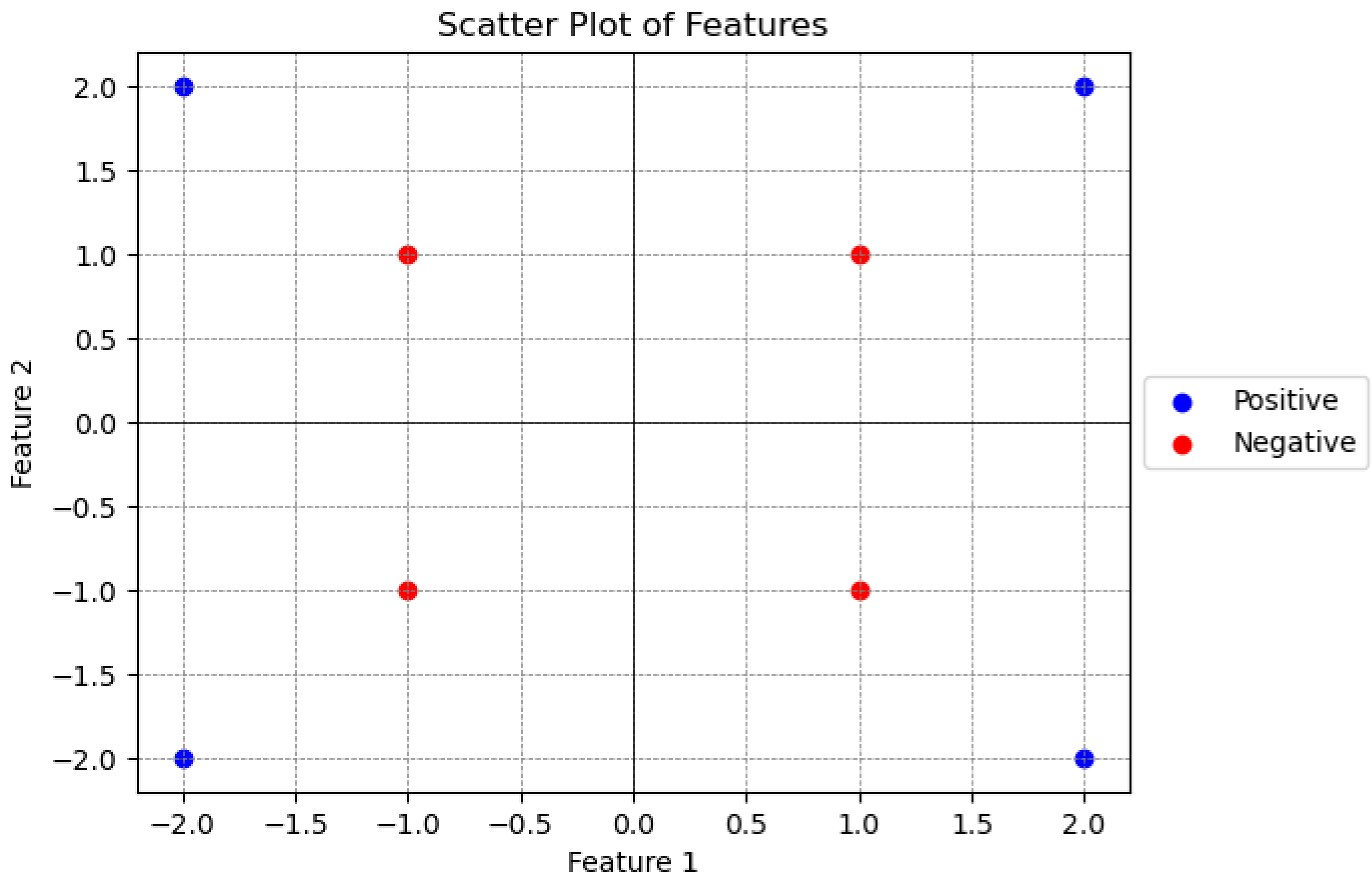
Sl. No.	Feature 1	Feature 2	Class/ Labels
1	2	2	Positive
2	2	-2	Positive
3	-2	-2	Positive
4	-2	2	Positive
5	1	1	Negative
6	1	-1	Negative
7	-1	-1	Negative
8	-1	1	Negative

Positive Labels (+1)

$$\left\{\begin{pmatrix}2\\2\end{pmatrix}\begin{pmatrix}2\\-2\end{pmatrix}\begin{pmatrix}-2\\-2\end{pmatrix}\begin{pmatrix}-2\\2\end{pmatrix}\right\}$$

Negative Labels (-1)

$$\left\{\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix}\begin{pmatrix}-1\\-1\end{pmatrix}\begin{pmatrix}-1\\1\end{pmatrix}\right\}$$



SVM Example 2(Non-Linear)

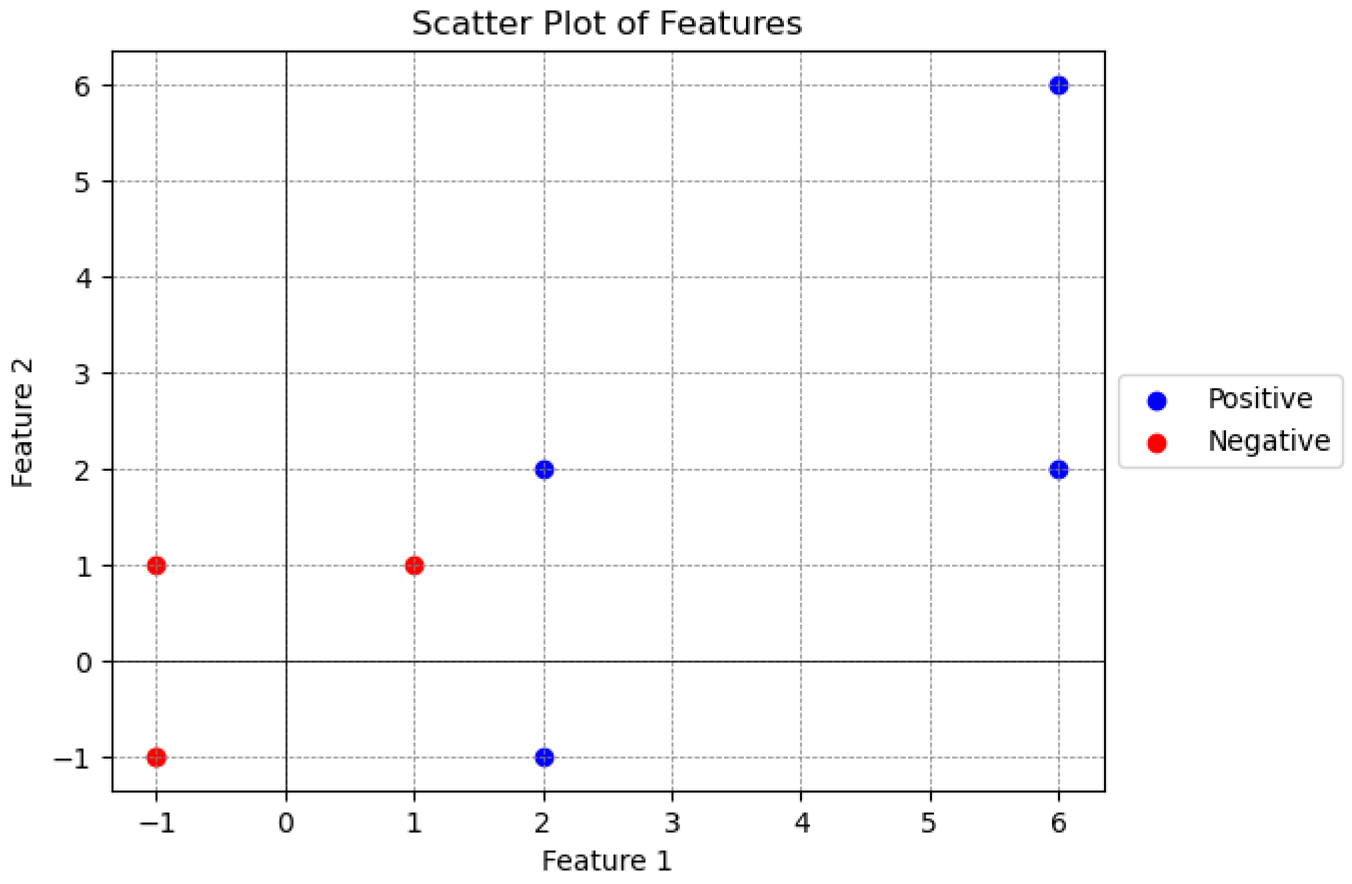
Sl. No.	Feature 1	Feature 2	Class/ Labels
1	2	2	Positive
2	6	2	Positive
3	6	6	Positive
4	2	6	Positive
5	1	1	Negative
6	1	-1	Negative
7	-1	-1	Negative
8	-1	1	Negative

Positive Labels (+1)

$$\left\{\binom{2}{2}\binom{6}{2}\binom{6}{6}\binom{2}{6}\right\}$$

Negative Labels (-1)

$$\left\{\binom{1}{1}\binom{1}{-1}\binom{-1}{-1}\binom{-1}{1}\right\}$$



Thank You