ED5340 - Data Science: Theory and Practise

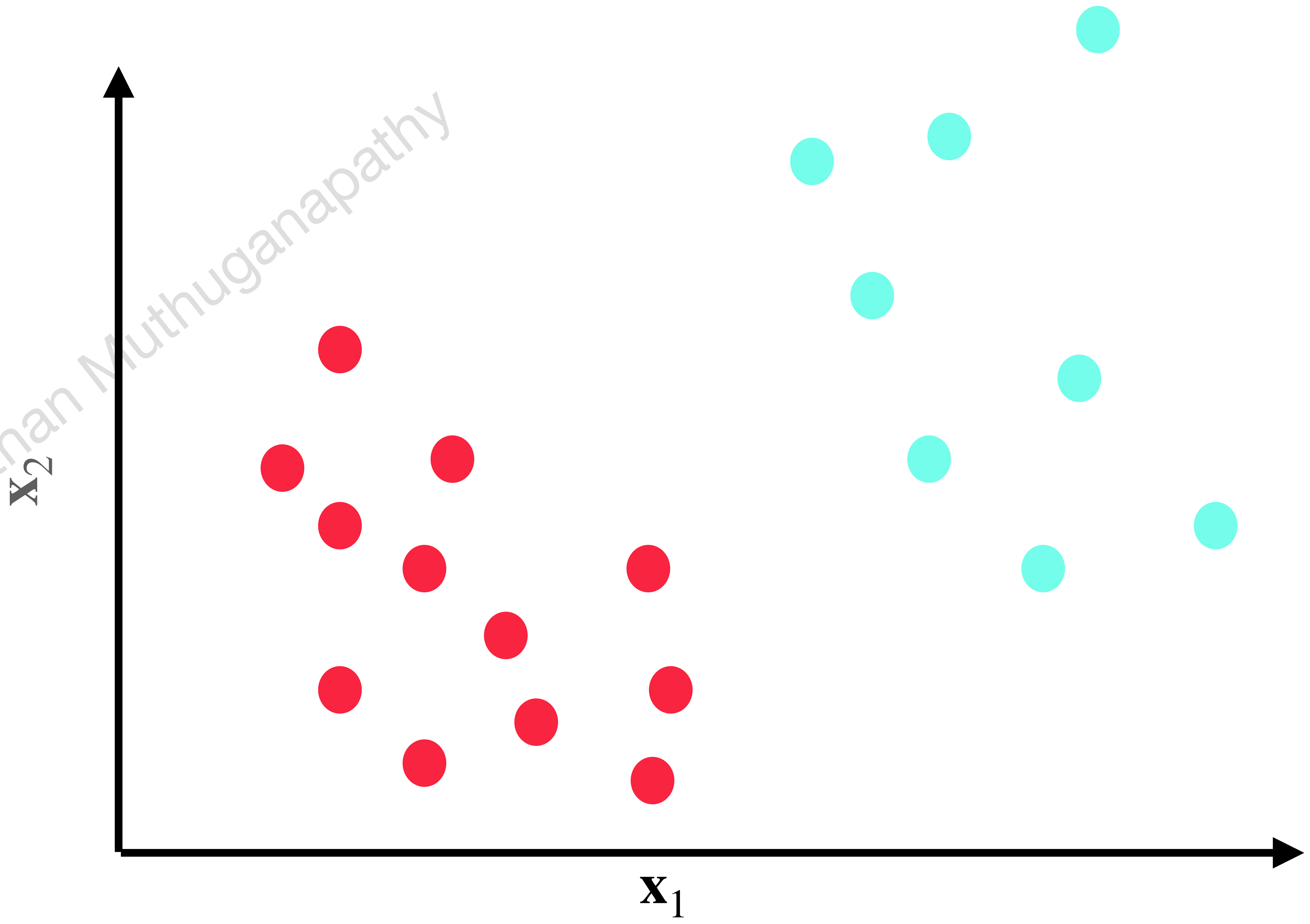
L20 - Support vector machine (SVM)

Ramanathan Muthuganapathy (https://ed.iitm.ac.in/~raman)
Course web page: https://ed.iitm.ac.in/~raman/datascience.html
Moodle page: Available at https://courses.iitm.ac.in/

Support vector machine

Decision boundary

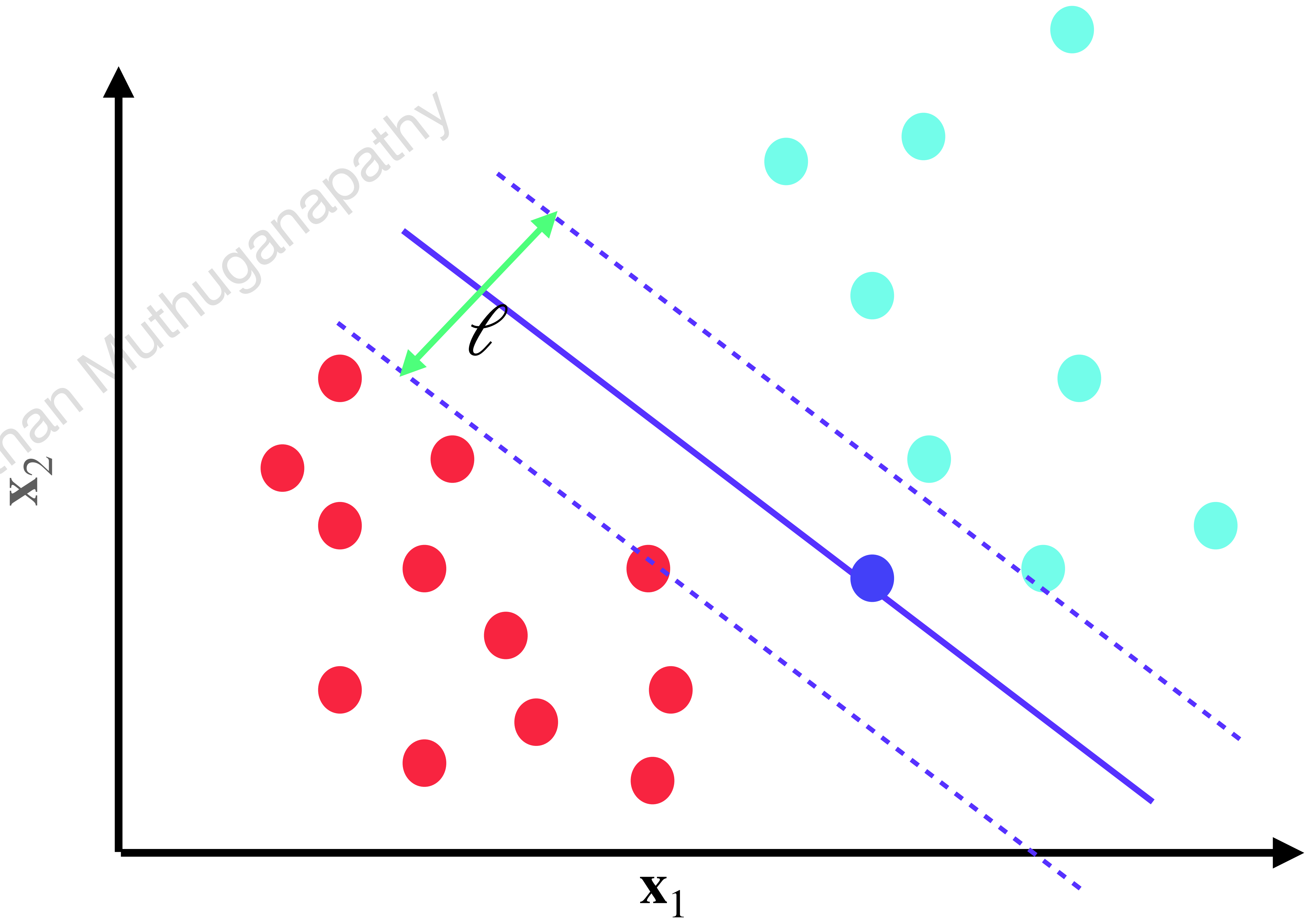
- Find a linear separable boundary
- with a margin
- looks very similar to logistic regression.



Support vector machine

Decision boundary

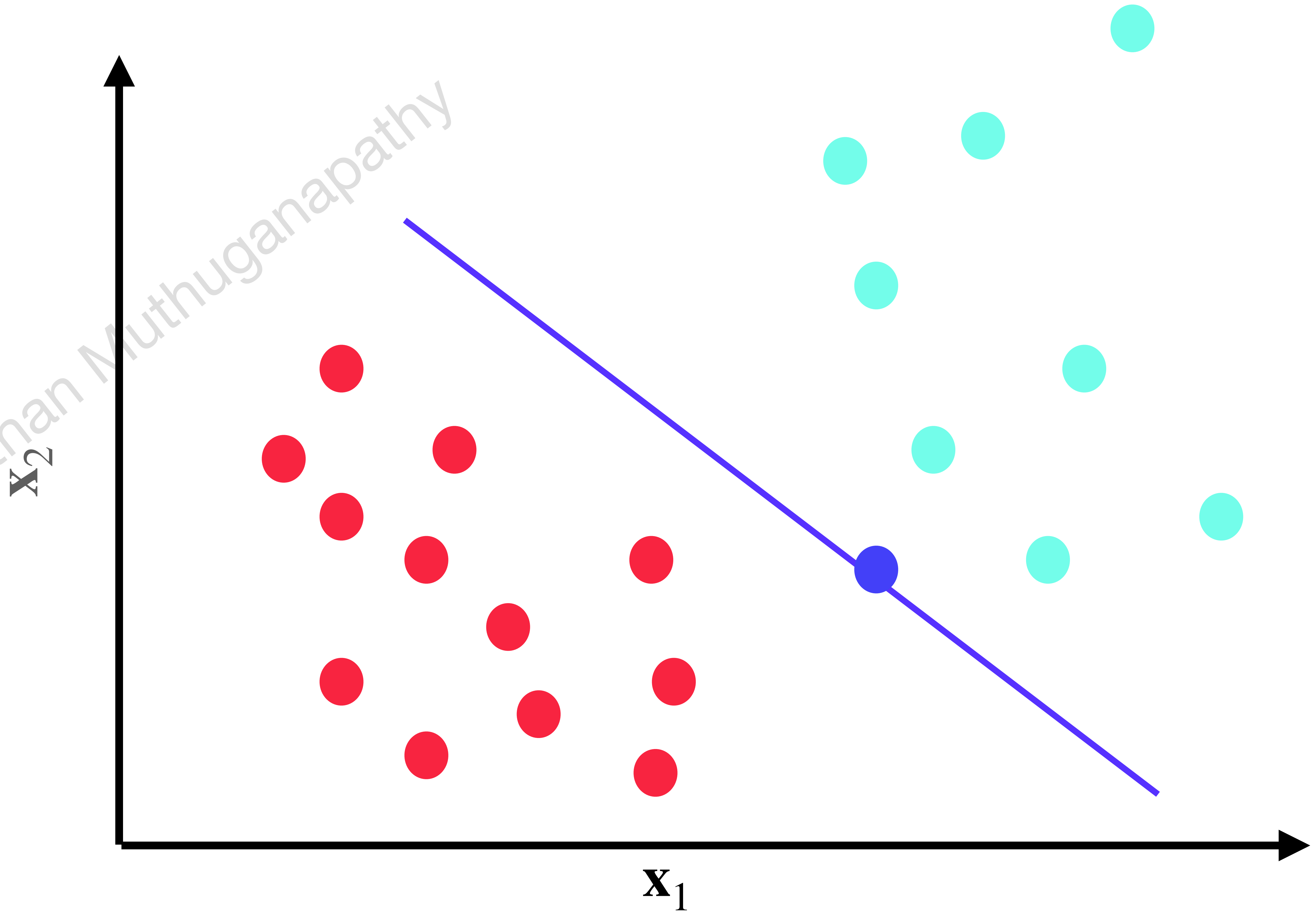
- Find a hyperplane that separates the two classes
- ℓ - the margin is the largest



Support vector machine

Recall vector notation

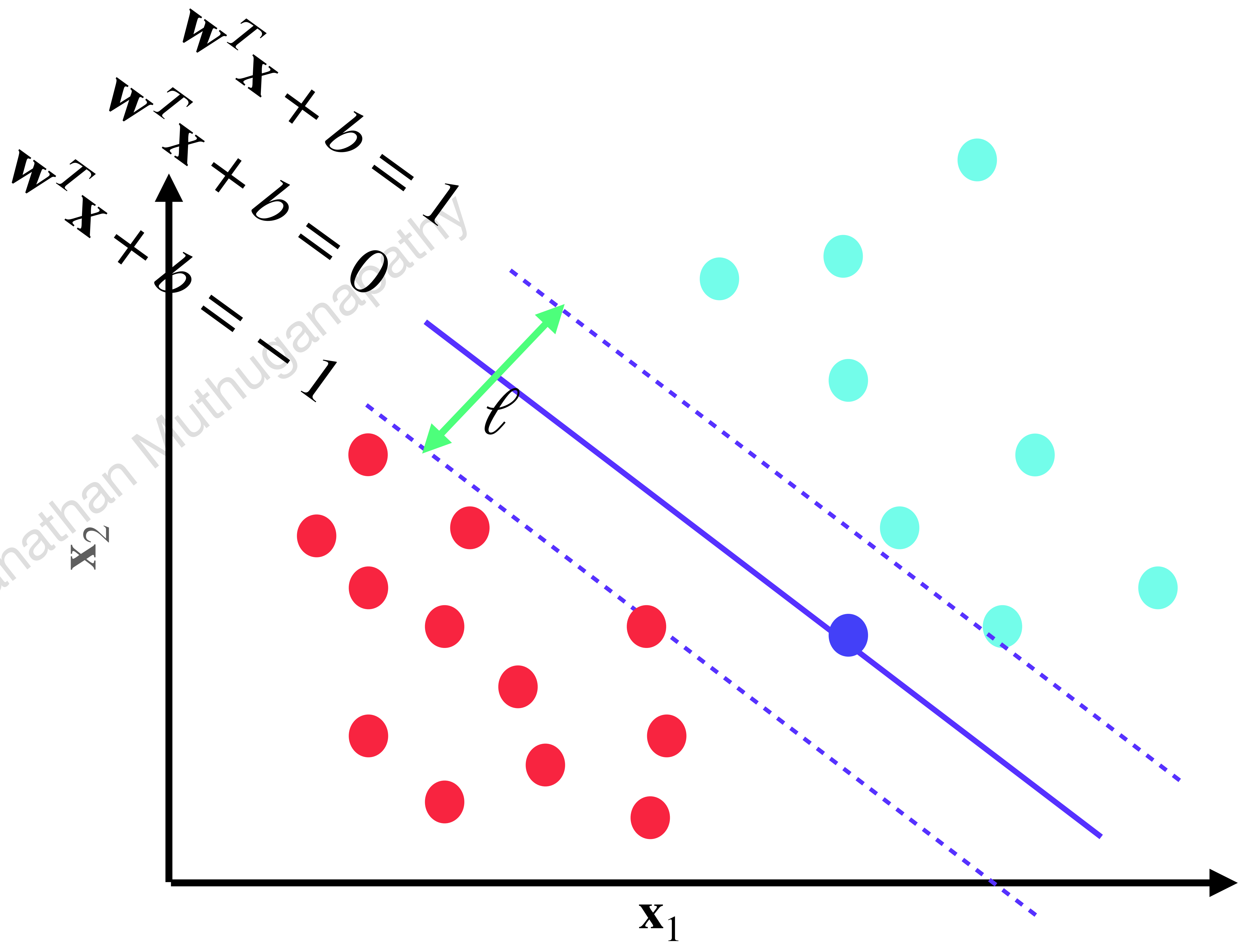
- $h_w(x) = w_0x_0 + w_1x_1 + w_1x_2$
- $\mathbf{w}^T = [w_1 \quad w_2]$
- $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$
- $b = w_0$ (bias)
- $h_w(x) = \mathbf{w}^T\mathbf{x} + b$



Support vector machine

Recall vector notation

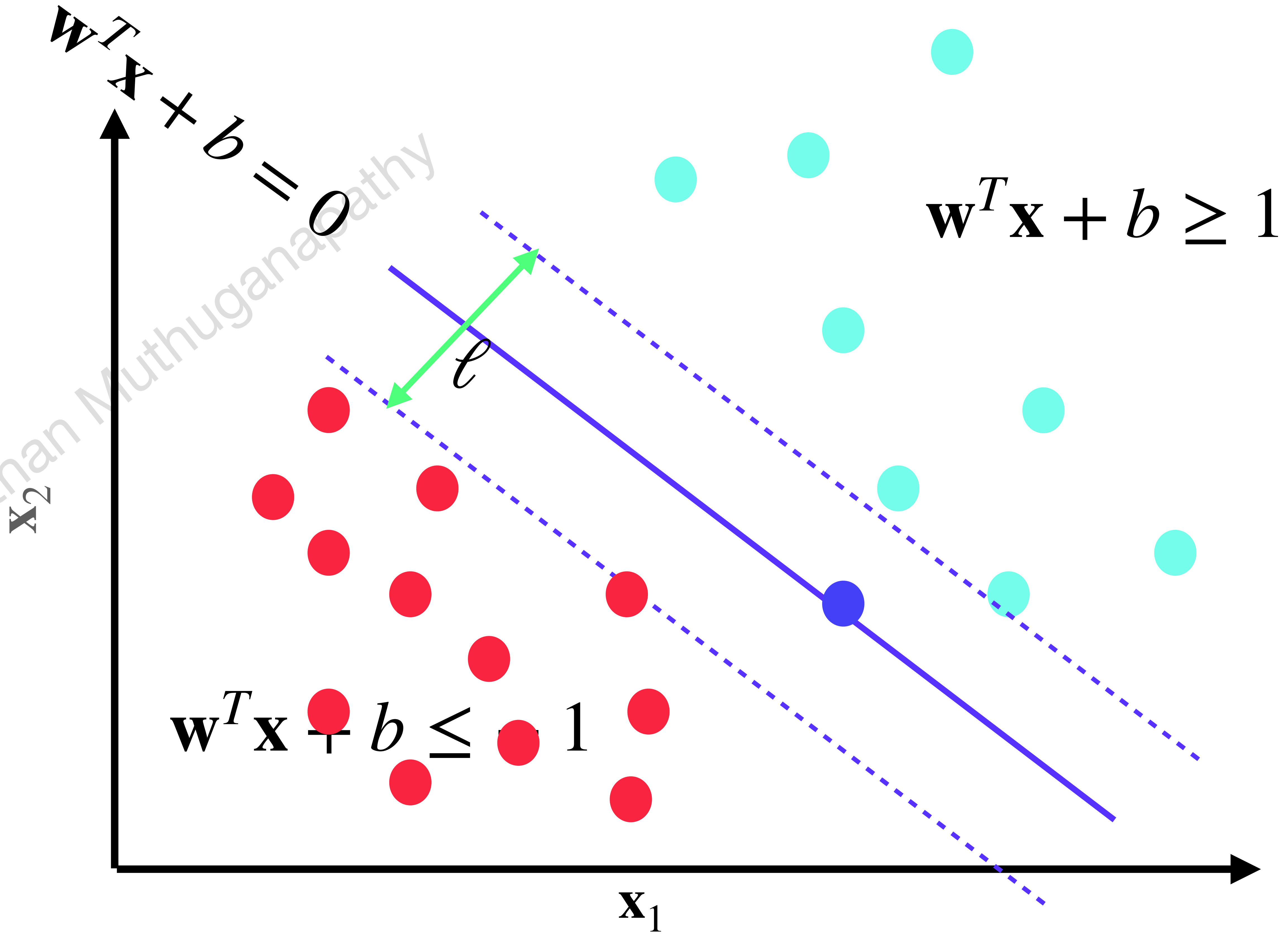
- $h_w(x) = w_0 x_0 + w_1 x_1 + w_1 x_2$
- $\mathbf{w}^T = [w_1 \quad w_2]$
- $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$
- $b = w_0$ (bias)
- $h_w(x) = \mathbf{w}^T \mathbf{x} + b$



Support vector machine

Recall vector notation

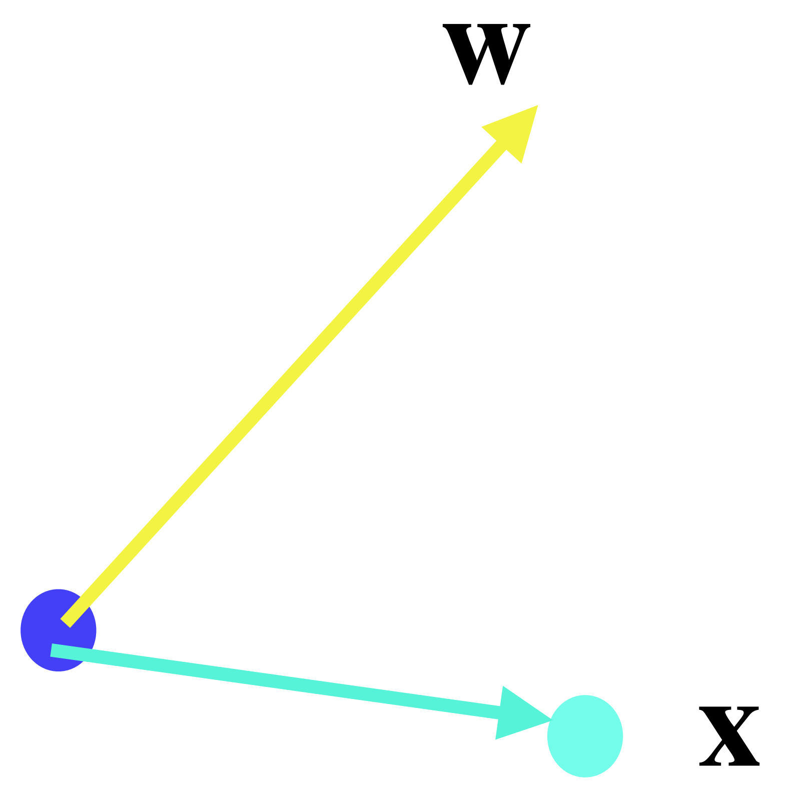
- $h_w(x) = w_0x_0 + w_1x_1 + w_1x_2$
- $\mathbf{w}^T = [w_1 \quad w_2]$
- $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$
- $b = w_0$ (bias)
- $h_w(x) = \mathbf{w}^T\mathbf{x} + b$



Support vector machine

Basic vector calculus

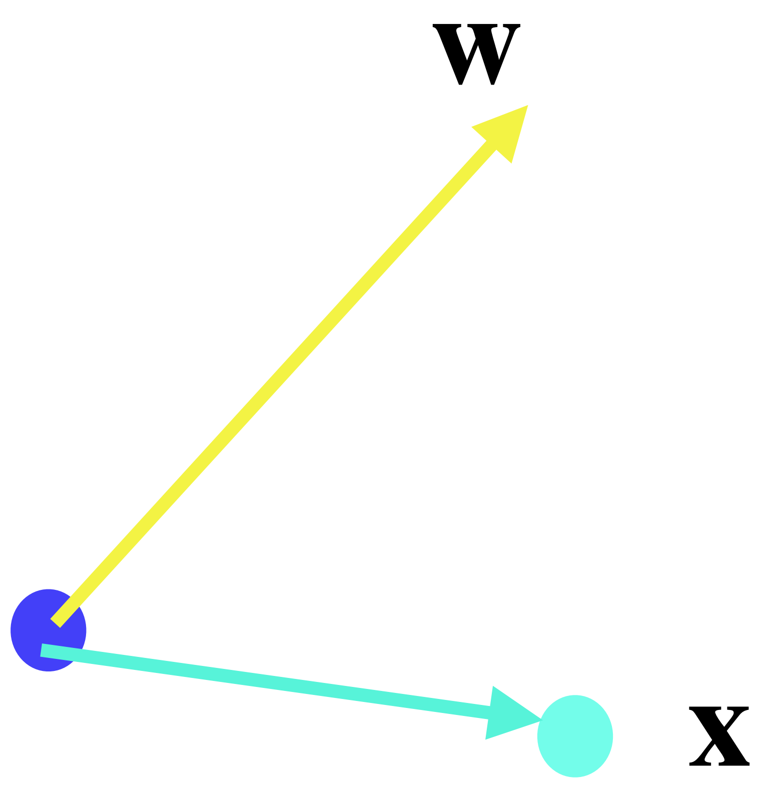
- Two vectors $\mathbf{w}$ and $\mathbf{x}$
- What is $\mathbf{w}^T\mathbf{x}$?



Support vector machine

Basic vector calculus

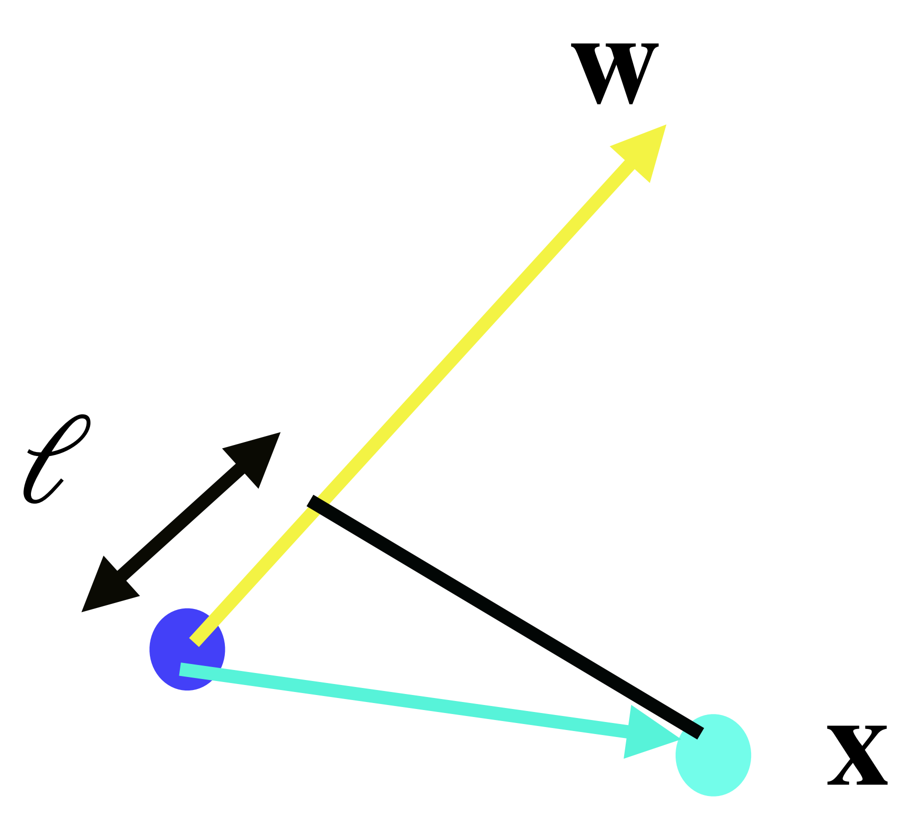
- What is $\mathbf{w}^T\mathbf{x}$?
- Dot product between the two vectors



Support vector machine

Basic vector calculus

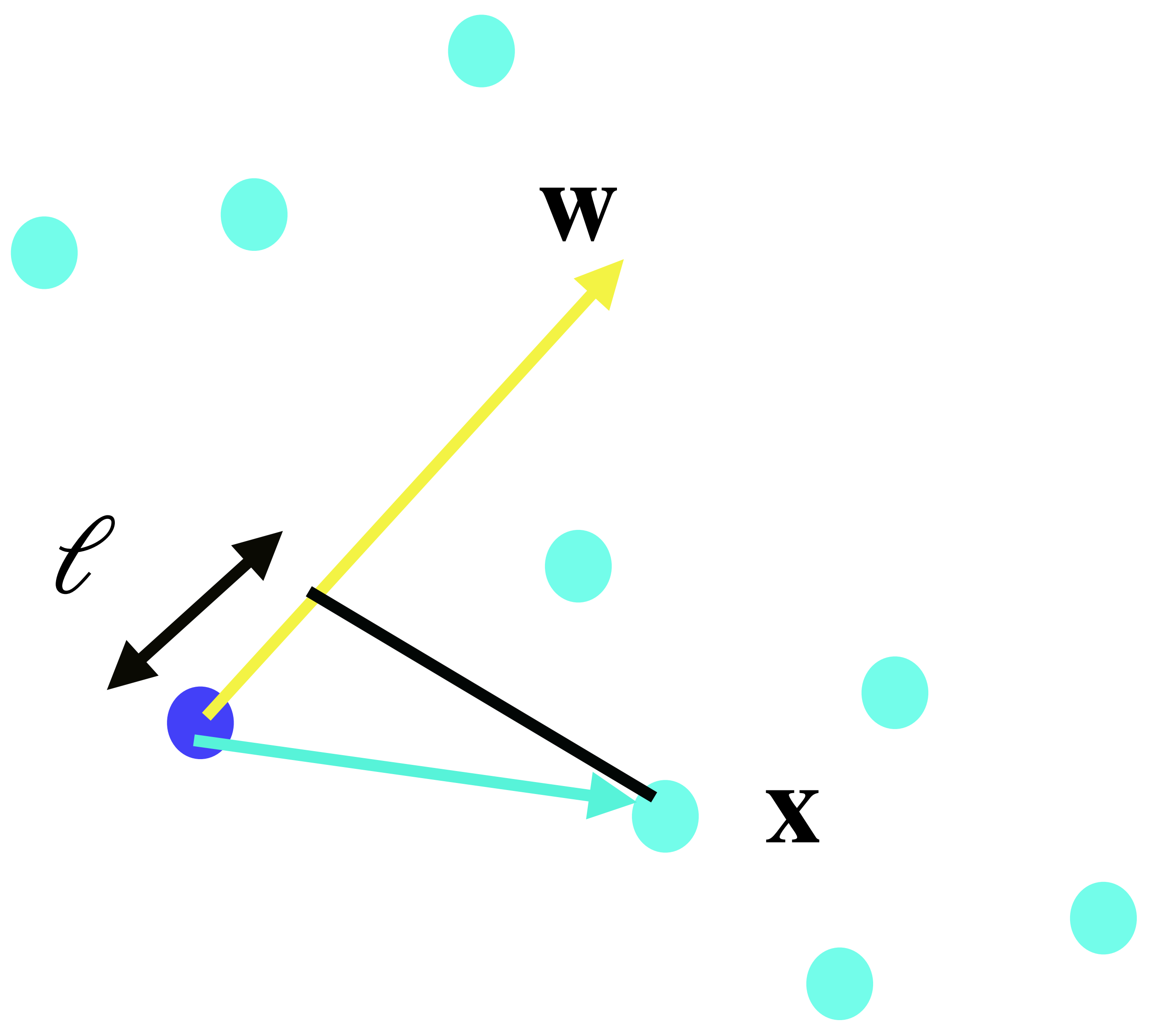
- What is $\mathbf{w}^T\mathbf{x}$?
- Dot product between the two vectors
- $\mathbf{w}^T\mathbf{x} = \ell||\mathbf{w}||$
- $\mathbf{w}^T\mathbf{x} \geq 0$



Support vector machine

Basic vector calculus

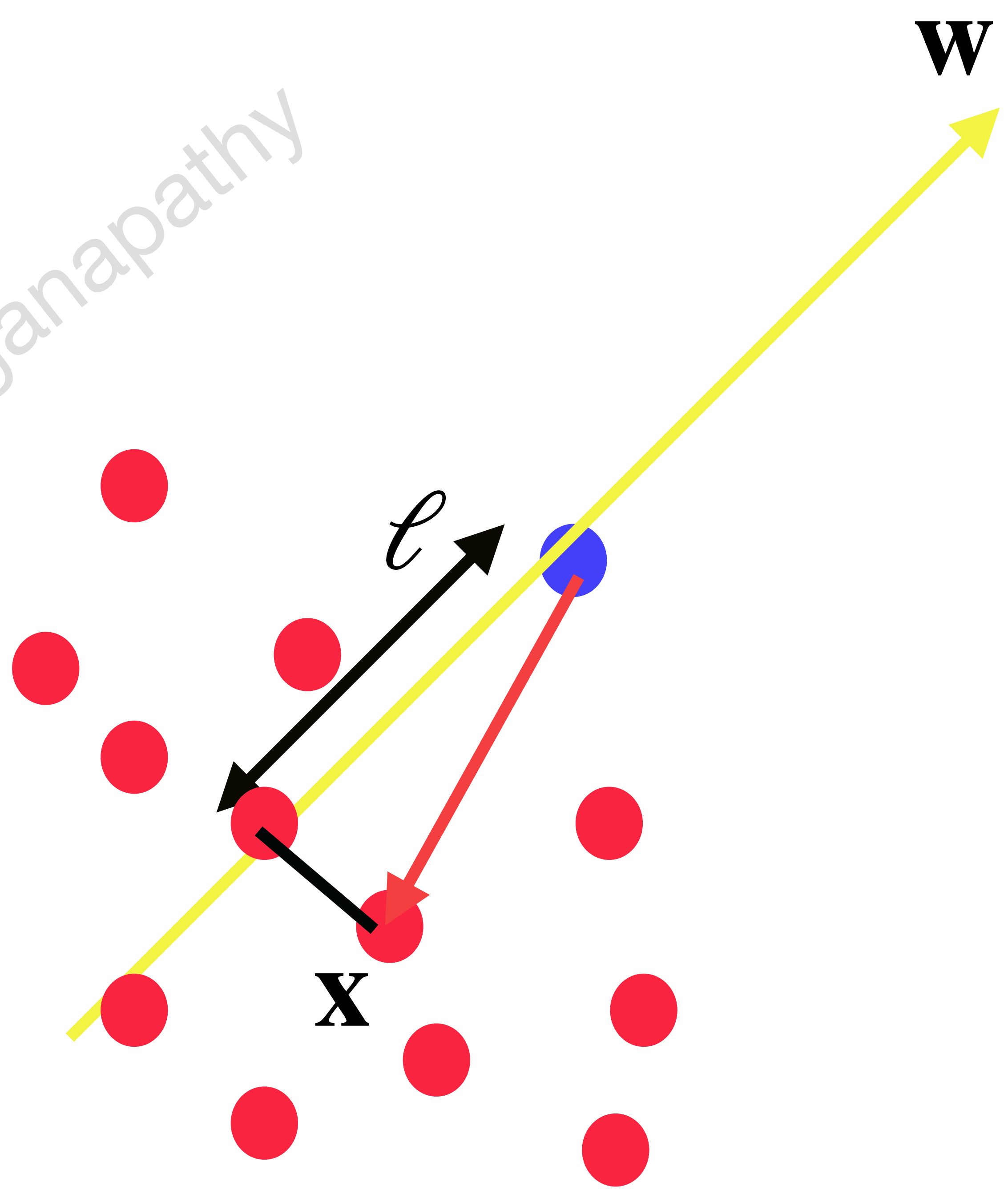
- $\mathbf{w}^T\mathbf{x} \geq 0$



Support vector machine

Basic vector calculus

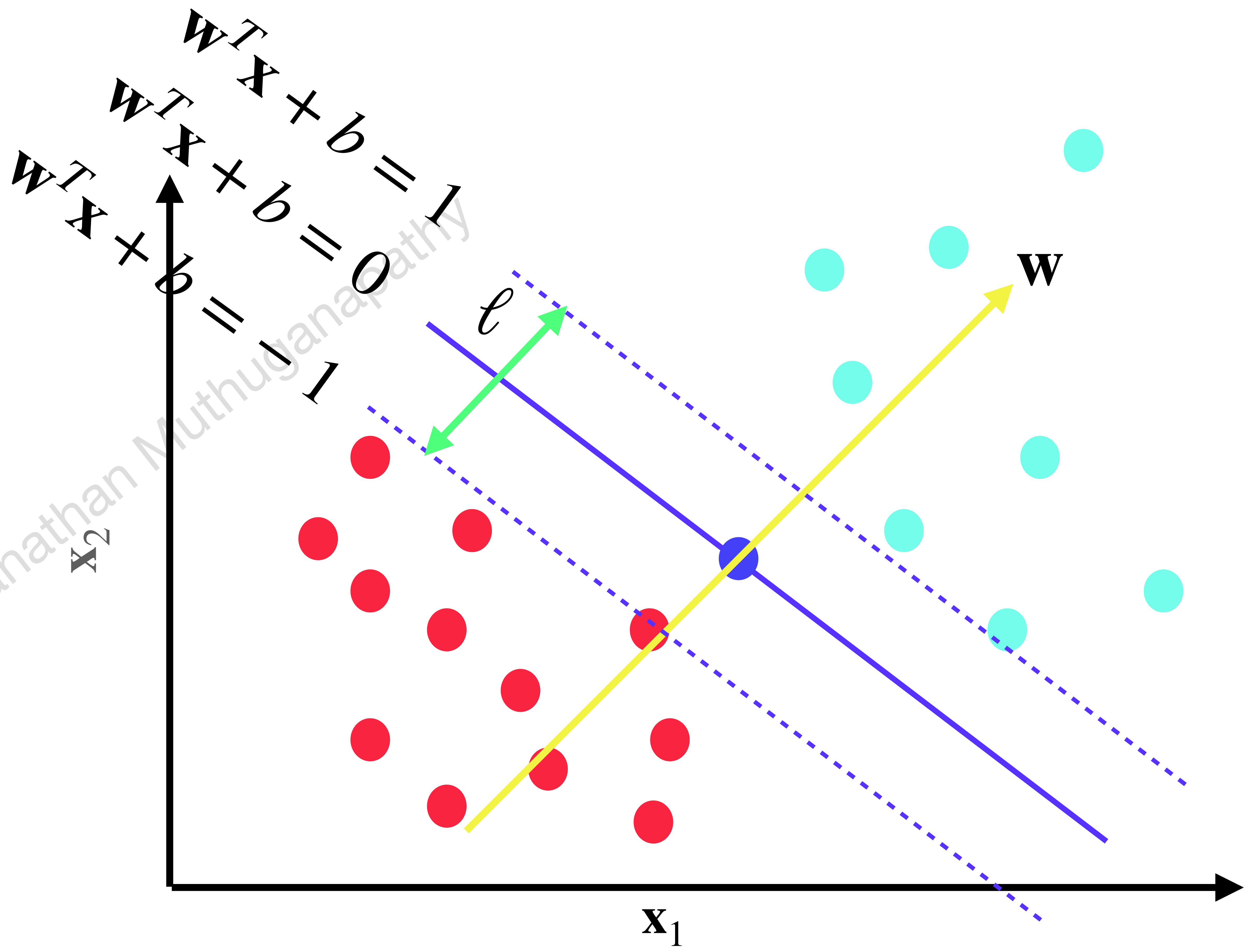
- $\mathbf{w}^T\mathbf{x} \leq 0$



Support vector machine

Putting together

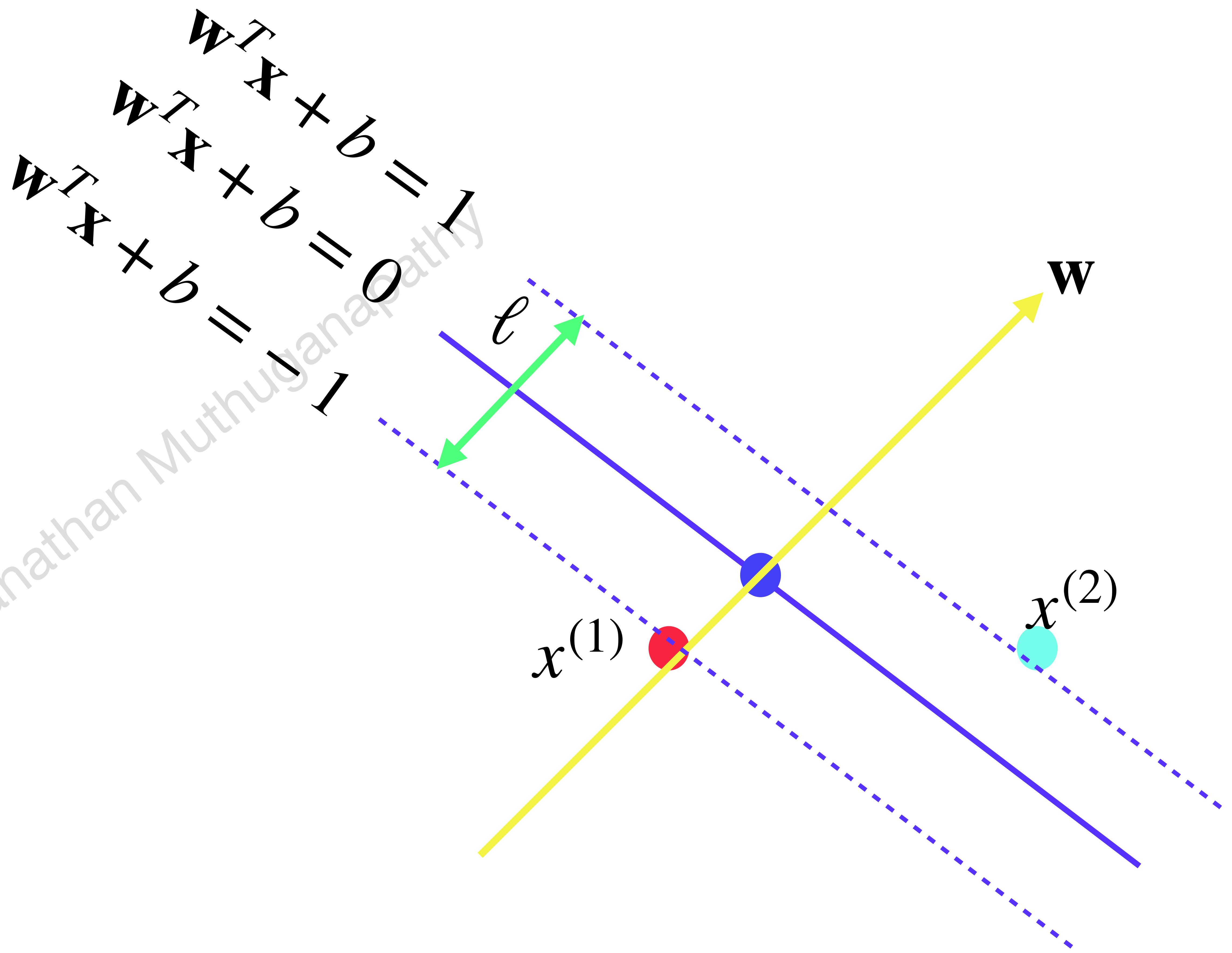
- Hyperplane
- perpendicular to $\mathbf{w}$
- with constraints



Support vector machine

Putting together

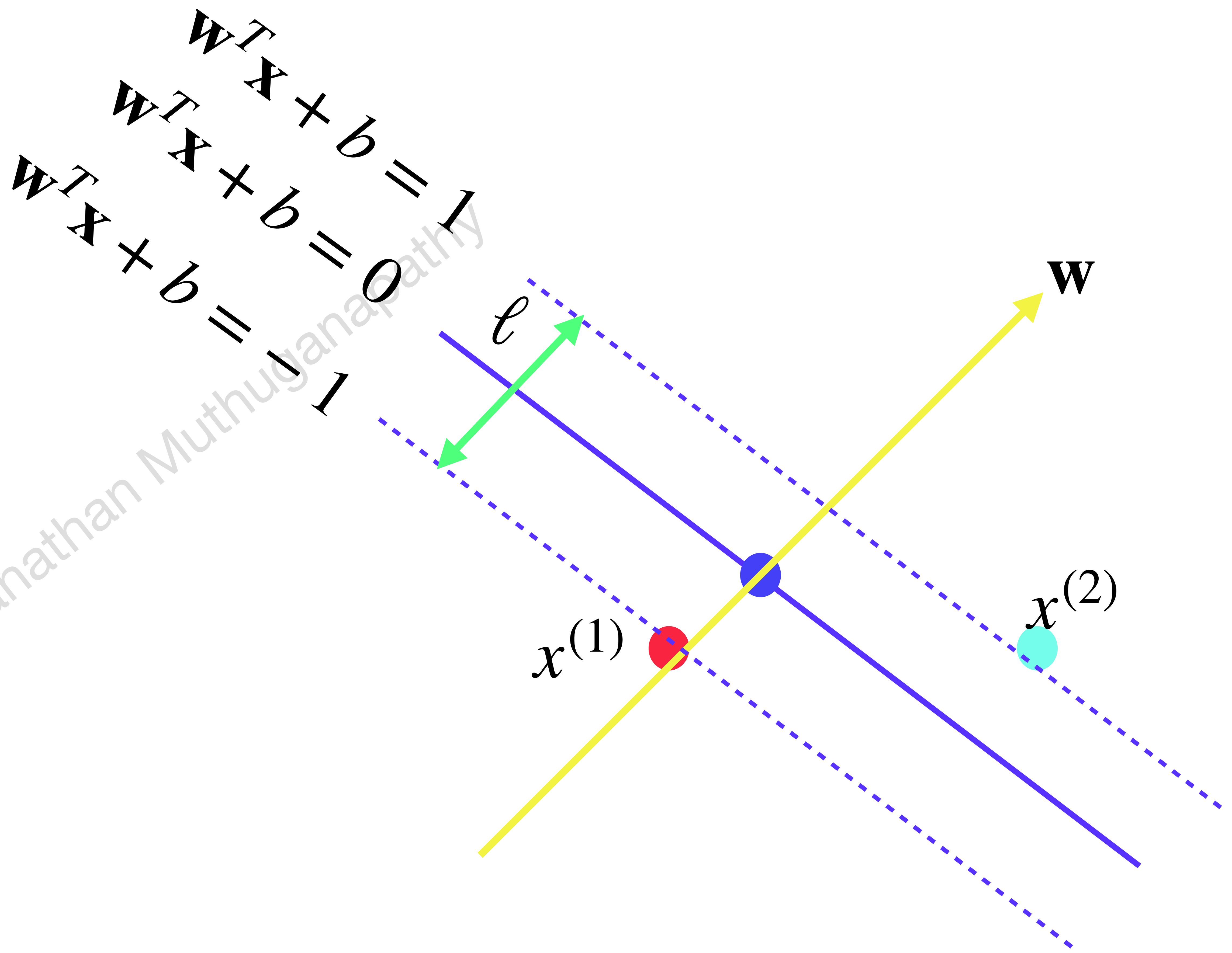
- $\mathbf{w}^T.(x^{(1)} - x^{(2)})$



Support vector machine

Putting together

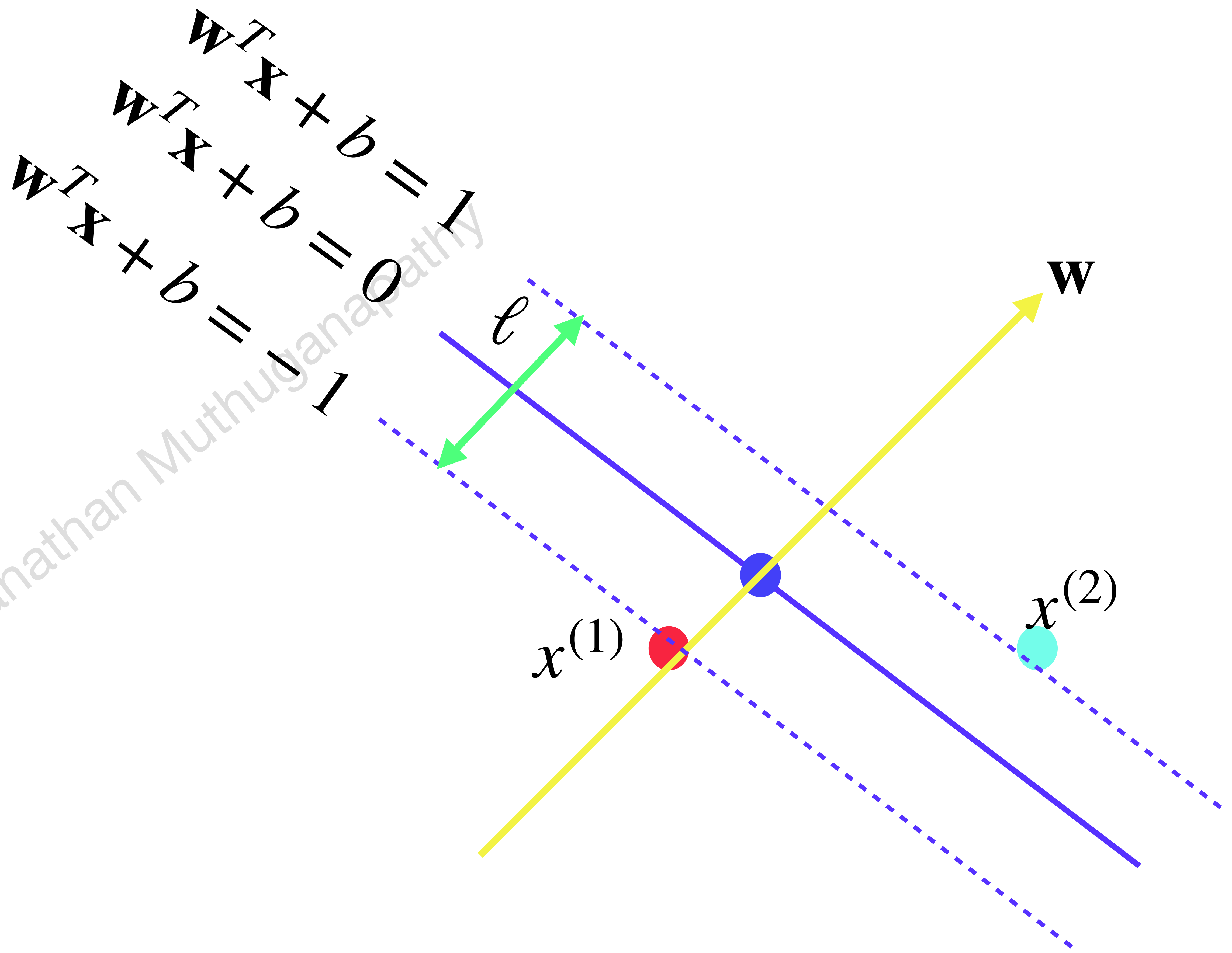
- $\mathbf{w}^T.(x^{(1)} - x^{(2)})$
- $\mathbf{w}^T.(x^{(1)} - x^{(2)}) = \ell \|\mathbf{w}\|$



Support vector machine

Putting together

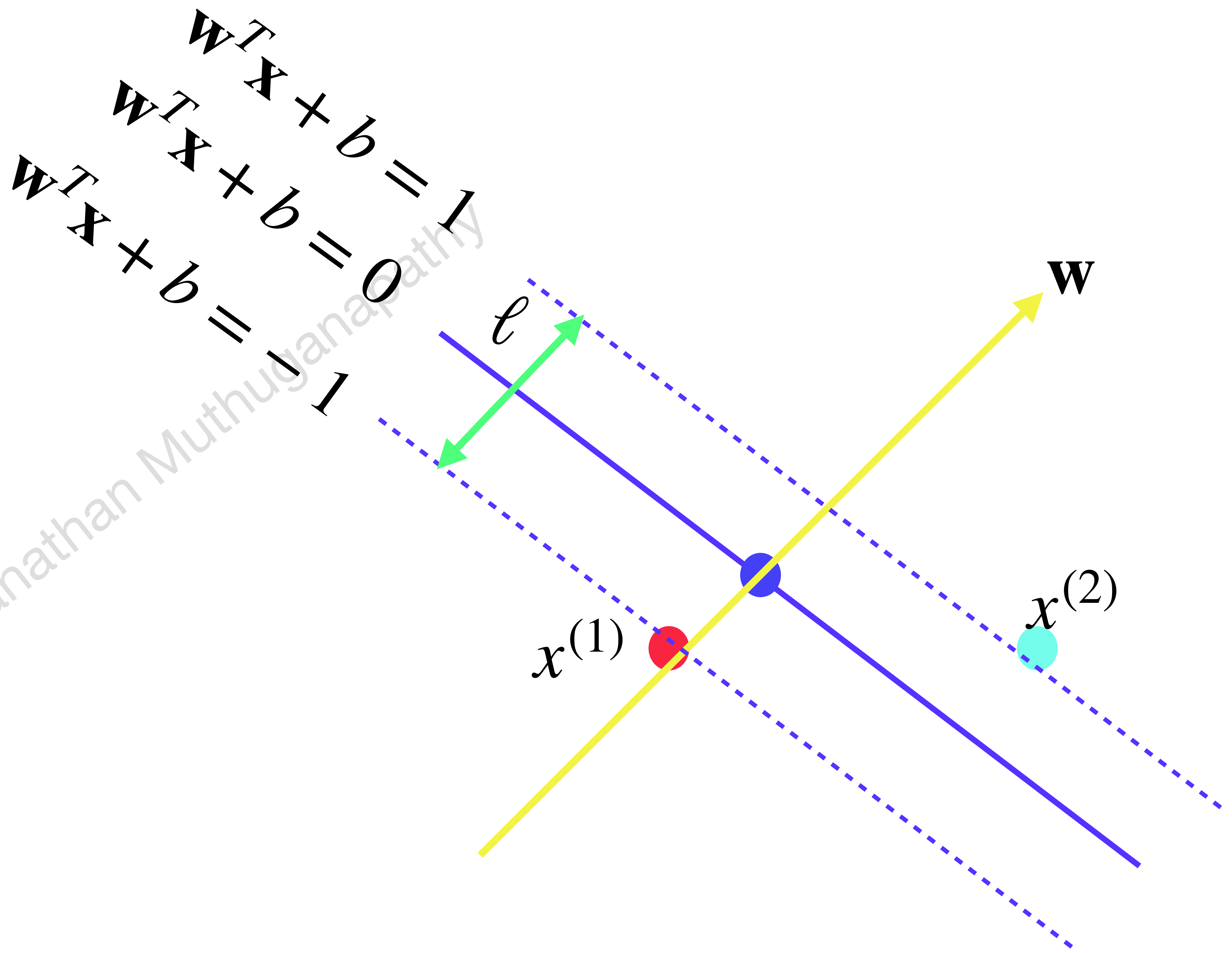
- $\mathbf{w}^T.(x^{(1)} - x^{(2)})$
- $\mathbf{w}^T.(x^{(1)} - x^{(2)}) = \ell \|\mathbf{w}\|$
- $(1 - b) - (-1 - b) = \ell \|\mathbf{w}\|$



Support vector machine

Putting together

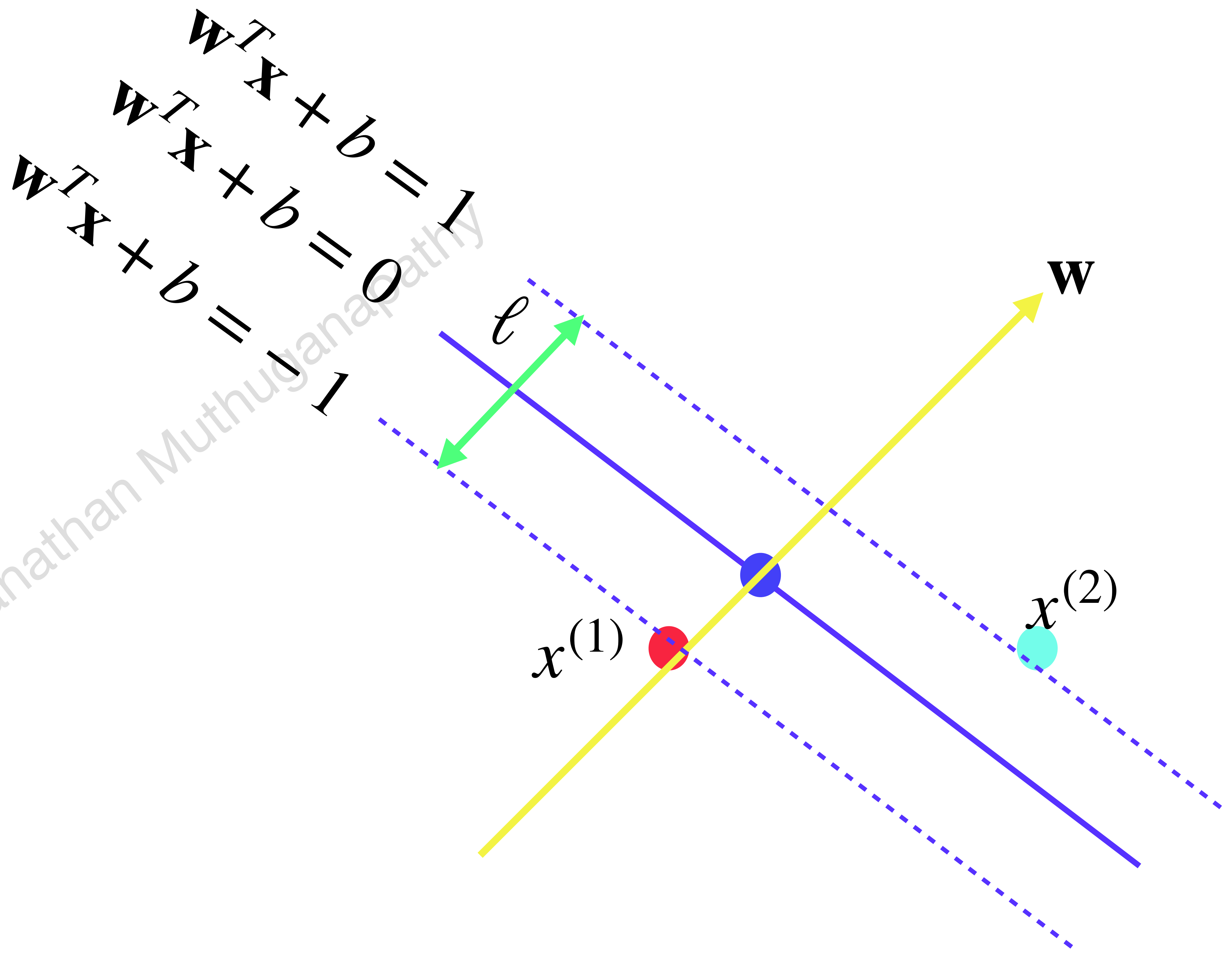
- $\mathbf{w}^T.(x^{(1)} - x^{(2)})$
- $\mathbf{w}^T.(x^{(1)} - x^{(2)}) = \ell \|\mathbf{w}\|$
- $(1 - b) - (-1 - b) = \ell \|\mathbf{w}\|$
- $2 = \ell \|\mathbf{w}\|$



Support vector machine

Putting together

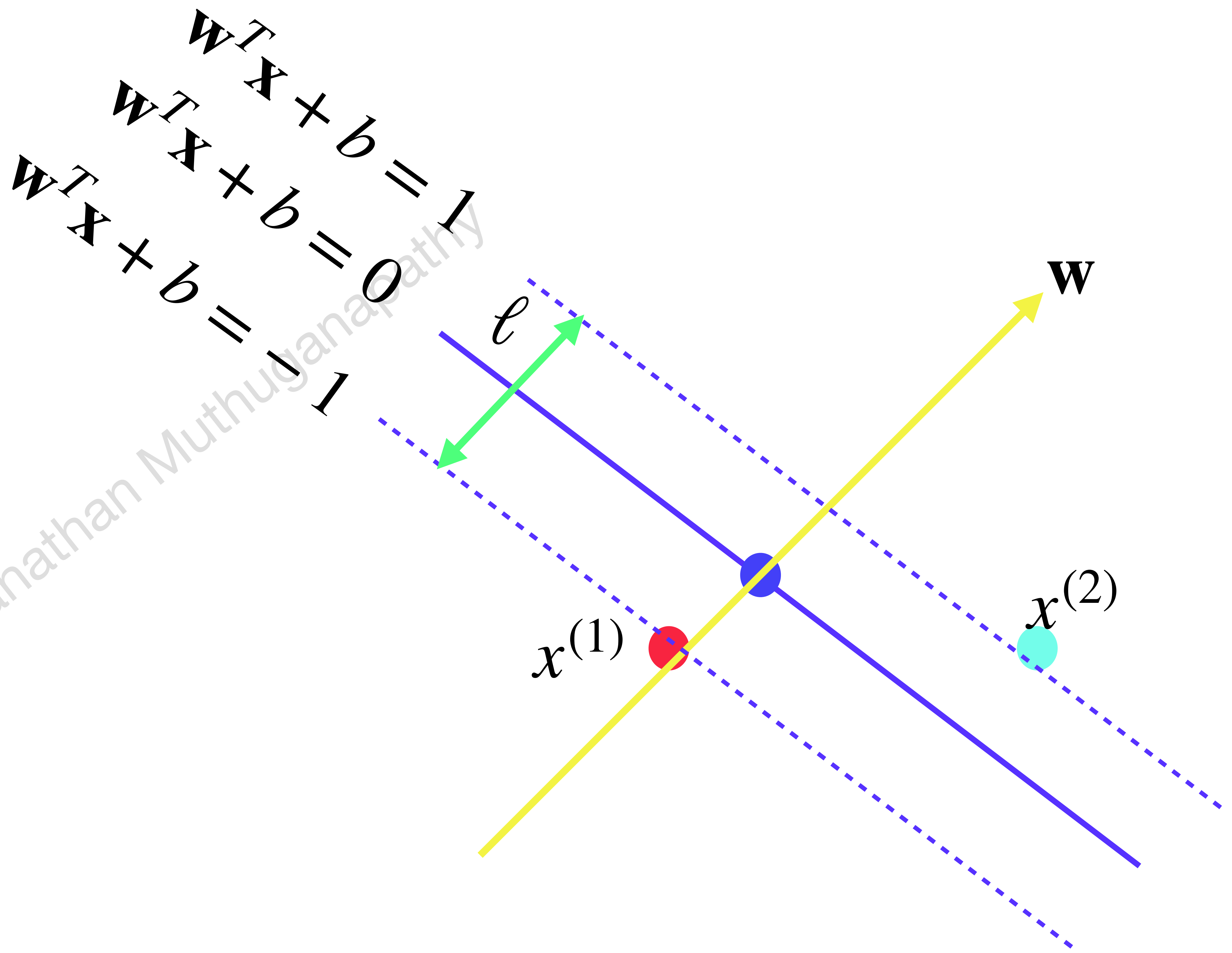
- $\mathbf{w}^T.(x^{(1)} - x^{(2)})$
- $\mathbf{w}^T.(x^{(1)} - x^{(2)}) = \ell \|\mathbf{w}\|$
- $(1 - b) - (-1 - b) = \ell \|\mathbf{w}\|$
- $2 = \ell \|\mathbf{w}\|$
- $\ell = \dfrac{2}{\|\mathbf{w}\|}$



Support vector machine

cost function

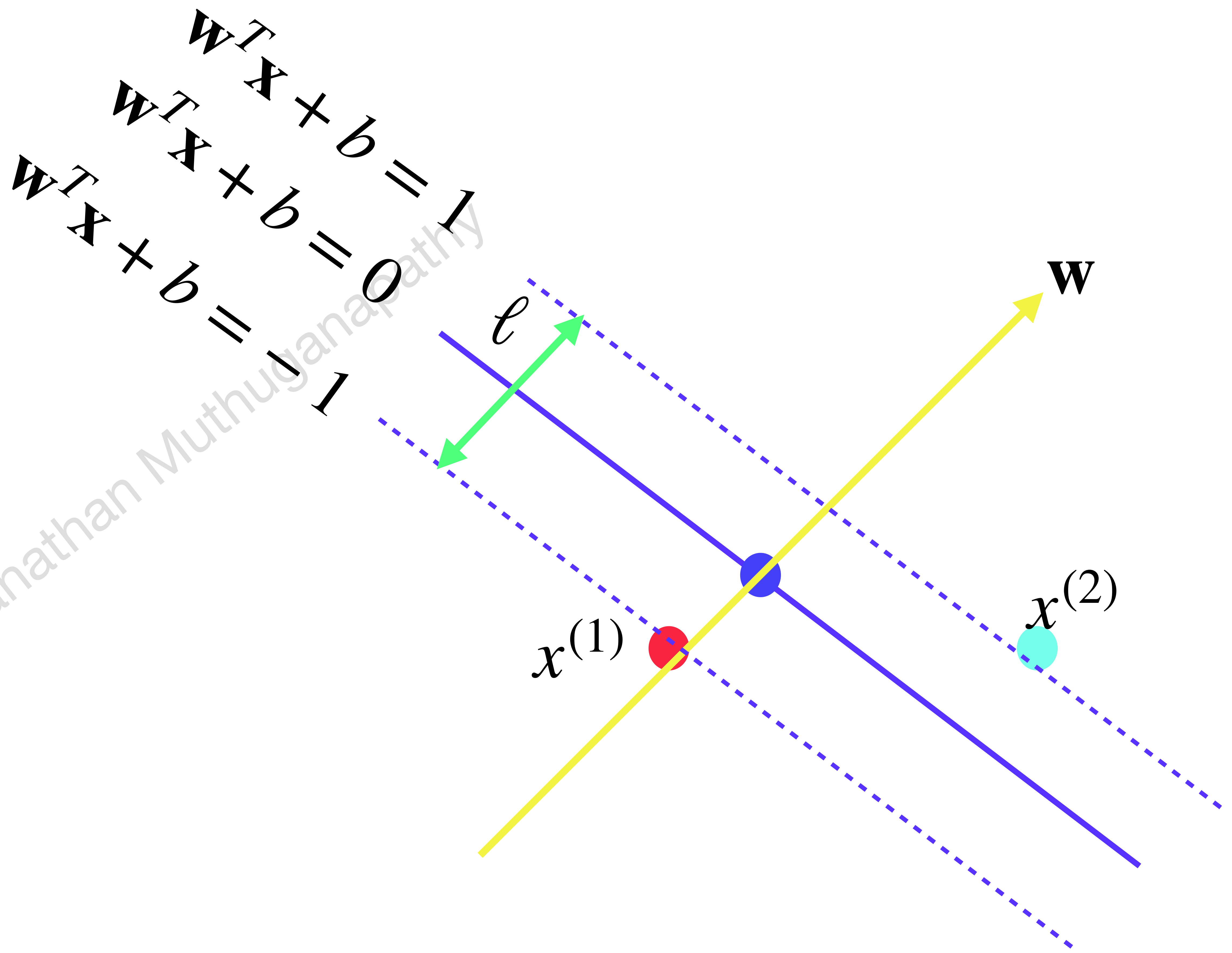
- $\ell = \dfrac{2}{\|\mathbf{w}\|^2}$



Support vector machine

Optimization

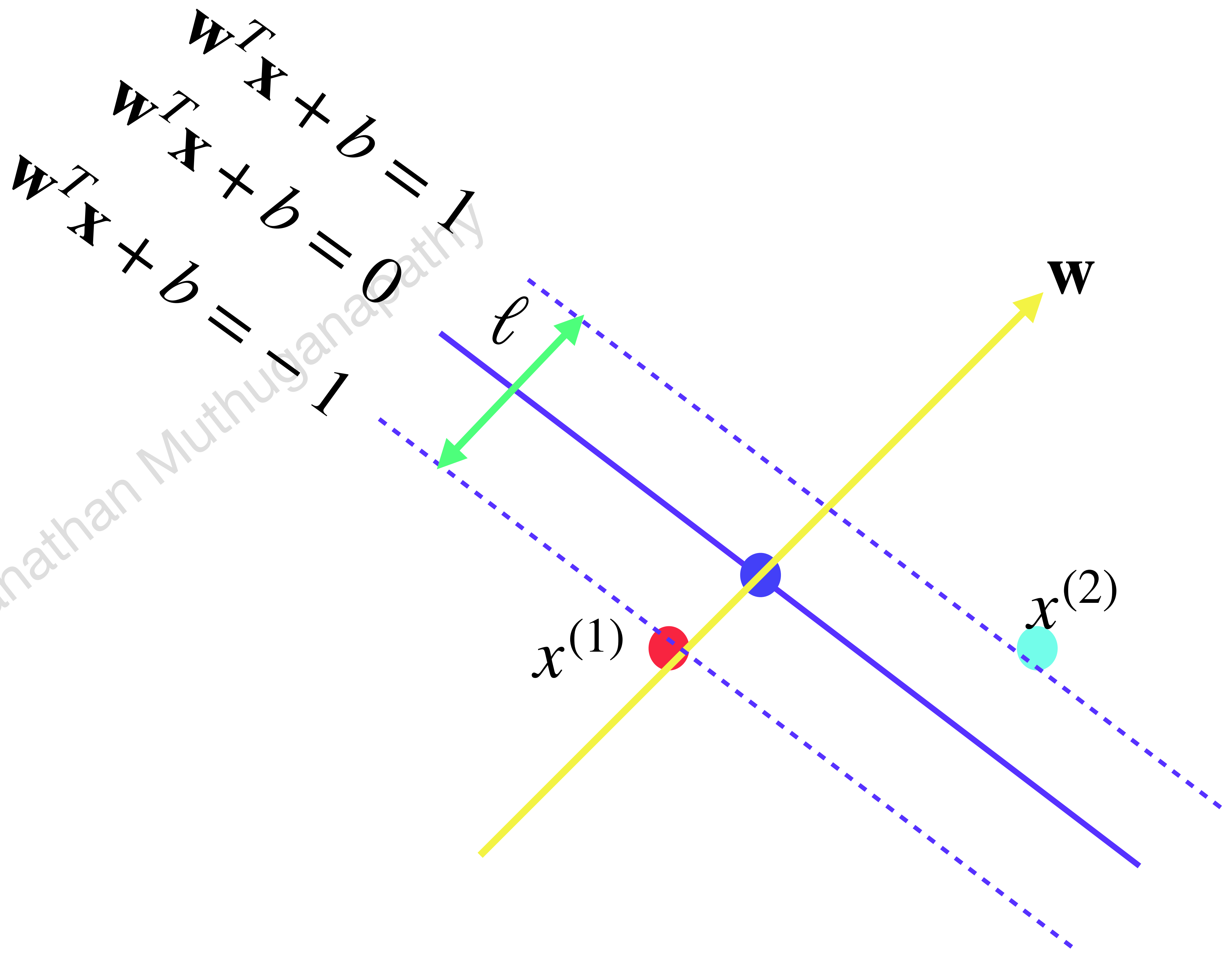
- $\max \ell = max \dfrac{2}{\|\mathbf{w}\|^2}$



Support vector machine

Optimization Objective

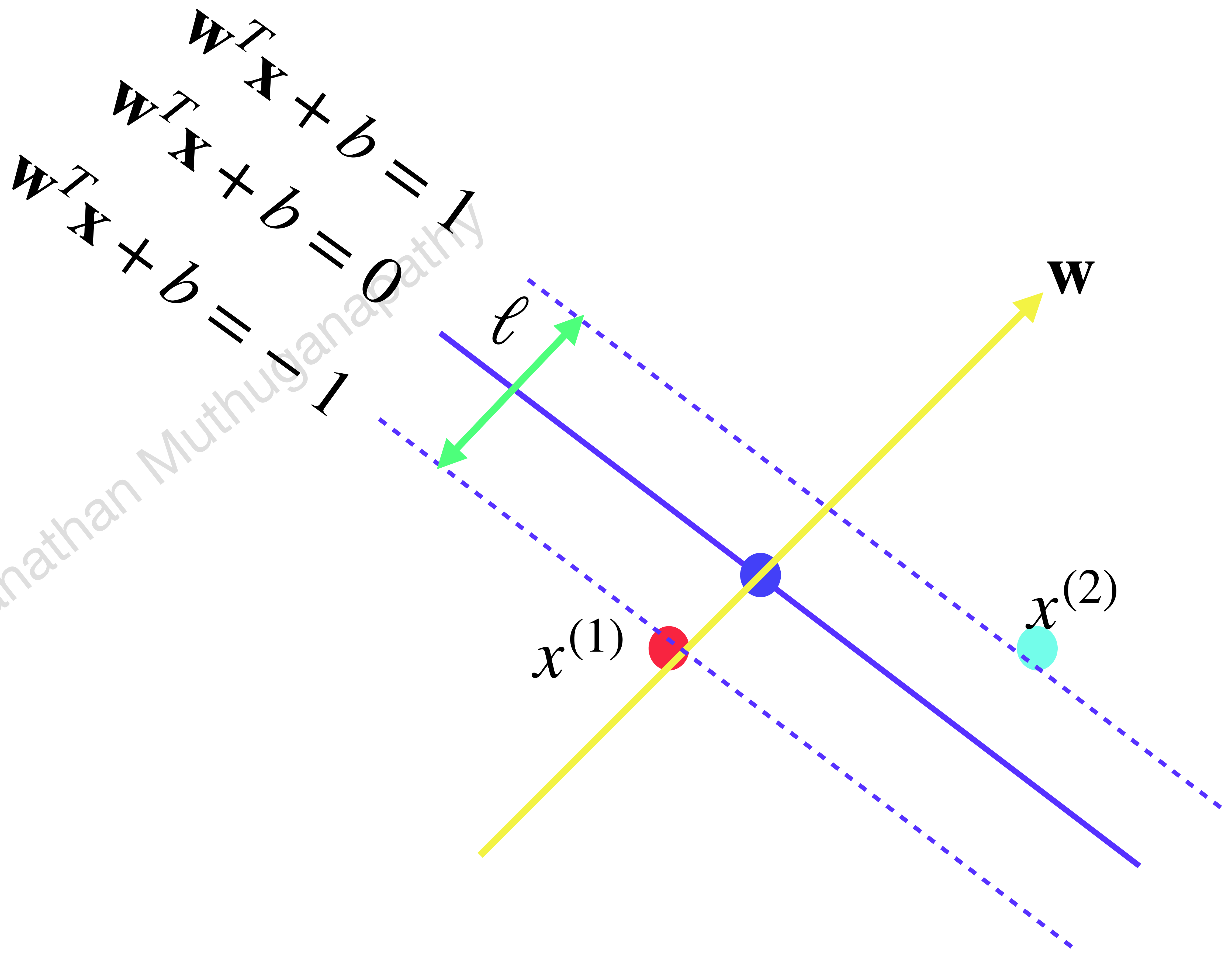
- $min \frac{\|\mathbf{w}\|^2}{2}$



Support vector machine

Optimization Objective

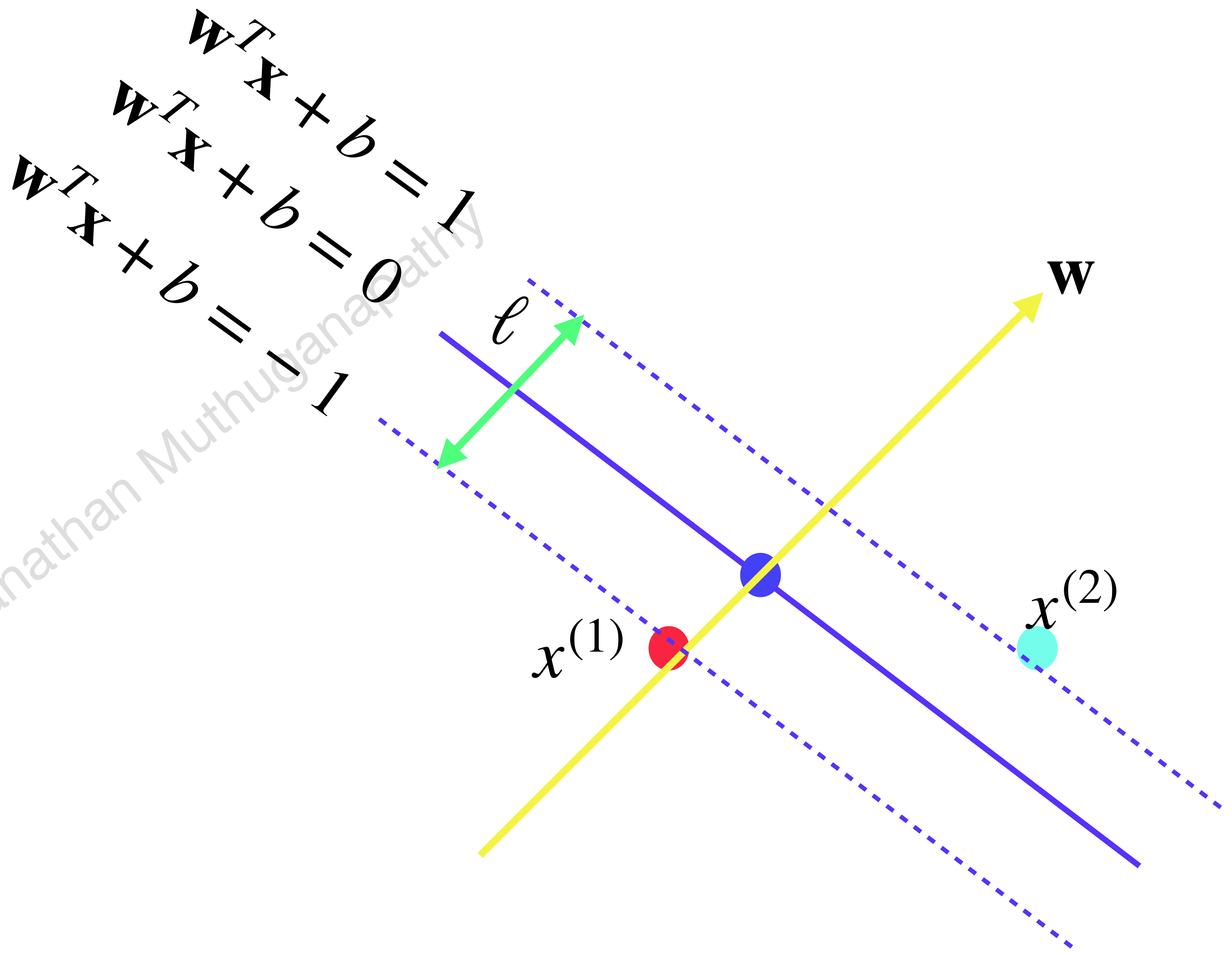
- $min \dfrac{\|\mathbf{w}\|^2}{2}$
- $\mathbf{w}^T\mathbf{x} + b \leq -1$
- $\mathbf{w}^T\mathbf{x} + b \geq -1$



Support vector machine

Optimization Objective

- $min \frac{\|\mathbf{w}\|^2}{2}$
- $1 - y^{(i)}(\mathbf{w}^T\mathbf{x} + b) \leq 0$



Support vector machine

Constrained Optimization Problem

- $min \dfrac{\|\mathbf{w}\|^2}{2}$
- $y^{(i)}(\mathbf{w}^T\mathbf{x} + b) - 1 \geq 0$