● 1. SUPPORT VECTOR MACHINE (SVM)

---

1 Problem Formulation

Input space:

$$\mathcal{X} \subseteq \mathbb{R}^d$$

Output space:

Binary classification:

$$\mathcal{Y} = \{-1, +1\}$$

Data distribution:

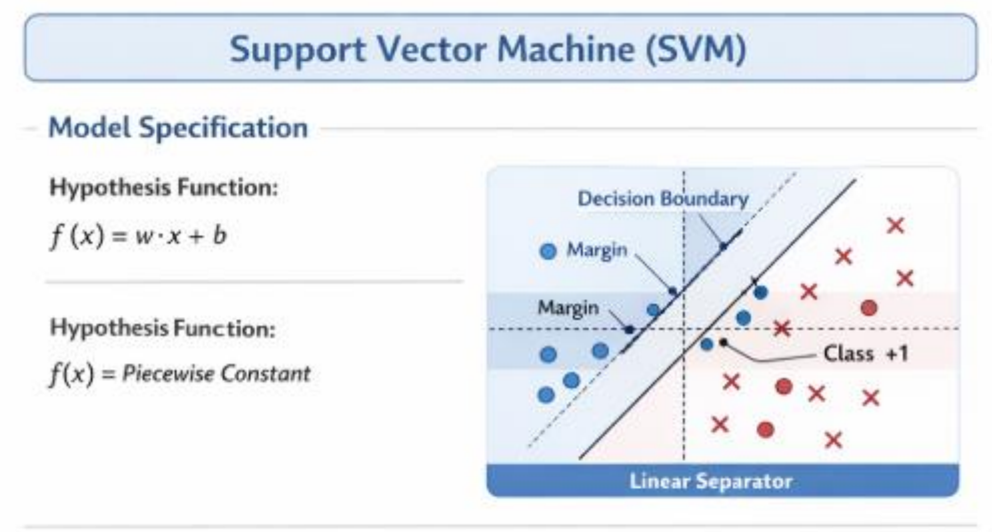
$$(x_i, y_i) \sim \mathcal{D}$$

Learning Objective (Expected Risk Minimization):

$$R(f) = \mathbb{E}_{(x,y)\sim\mathcal{D}}[\ell(y, f(x))]$$

---

2 Model Specification



Hypothesis Function

Linear SVM:

$$f(x) = w^T x + b$$

Kernel SVM:

$$f(x) = \sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b$$

Parameter Space

$$w \in \mathbb{R}^d, \quad b \in \mathbb{R}$$

Structural Assumptions

- Linear separability (hard margin)
- Non-linear separability via kernel trick
- Large margin principle

---

3 Loss Function

Hinge Loss

$$\ell(y, f(x)) = \max(0, 1 - y f(x))$$

Why Hinge Loss?

- Upper bound on 0–1 loss
- Encourages margin maximization

Convexity

- Convex but non-differentiable
- Global optimum guaranteed

---

4 Objective Function

Empirical Risk

$$\frac{1}{n}\sum_{i=1}^{n}\max(0, 1 - y_i(w^T x_i + b))$$

Regularized Form

$$\min_{w,b}\frac{1}{2}||w||^2 + C\sum_{i=1}^{n}\max(0, 1 - y_i f(x_i))$$

---

5 Optimization Method

- Quadratic Programming (QP)
- SMO algorithm (iterative)

Gradient (Subgradient)

If margin violated:

$$\nabla_w = -y_i x_i$$

Convergence

- Convex optimization → global optimum
- Complexity: O(n3)O(n^3)O(n3) worst case (QP)

---

6 Statistical Interpretation

- Equivalent to L2-regularized hinge risk minimization
- Can be interpreted as MAP under specific priors
- Output is margin distance (not probability)

---

7 Regularization & Generalization

- Controlled by CCC
- Large margin → lower VC dimension
- Bias–variance tradeoff via margin width

---

8 Theoretical Properties

- Convex optimization
- Unique global minimum
- Consistent under separability

9 Computational Complexity

Aspect	Complexity
Training	O(n3)O(n^3)O(n3)
Inference	O(sv)O(sv)O(sv)
Memory	O(n2)O(n^2)O(n2)

(sv = number of support vectors)

🔟 Limitations

- Not scalable for very large datasets
- No probabilistic output (unless calibrated)
- Sensitive to kernel choice

📊 SVM Margin Visualization

```
   +      +      +
        |  Margin  |
----------------------------  Decision boundary
   -      -      -
```