

Convex Optimization

Lab 8: Solve SVM via Quadratic Optimization

Lecturer: Dr. Wan-Lei Zhao

Autumn Semester 2025

Outline

- 1 Support Vector Machine: a review
- 2 Solve SVM by Quadratic Programming

Overview of discriminative classifier (1)

- Given a training set $(x_i, y_i)_{i=1 \dots m}$
- $x_i \in R^n$ is the observation, $y_i \in [-1, 1]$ is the class label
- A classifier is trained with this set
- Given a new instance $u \in R^n$
- The classifier makes the prediction whether it is -1 or 1

SVM: the model (1)

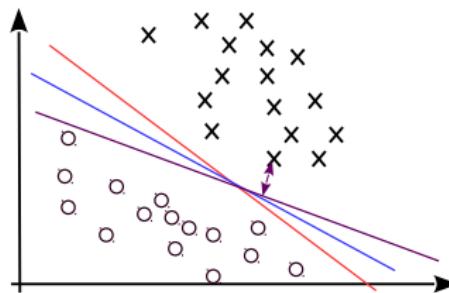


Figure: Searching for maximum γ

$$\gamma = \underset{i=1, \dots, m}{\operatorname{argmin}} \gamma^{(i)} \quad (1)$$

- Searching for w, b that maximizes γ

$$\begin{aligned} & \underset{w, b, \gamma}{\operatorname{Max.}} && \gamma \\ & \text{s. t.} && y^{(i)}(w^T x^{(i)} + b) \geq \gamma, i = 1, \dots, m \\ & && \|w\| = 1. \end{aligned}$$

SVM: the model (2)

- Searching for w, b that maximizes γ

$$\begin{aligned} \text{Max. } & \gamma \\ \text{s. t. } & y^{(i)}(w^T x^{(i)} + b) \geq \gamma, i = 1, \dots, m \\ & \|w\| = 1. \end{aligned} \tag{2}$$

- Unfortunately, above problem is not solvable
- Constraint $\|w\| = 1$ is not convex

$$\begin{aligned} \text{Max. } & \hat{\gamma}/\|w\| \\ \text{s. t. } & y^{(i)}(w^T x^{(i)} + b) \geq \hat{\gamma}, i = 1, \dots, m. \end{aligned} \tag{3}$$

- Unfortunately, above problem is not solvable either
- $\hat{\gamma}$ is functional margin, it is valid to scale it to $\hat{\gamma} = 1$

SVM: the model (3)

$$\begin{aligned} \text{Max.}_{w,b,\hat{\gamma}} \quad & \hat{\gamma}/\|w\| \\ \text{s. t.} \quad & y^{(i)}(w^T x^{(i)} + b) \geq \hat{\gamma}, i = 1, \dots, m. \end{aligned} \tag{3}$$

- $\hat{\gamma}$ is functional margin, it is valid to scale it to $\hat{\gamma} = 1$

$$\begin{aligned} \text{Max.}_{w,b} \quad & \frac{1}{\|w\|} \\ \text{s. t.} \quad & y^{(i)}(w^T x^{(i)} + b) \geq 1, i = 1, \dots, m. \end{aligned} \tag{4}$$

- This is equivalent to solving following **quadratic optimization** problem

$$\begin{aligned} \text{Min.}_{w,b} \quad & \frac{1}{2}\|w\|^2 \\ \text{s. t.} \quad & y^{(i)}(w^T x^{(i)} + b) \geq 1, i = 1, \dots, m. \end{aligned} \tag{5}$$

SVM: the model (4)

- Binary classification problem is now modeled as **quadratic optimization** problem

$$\begin{array}{ll} \text{Min.} & \frac{1}{2} ||w||^2 \\ w, b & \\ \text{s. t. } & 1 - y^{(i)}(w^T x^{(i)} + b) \leq 0, i = 1, \dots, m. \end{array} \quad (6)$$

- The unknowns are w and \mathbf{b}
- The target function is quadratic
- m constraints are linear
- In the lecture, we solve it with **Lagrange multiplier** method
- Here we solve it with **quadratic programming (QP)**

Outline

- 1 Support Vector Machine: a review
- 2 Solve SVM by Quadratic Programming

Quadratic Programming: standard form (1)

- Given the model for a problem:

$$\text{Min. } 2x_1^2 + x_2^2 - x_1x_2 - 4x_1 - 3x_2$$

s. t.
$$\begin{cases} x_1 + x_2 \leq 4 \\ -x_1 + 3x_2 \leq 3 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

$$H = \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}, c = [-4 \ -3]^T, A = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}, b = [4 \ 3]^T$$

$$Aeq = [], Beq = [], lb = [0 \ 0], ub = [].$$

- Standard Quadratic Optimization form:

$$\text{Min. } \frac{1}{2}x^T Hx + c^T x$$

$$\text{s. t. } Ax \leq b$$

$$Aeqx = Beq$$

$$lb \leq x \leq ub$$

- Solve it by Matlab: `[x, fval]=quadprog(H,c,A,b,Aeq,Beq,lb,ub)`

Quadratic Programming: standard form (2)

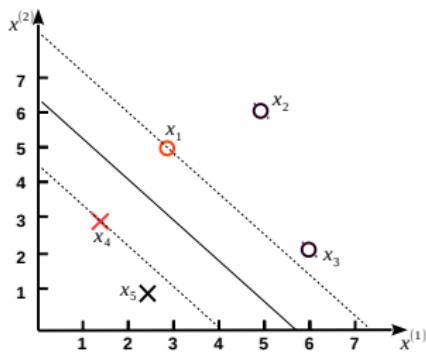
$$H = \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}, c = [-4 \ -3]^T, A = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}, b = [4 \ 3]^T$$

$$Aeq = [], Beq = [], lb = [0 \ 0], ub = [].$$

- Solve it by Matlab: **quadprog**

- ① $H=[4 \ -1; -1 \ 2];$
- ② $c=[-4 \ -3]';$
- ③ $A=[1 \ 1; -1 \ 3];$
- ④ $b=[4 \ 3]';$
- ⑤ $Aeq=[]; Beq=[];$
- ⑥ $lb=[0 \ 0]'; ub=[]$
- ⑦ $[x \ fval]=\text{quadprog}(H,c,A,b,Aeq,Beq,lb,ub);$

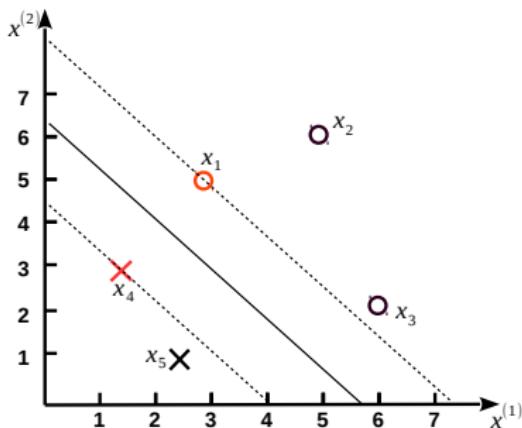
Simplified SVM: 2D case



$$\begin{aligned} \text{Min. } & \frac{1}{2} \|w\|^2 \\ \text{s. t. } & 1 - y^{(i)}(w^T x^{(i)} + b) \leq 0 \end{aligned} \quad (7)$$

- Positive: $x_1(3, 5); x_2(5, 6); x_3(6, 2.3)$
- Negative: $x_4(1.5, 3); x_5(2.5, 1)$
- **Hints:** [x fval]=quadprog(H,c,A,B,Aeq,Beq,lb,ub);
- In 2D case, $w = [w_1, w_2]$
- $y=1$ for positive example
- $y=-1$ for negative example
- Try to solve this problem with '**quadprog**'
- Display your result

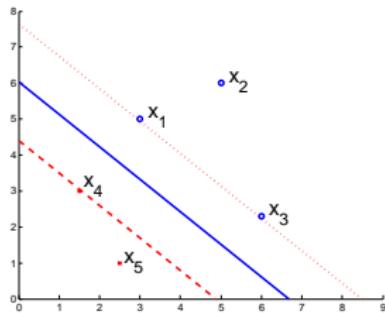
Simplified SVM: solution (1)



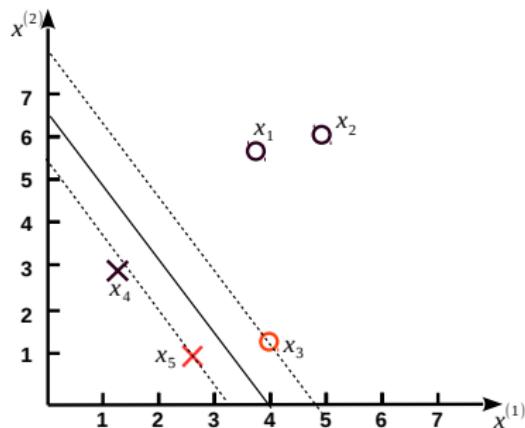
$$\begin{aligned}
 & \underset{w,b}{\text{Min.}} && \frac{1}{2}(w_1^2 + w_2^2) \\
 & \text{s.t.} && -3w_1 - 5w_2 - b \leq -1 \\
 & && -5w_1 - 6w_2 - b \leq -1 \\
 & && -6w_1 - 2.3w_2 - b \leq -1 \\
 & && 1.5w_1 + 3w_2 + b \leq -1 \\
 & && 2.5w_1 + w_2 + b \leq -1
 \end{aligned} \tag{8}$$

- Positive: $x_1(3, 5); x_2(5, 6); x_3(6, 2.3)$
- Negative: $x_4(1.5, 3); x_5(2.5, 1)$
- Hints:** [x fval]=quadprog(H, c, A, B, Aeq, Beq, lb, ub);

Simplified SVM: result



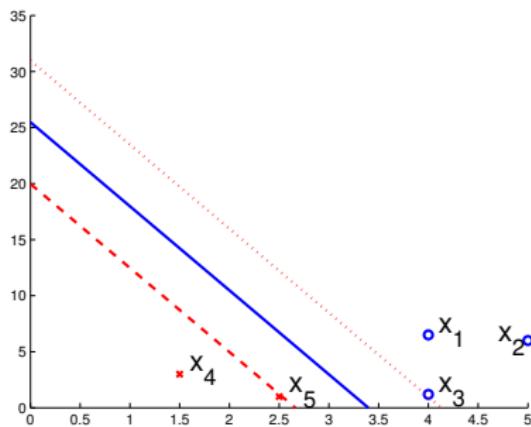
Change training points, see what happens



$$\begin{array}{ll} \text{Min.} & \frac{1}{2}(w_1^2 + w_2^2) \\ \text{s. t.} & -4w_1 - 6.5w_2 - b \leq -1 \\ & -5w_1 - 6w_2 - b \leq -1 \\ & -4w_1 - 1.2w_2 - b \leq -1 \\ & 1.5w_1 + 3w_2 + b \leq -1 \\ & 2.5w_1 + w_2 + b \leq -1 \end{array} \quad (9)$$

- Positive: $x_1(4, 6.5); x_2(5, 6); x_3(4, 1.2)$
- Negative: $x_4(1.5, 3); x_5(2.5, 1)$
- Hints:** [x fval]=quadprog(H, c, A, B, Aeq, Beq, lb, ub);

Simplified SVM: result updated



- This time x_3 and x_5 become the support vectors

Food for thought

- ① Are all the training points used to determine the splitting line?
- ② Is it possible to train the model with only negative or positive points?
- ③ The more training points the better?
- ④ Why in practice we do not solve SVM by QP?