

Outline

- Maximizing the margin
- The solution
- Nonlinear transforms

Constrained optimization

Minimize $\frac{1}{2} \mathbf{w}^T \mathbf{w}$

subject to $y_n (\mathbf{w}^T \mathbf{x}_n + b) \geq 1$ for $n = 1, 2, \dots, N$

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

Lagrange? inequality constraints \implies KKT

We saw this before

Remember regularization?

Minimize $E_{\text{in}}(\mathbf{w}) = \frac{1}{N} (\mathbf{Z}\mathbf{w} - \mathbf{y})^T (\mathbf{Z}\mathbf{w} - \mathbf{y})$

subject to: $\mathbf{w}^T \mathbf{w} \leq C$

∇E_{in} normal to constraint

optimize

constrain

Regularization:

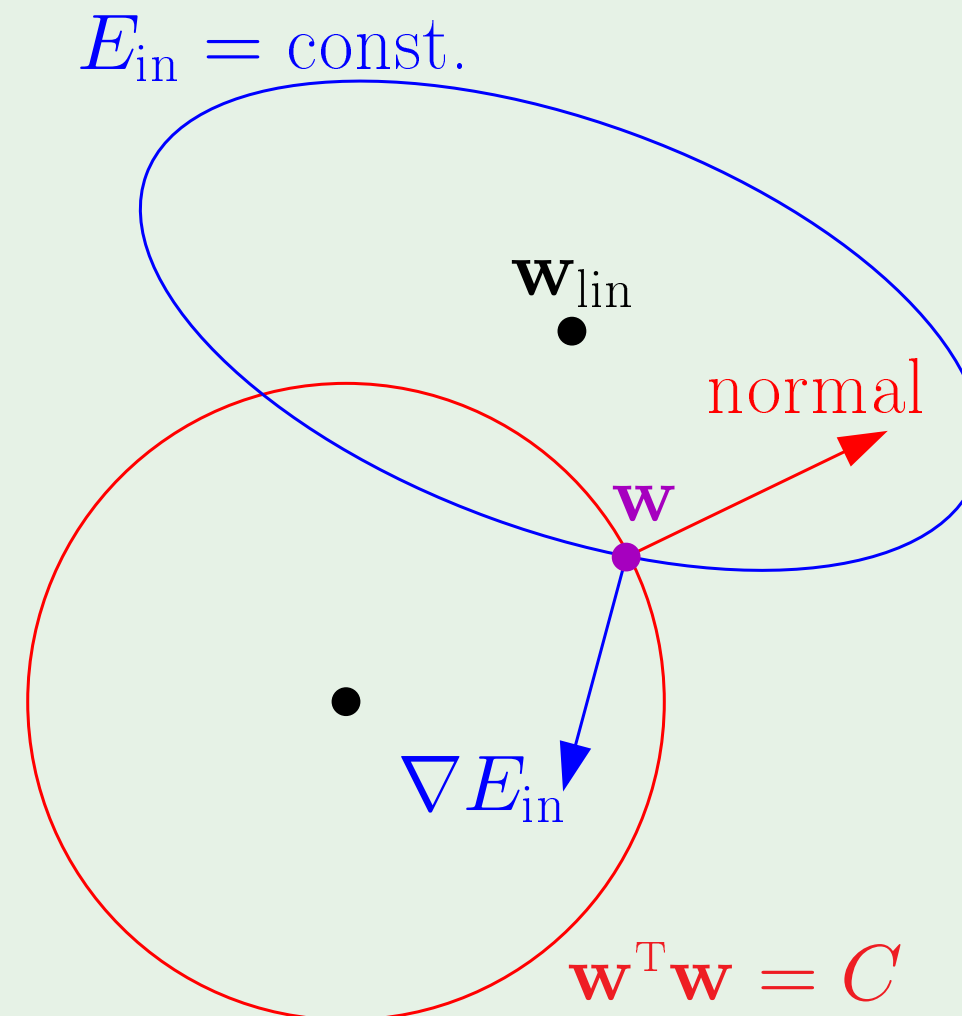
$$E_{\text{in}}$$

$$\mathbf{w}^T \mathbf{w}$$

SVM:

$$\mathbf{w}^T \mathbf{w}$$

$$E_{\text{in}}$$



Lagrange formulation

$$\text{Minimize } \mathcal{L}(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} - \sum_{n=1}^N \alpha_n (y_n (\mathbf{w}^T \mathbf{x}_n + b) - 1)$$

w.r.t. \mathbf{w} and b and maximize w.r.t. each $\alpha_n \geq 0$

$$\nabla_{\mathbf{w}} \mathcal{L} = \mathbf{w} - \sum_{n=1}^N \alpha_n y_n \mathbf{x}_n = \mathbf{0}$$

$$\frac{\partial \mathcal{L}}{\partial b} = - \sum_{n=1}^N \alpha_n y_n = 0$$

Substituting ...

$$\mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{x}_n \quad \text{and} \quad \sum_{n=1}^N \alpha_n y_n = 0$$

in the Lagrangian

$$\mathcal{L}(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} - \sum_{n=1}^N \alpha_n (y_n (\mathbf{w}^T \mathbf{x}_n + b) - 1)$$

we get

$$\mathcal{L}(\boldsymbol{\alpha}) = \sum_{n=1}^N \alpha_n - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N y_n y_m \alpha_n \alpha_m \mathbf{x}_n^T \mathbf{x}_m$$

Maximize w.r.t. to $\boldsymbol{\alpha}$ subject to $\alpha_n \geq 0$ for $n = 1, \dots, N$ and $\sum_{n=1}^N \alpha_n y_n = 0$

The solution - quadratic programming

$$\min_{\alpha} \frac{1}{2} \alpha^T \underbrace{\begin{bmatrix} y_1 y_1 \mathbf{x}_1^T \mathbf{x}_1 & y_1 y_2 \mathbf{x}_1^T \mathbf{x}_2 & \dots & y_1 y_N \mathbf{x}_1^T \mathbf{x}_N \\ y_2 y_1 \mathbf{x}_2^T \mathbf{x}_1 & y_2 y_2 \mathbf{x}_2^T \mathbf{x}_2 & \dots & y_2 y_N \mathbf{x}_2^T \mathbf{x}_N \\ \dots & \dots & \dots & \dots \\ y_N y_1 \mathbf{x}_N^T \mathbf{x}_1 & y_N y_2 \mathbf{x}_N^T \mathbf{x}_2 & \dots & y_N y_N \mathbf{x}_N^T \mathbf{x}_N \end{bmatrix}}_{\text{quadratic coefficients}} \alpha + \underbrace{(-\mathbf{1}^T)}_{\text{linear}} \alpha$$

subject to

$$\underbrace{\mathbf{y}^T \alpha = 0}_{\text{linear constraint}}$$

$$\underbrace{\mathbf{0}}_{\text{lower bounds}} \leq \alpha \leq \underbrace{\infty}_{\text{upper bounds}}$$

QP hands us α

Solution: $\alpha = \alpha_1, \dots, \alpha_N$

$$\implies \mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{x}_n$$

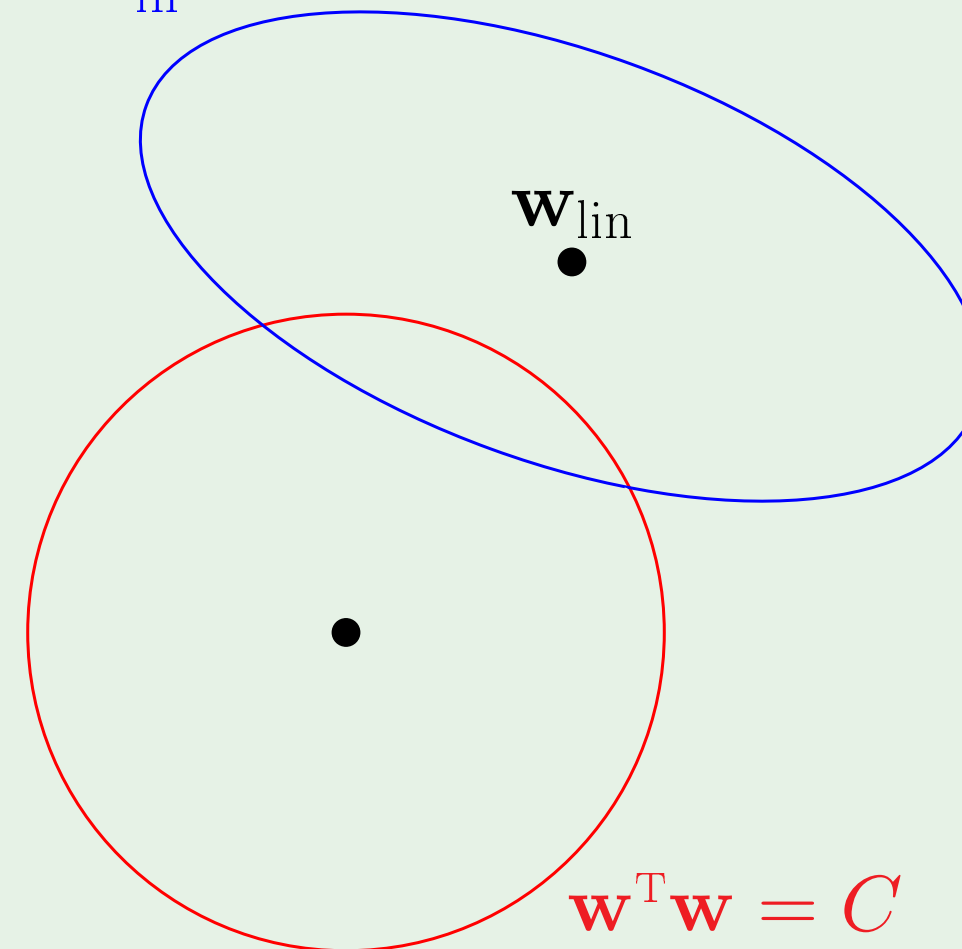
KKT condition: For $n = 1, \dots, N$

$$\alpha_n (y_n (\mathbf{w}^T \mathbf{x}_n + b) - 1) = 0$$

We saw this before!

$\alpha_n > 0 \implies \mathbf{x}_n$ is a support vector

$E_{\text{in}} = \text{const.}$



Support vectors

Closest \mathbf{x}_n 's to the plane: achieve the margin

$$\implies y_n (\mathbf{w}^\top \mathbf{x}_n + b) = 1$$

$$\mathbf{w} = \sum_{\mathbf{x}_n \text{ is SV}} \alpha_n y_n \mathbf{x}_n$$

Solve for b using any SV:

$$y_n (\mathbf{w}^\top \mathbf{x}_n + b) = 1$$

