Tuesday, March 24, 2020 2:30 PM

ECE 523
Homework 3

Rafael Antonio Garcia Mar
RafaelGarcia
03/24/2020

Part A: Theory

Support Vector Machines

ℓ_2-norm soft margin SVM

The optimization problem is:

$$\underset{\vec{w}, b, \xi_i}{\arg\min} \ \frac{1}{2}\vec{w}^T\vec{w} + \frac{C}{2}\sum_{i=1}^{n} \xi_i^2$$

$$\text{s.t.} \quad y_i(\vec{w}^T\vec{x}_i + b) \geq 1 - \xi_i \quad , i = 1, \dots, n$$

We use Lagrange multipliers

$$L = \frac{1}{2}\vec{w}^T\vec{w} + \frac{C}{2}\sum_{i=1}^{n} \xi_i^2 - \sum_{i=1}^{n} \alpha_i\left[y_i(\vec{w}^T\vec{x}_i + b) - 1 + \xi_i\right]$$

$$\frac{\partial L}{\partial \vec{w}} = \vec{w} - \sum^{n} \alpha_i y_i \vec{x}_i = 0 \Rightarrow \vec{w} = \sum^{n} \alpha_i y_i \vec{x}_i$$

$$\frac{\partial L}{\partial \vec{w}^T} = \vec{w} - \sum_{i=1}^{n} \alpha_i y_i \vec{x}_i = 0 \quad \Rightarrow \quad \vec{w} = \sum_{i=1}^{n} \alpha_i y_i \vec{x}_i$$

$$\frac{\partial L}{\partial b} = -\sum_{i=1}^{n} \alpha_i y_i = 0 \quad \Rightarrow \quad \sum_{i=1}^{n} \alpha_i y_i = 0$$

$$\frac{\partial L}{\partial \xi_i} = C\xi_i - \alpha_i = 0 \quad \Rightarrow \quad \xi_i = \frac{\alpha_i}{C}$$

Substitution in the expression for L yields

$$L = \frac{1}{2}\sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j \vec{x}_i^T \vec{x}_j + \frac{C}{2}\sum_{i=1}^{n} \frac{\alpha_i^2}{C^2} - \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j \vec{x}_i^T \vec{x}_j$$
$$- b\cancelto{0}{\sum_{i=1}^{n} \alpha_i y_i} + \sum_{i=1}^{n} \alpha_i - \sum_{i=1}^{n} \alpha_i \frac{\alpha_i}{C}$$

$$L = -\frac{1}{2}\sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j \vec{x}_i^T \vec{x}_j + \sum_{i=1}^{n} \alpha_i - \frac{1}{2C}\sum_{i=1}^{n} \alpha_i^2$$

s.t. $\alpha_i \geq 0 \ , \ i = 1, \ldots, n$

$$\sum_{i=1}^{n} \alpha_i y_i = 0$$

$\boxed{\text{Support Vector Machines (Revisited)}}$

Domain adaptation SVM

The optimization problem is

$$\underset{\vec{w}_T, b, \xi}{\operatorname{argmin}} \ \frac{1}{2}\vec{w}_T^T \vec{w}_T + C\sum_{i=1}^{n} \xi_i - B\, \vec{w}_T^T \vec{w}_S$$

s.t. $y_i(\vec{w}_T^T \vec{x}_i + b) \geq 1 - \xi_i \ , \ i = 1, \ldots, n$

$\quad \xi_i \geq 0 \ , \ i = 1, \ldots, n$

We use Lagrange multipliers

$$L = \frac{1}{2}\vec{w}_T^T\vec{w}_T + C\sum_{i=1}^{n} \xi_i - B\vec{w}_T^T\vec{w}_s - \sum_{i=1}^{n} \alpha_i\left[y_i(\vec{w}_T^T\vec{x}_i + b) - 1 + \xi_i\right] - \sum_{i=1}^{n} \mu_i\xi_i$$

$$\frac{\partial L}{\partial \vec{w}_T} = \vec{w}_T - B\vec{w}_s - \sum_{i=1}^{n}\alpha_i y_i \vec{x}_i = 0 \implies \vec{w}_T = \sum_{i=1}^{n}\alpha_i y_i \vec{x}_i + B\vec{w}_s$$

$$\frac{\partial L}{\partial b} = -\sum_{i=1}^{n}\alpha_i y_i = 0 \implies \sum_{i=1}^{n}\alpha_i y_i = 0$$

$$\frac{\partial L}{\partial \xi_i} = C - \alpha_i - \mu_i = 0 \implies C = \alpha_i + \mu_i \implies \alpha_i \geq 0$$

Substitution in the expression for L yields

$$\begin{aligned} L = {} & \frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j \vec{x}_i^T\vec{x}_j + \frac{1}{2}B^2\vec{w}_s^T\vec{w}_s + B\sum_{i=1}^{n}\alpha_i y_i \vec{w}_s^T\vec{x}_i \\ & + \sum_{i=1}^{n}(\alpha_i \cancel{+\mu_i})\xi_i - B\sum_{i=1}^{n}\alpha_i y_i\vec{w}_s^T\vec{x}_i - B^2\vec{w}_s^T\vec{w}_s \\ & - \sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j\vec{x}_i^T\vec{x}_j - B\sum_{i=1}^{n}\alpha_i y_i\vec{w}_s^T\vec{x}_i - b\cancelto{0}{\sum_{i=1}^{n}\alpha_i y_i} + \sum_{i=1}^{n}\alpha_i - \sum_{i=1}^{n}\cancel{\alpha_i}\xi_i \\ & - \sum_{i=1}^{n}\cancel{\mu_i}\xi_i \end{aligned}$$

$$L = \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j\vec{x}_i^T\vec{x}_j - B\sum_{i=1}^{n}\alpha_i y_i\vec{w}_s^T\vec{x}_i \underbrace{-\frac{1}{2}B^2\vec{w}_s^T\vec{w}_s}_{\text{can be ignored}}$$

s.t. $0 \leq \alpha_i \leq C$, $i = 1, \dots, n$

$$\sum_{i=1}^{n}\alpha_i y_i = 0$$