# #

SVM

using slack variable

$$\xi_i = 1 - y_i(w^T x_i - b)$$

or

$$\min_{w,b} \sum_{i=1}^{m} \max\left(0, 1 - y_i(w^T x_i + b)\right) + \lambda \|w\|^2$$

Soln

We begin by taking the minimization function $\forall_i$ as

$$\Rightarrow \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{n} \xi_i \qquad \forall_i, \; i = 1, \ldots n$$

which is subject to

SVM

using slack variable

$$\xi_i = 1 - y_i(w^T x_i - b)$$

on $\min_{w,b} \sum_{i=1}^{m} \max\left(0, 1 - y_i(w^T x_i + b)\right) + \lambda \|w\|^2$

Soln

We begin by taking the minimization function $\forall_i$ as

$$\Rightarrow \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \xi_i \qquad \forall_i, \; i = 1, \ldots n$$

which is subject to

$$\Rightarrow \xi_i = 1 - y_i(w^T x_i - b) \qquad \forall_i \quad i = 1, \ldots, n$$

Thus for $\xi_i \geq 0$, we have the above written as

$$1 - \xi_i \leq y_i(w^T x_i - b) \qquad \forall_i \quad i = 1, \ldots n$$

We begin by taking the minimization function $\forall_i$ as

$$\Rightarrow \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n} \xi_i \qquad \forall_i, \ i=1,\dots n$$

which is subject to

$$\Rightarrow \xi_i = 1 - y_i(\omega^T x_i - b) \qquad \forall_i \quad i=1,\dots,n$$

$\Rightarrow$ Thus for $\xi_i \geq 0$, we have the above written as

$$1 - \xi_i \leq y_i(\omega^T x_i - b) \qquad \forall_i \quad i=1,\dots n$$

$\Rightarrow$ In this case the opt variable are ω, b, ξ_i

Next we look at all possible deviations for Change in $\omega^T x$ and b as follows

$$\omega^T x + b = 1$$
$$\omega^T x + b = 0$$

subject to

$\Rightarrow \xi_i = 1 - y_i(w^T x_i - b) \quad \forall_i \quad i = 1, \dots, n$

$\Rightarrow$ Thus for $\xi_i \geq 0$, we have the above written as

$$1 - \xi_i \leq y_i(w^T x_i - b) \quad \forall_i \quad i = 1, \dots n$$

$\Rightarrow$ In this case the opt variable are w, b, ξ_i

$\Rightarrow$ Next we look at all possible deviations for Change in $w^T x$ and b as follows

$$w^T x + b = 1$$
$$w^T x + b = 0$$
$$w^T x + b = -1$$

$\Rightarrow$ We let $w^T x + b = 0$, then the optimization problem will now become

$$\min_{w, b, \xi_i} \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{n} \xi_i$$

which is subject to

$\Rightarrow \xi_i = 1 - y_i(\omega^T x_i - b) \quad \forall_i \quad i = 1, \dots, n$

$\Rightarrow$ Thus for $\xi_i \geq 0$, we have the above written as

$$1 - \xi_i \leq y_i(\omega^T x_i - b) \quad \forall_i \quad i = 1, \dots n$$

$\Rightarrow$ In this case the opt variable are ω, b, ξ_i

$\Rightarrow$ Next we look at all possible deviations for Change in $\omega^T x$ and b as follows

$$\omega^T x + b = 1$$
$$\omega^T x + b = 0$$
$$\omega^T x + b = -1$$

$\Rightarrow$ We let $\omega^T x + b = 0$, then the optimization problem will now become

$$\min_{\omega, b, \xi_i} \frac{1}{2}||\omega||^2 + C \sum_{i=1}^{n} \xi_i$$

which is subject to

$$1 - y_i(w^T x_i - b) - \xi_i \leq 0 \quad \forall_i \quad i = 1, \ldots, n$$

with $-\xi_i \leq 0 \quad \forall_i$

=) Next we compute the associated langriagian L in relation to the optimization problem as follows

$$L(\mu, \lambda, \xi, w, b) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \xi_i + \sum_{i=1}^{n} \mu_i (1 - \xi_i - y_i(w^T x_i - b)) - \sum_{i=1}^{n} \lambda_i \xi_i$$

where

μ_i are the langrage multipliers for the separability constraint

=> Next we compute the associated langriagian L in relation to the optimization problem as follows

$$L(\mu, \lambda, \xi, \omega, b) = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n} \xi_i + \sum_{i=1}^{n} \mu_i (1 - \xi_i - y_i(\omega^T x_i - b))$$
$$- \sum_{i=1}^{n} \lambda_i \xi_i$$

where

=> μ_i are the langrage multipliers for the separability constraint

=> λ_i are the langrage multiplier for the constraints $-\xi_i \leq 0$

=> We can now take the derivative of the above langragian function w.r.t to both ξ_i and b

=> $\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1} \mu_i y_i = 0$

$$- \sum_{i=1}^{n} \lambda_i \xi_i$$

where

$\Rightarrow$ α_i are the language multipliers for the separability constraint

$\Rightarrow$ λ_i are the language multipliers for the constraints $-\xi_i \leq 0$

$\Rightarrow$ We can now take the derivative, of the above langragian function w.r.t to both ξ_i and b

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

$$\Rightarrow \frac{\partial L}{\partial \xi_i} = 0 \Rightarrow \alpha_i \lambda_i = C \quad \forall_i \quad i = 1, \ldots, n$$

note that $-\xi_i - y_i(w^T x_i - b) + 1 \leq 0$; $\xi_i \geq 0$ $\forall_i$

$\Rightarrow$ Using the complementary slackness ... in interv...

$\Rightarrow$ u_i are the language multipli...
constraint

$\Rightarrow$ λ_i are the language multipliers for the constraints
$-\xi_i \leq 0$

$\Rightarrow$ We can now take the derivative, of the above langragian function w.r.t to both ξ_i and b

$$\Rightarrow \frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1} u_i y_i = 0$$

$$\Rightarrow \frac{\partial L}{\partial \xi_i} = 0 \Rightarrow u_i \lambda_i = C \quad \forall_i \quad i = 1, \ldots, n$$

note that $-\xi_i - y_i(w^T x_i - b) + 1 \leq 0$; $\xi_i \geq 0$ $\forall_i$

$\Rightarrow$ Using the complementary slackness condition in interval $0 \leq u_i < C$

i.e

$1 - y_i(w^T x_i - b) = 0$, we need to find the dual function for all μ_i which is given as

$$q(\mu,\lambda) = \inf_{\omega, b, \xi_i} L(\mu,\lambda,\xi,\omega,b)$$

=> But from the langrangian function we have the term

$$\sum_i (C - \mu_i - \lambda_i)\xi_i$$

Taking the infimum w.r.t ξ_i and imposi[ng] on the in[...] the condition $C = \mu_i + \lambda_i \quad \forall_i$

the dual function for all μ_i which is given by as

$$q(\mu,\lambda) = \inf_{w,b,\xi_i} L(\mu,\lambda,\xi,w,b)$$

$\Rightarrow$ But from the langragian function we have the term

$$\sum_i (C-\mu_i-\lambda_i)\xi_i$$

$\Rightarrow$ Taking the infimum w.r.t ξ_i and imposing the condition $C = \mu_i + \lambda_i \quad \forall_i$ on the interval $0 \le \mu_i \le C$, the dual problem becomes

$$\max_{\mu} q(\mu) = \sum_{i=1}^{n} \mu_i - \frac{1}{2} \sum_{i,j=1}^{n} \mu_i \mu_j y_i y_j x_i^T x_j$$

Subject to $0 \le \mu_i \le C \quad \forall_i \quad i = 1 \ldots n$

term

$$\sum_i (C - \mu_i - \lambda_i)\, \xi_i$$

$(w^T x_i - b)$

$\Rightarrow$ Taking the infimum w.r.t ξ_i and imposing the condition $C = \mu_i + \lambda_i \ \forall_i$ on the interval $0 \le \mu_i \le C$, the dual problem becomes

$$\max_{\mu} \ q(\mu) = \sum_{i=1}^{n} \mu_i - \frac{1}{2} \sum_{i,j=1}^{n} \mu_i \mu_j y_i y_j x_i^T x_j$$

Subject to $0 \le \mu_i \le C \quad \forall_i \ i = 1, \ldots, n$

and

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

$\Rightarrow$ Associated primal problem is given by

$$\text{minimize} \ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i$$

Subject to

=> Taking the infimum w.r.t ξ_i and imposing the condition $C = \mu_i + \lambda_i$ $\forall i$ on the interval $0 \le \mu_i \le C$, the dual problem becomes

$$\max_{\mu} \; q(\mu) = \sum_{i=1}^{n} \mu_i - \frac{1}{2} \sum_{i,j=1}^{n} \mu_i \mu_j y_i y_j x_i^T x_j$$

Subject to $0 \le \mu_i \le C \quad \forall i \; i = 1, \dots, n$

and

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

=> Associated primal problem is given by

$$\text{minimize} \quad \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{n} \xi_i$$

Subject to

$$1 - \xi_i \le y_i (\omega^T x_i - b) \quad \forall i \; i = 1, \dots, n$$