Formulas from Support Vector Machines Document

1 Distance from Hyperplane

$$d(x_i) = \frac{|z_i(w^T x_i + b)|}{\|w\|} \tag{1}$$

2 Maximum Margin Hyperplane

$$w^*, b^* = \arg\max_{w,b} \min_{i \in \{1 \dots n\}} \frac{|z_i(w^T x_i + b)|}{\|w\|} \tag{2}$$

subject to $z_i(w^T x_i + b) > 0$ for all i.

3 Final Primal Optimization Formulation

$$\arg\min_{w,b} \frac{1}{2}\|w\|^2 \quad \text{subject to } z_i(w^T x_i + b) \geq 1, \forall i \tag{3}$$

4 Dual SVM Problem

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j z_i z_j x_i^T x_j \tag{4}$$

subject to $\alpha_i \geq 0$ and $\sum_{i=1}^{n} \alpha_i z_i = 0$.

5 Soft Margin SVM Primal Formulation

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{n} \xi_i \tag{5}$$

subject to $z_i(w^T x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \forall i.$

Or, equivalently:

$$\min_{\boldsymbol{w},b} \frac{1}{2}\|\boldsymbol{w}\|^2 + C \sum_{i=1}^{n} \max\left[0, 1 - z_i(\boldsymbol{w}^T \mathbf{x}_i + b)\right]$$

6 Hinge Loss Function

$$f(s) = \max(0, 1 - s) \tag{6}$$

Alternative notation:

$$f(s) = [1 - s]_+ \tag{7}$$

7 Soft Margin SVM Dual Formulation

$$\max_{\boldsymbol{\alpha}} L_D(\boldsymbol{\alpha}) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j z_i z_j \mathbf{x}_i^T \mathbf{x}_j$$

s.t.

$$0 \leq \alpha_i \leq C, \quad i = 1, \ldots, n$$

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

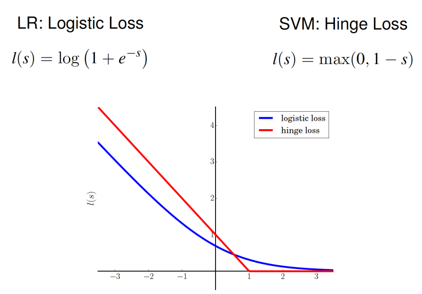
8 SVM vs LR

SVM Objective:

$$\min_{\boldsymbol{w}, b} \frac{\lambda}{2} \|\boldsymbol{w}\|^2 + \frac{1}{n} \sum_{i=1}^{n} \max\left[0, 1 - z_i(\boldsymbol{w}^T \mathbf{x}_i + b)\right]$$

Logistic Regression Objective:

$$\min_{\boldsymbol{w}, b} \frac{\lambda}{2} \|\boldsymbol{w}\|^2 + \frac{1}{n} \sum_{i} \log\left[1 + e^{-z_i(\boldsymbol{w}^T \mathbf{x}_i + b)}\right]$$



9 Kernelized Dual SVM Problem

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j z_i z_j k(x_i, x_j) \tag{8}$$

where $k(x_i, x_j) = \Phi(x_i)^T \Phi(x_j)$.

10 Kernelized Scoring

$$s(\mathbf{x}_t) = \sum_{i=1|\alpha_i>0} \alpha_i z_i \Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}_t) + b = \sum_{i=1|\alpha_i>0} \alpha_i z_i k(\mathbf{x}_i, \mathbf{x}_t) + b$$

11 Radial Basis Function (RBF) Kernel

$$k(x_1, x_2) = e^{-\gamma \|x_1 - x_2\|^2} \tag{9}$$

12 Polynomial Kernel

$$k(x_1, x_2) = (x_1^T x_2 + 1)^d \tag{10}$$