Duality Forumlation of Max-Margin Classifier*

Arash Rouhani

March 1, 2011

Abstract

In this fake-paper I write a short mathematical explenation of how to formulate the Max-Margin classifier in it's dual form. I go through it step by step.

1 Primal formulation

To be written ...

2 Lagrangian function

There will be as always 3 kinds of terms, coming from either the *the objective function*,*a constraint* or *a variable constraint*.

$$L(w, b, \xi, \mu, \nu) = \frac{1}{2}||w||_2^2 + \sum_{i=1}^{n} C\xi_i + \mu_i(y_i b + 1 - y_i w^T x_i - \xi_i) - \nu_i \xi_i \quad (1)$$

We now derivativs for each variable-type and get an equation. There are 3 'types' of variables here, so we will get 3 equations.

2.1 Derivative with respect to w

$$\frac{dL}{dw} = \mathbf{0} = w - \sum_{i=1}^{n} \mu_i y_i x_i \quad (2)$$

2.2 Derivative with respect to b

$$\frac{dL}{db} = 0 = \sum_{i=1}^{n} \mu_i y_i \quad (3)$$

*This is the first time I use latex

2.3 Derivative with respect to ξ

$$\frac{dL}{db} = 0 = C - \mu_i - \nu_i \tag{4}$$

3 Reducing the lagrangian by substitution

First we substitute (4) which is clearly eliminating the ξ terms.

$$L(w, b, \mu) = \frac{1}{2}||w||_2^2 + \sum_{i=1}^{n} \mu_i(y_i b + 1 - y_i w^T x_i) \tag{5}$$

Let's rewrite the function using $||w||_2^2 = w^T w$ and other simple rewritings.

$$L(w, b, \mu) = w^T(\frac{1}{2}w - \sum_{i=1}^{n} \mu_i y_i x_i) + b\sum_{i=1}^{n} \mu_i y_i + \sum_{i=1}^{n} \mu_i \tag{6}$$

Now we use (2) and (3) in order to reduce once more.

$$L(w, \mu) = \sum_{i=1}^{n} \mu_i - \frac{1}{2}w^T \sum_{i=1}^{n} \mu_i y_i x_i \tag{7}$$

Now we use (2) again to rewrite w^T, finally eliminating all free varaibles we had in the primal.

$$L(\mu) = \sum_{i=1}^{n} \mu_i - \frac{1}{2}\sum_{j=1}^{n} \mu_j y_j x_j^T \sum_{i=1}^{n} \mu_i y_i x_i \tag{8}$$

Now we finally rewrite the lagrangian to it's well-recognized form.

$$L(\mu) = \sum_{i=1}^{n} \mu_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mu_i \mu_j y_i y_j x_i^T x_j \tag{9}$$