Modern Multivariate Statistical Techniques

Jinhong Du,15338039

November 26, 2018

Content

1 Ex 11.3

The "2-norm soft margin" optimization problem for SVM classification: Consider the regularization problem of minimizing $\frac{1}{2}\|\beta\|^2 + C\sum_{i=1}^{n}\xi_i^2$ subject to the constraints $y_i(\beta_0 + \mathbf{x}_i^\top \beta) \geq 1 - \xi_i$, and $\xi_i \geq 0$, for $i = 1, 2, \cdots, n$.

(a) Show that the same optimal solution to this problem is reached if we remove the constraints $\xi_i \geq 0$, $i = 1, 2, \cdots, n$, on the slack variables. (Hint: What is the effect on the objective functional if this constraint is violated?)

(b) Form the primal Lagrangian F_P, which will be a function of β_0, $\boldsymbol{\beta}$, $\boldsymbol{\xi}$, and the Lagrangian multipliers $\boldsymbol{\alpha}$. Differentiate F_P wrt β_0, $\boldsymbol{\beta}$, and $\boldsymbol{\xi}$, set the results equal to zero, and solve for a stationary solution.

(c) Substitute the results from (b) into the primal Lagrangian to obtain the dual objective functional F_D. Write out the dual problem (objective functional and constraints) in matrix notation. Maximize the dual wrt $\boldsymbol{\alpha}$. Use the Karush − Kuhn − Tucker complementary conditions $\alpha_i\{y_i(\beta_0 + \mathbf{x}_i^\top \boldsymbol{\beta})(1\xi_i)\} = 0$ for $i = 1, 2, \cdots, n$.

(d) If $\boldsymbol{\alpha}^*$ is the solution to the dual problem, find $\hat{\boldsymbol{\beta}}$ and its norm, which gives the width of the margin.