

Foundations of ML – Additional material for lesson 7

1. Proof of wavelet-based feature importance

If $a_i \leq b_i, \forall i$, then $\min a_i \leq \min b_i$. Otherwise, assume that for some j , we have $b_j < a_j, \forall i$. Then, this is a contradiction to $a_j \leq b_j$.

2. Further details for Support Vector Machine formulation (12.3) \Leftrightarrow (12.4)

$$\max_{\beta, \beta_0, \|\beta\|=1} M \quad \text{s.t.} \quad y_i (\langle x_i, \beta \rangle + \beta_0) \geq M, \quad i = 1, \dots, N \quad \Leftrightarrow$$

$$\max_{\beta, \beta_0, \|\beta\|=1} M \quad \text{s.t.} \quad y_i \left(\left\langle x_i, \frac{\beta}{M} \right\rangle + \frac{\beta_0}{M} \right) \geq 1, \quad i = 1, \dots, N \quad \Leftrightarrow_{\tilde{\beta} = \beta/M, \tilde{\beta}_0 = \beta_0/M}$$

$$\min_{\tilde{\beta}, \tilde{\beta}_0} \|\tilde{\beta}\| \quad \text{s.t.} \quad y_i (\langle x_i, \tilde{\beta} \rangle + \tilde{\beta}_0) \geq 1, \quad i = 1, \dots, N \quad \Leftrightarrow$$

$$\min_{\tilde{\beta}, \tilde{\beta}_0} \|\tilde{\beta}\|^2 \quad \text{s.t.} \quad y_i (\langle x_i, \tilde{\beta} \rangle + \tilde{\beta}_0) \geq 1, \quad i = 1, \dots, N.$$