Appendix E: The proof of Theorem 1

Proof. Let us consider Problem (6) directly, rewrite it as following

$$\begin{aligned}\min_{\mathbf{W},b,\boldsymbol{\zeta}} \quad & \frac{1}{2}\mathrm{tr}(\mathbf{W}^T\mathbf{W}) + C\sum_{i=1}^{n}\zeta_i + \tau||\mathbf{W}||_* \qquad (19)\\ \text{s.t.} \quad & y_i[\mathrm{tr}(\mathbf{W}^T\mathbf{X}_i) + b] \geq 1 - \zeta_i, \quad i = 1, \ldots, n\\ & \zeta_i \geq 0, \quad i = 1, \ldots, n.\end{aligned}$$

We construct the following Lagrangian function for (19)

$$\begin{aligned} & Q(\mathbf{W}, b, \boldsymbol{\zeta}, \boldsymbol{\beta}, \boldsymbol{\mu}) \qquad (20)\\ = \; & \frac{1}{2}\mathrm{tr}(\mathbf{W}^T\mathbf{W}) + C\sum_{i=1}^{n}\zeta_i + \tau||\mathbf{W}||_*\\ & - \sum_{i=1}^{n}\beta_i\{y_i[\mathrm{tr}(\mathbf{W}^T\mathbf{X}_i) + b] - 1 + \zeta_i\} - \sum_{i=1}^{n}\mu_i\zeta_i.\end{aligned}$$

Letting the first-order derivative of Q with respect to ζ_i be 0, we have

$$\mu_i = C - \beta_i, \quad i = 1, \ldots, n. \qquad (21)$$

Substituting (21) in (20) to eliminate r_i and ξ_i, we obtain

$$\begin{aligned} Q(\mathbf{W}, \boldsymbol{\alpha}) \; = \; & \frac{1}{2}\mathrm{tr}(\mathbf{W}^T\mathbf{W}) + \tau||\mathbf{W}||_*\\ & - \sum_{i=1}^{n}\alpha_i[y_i\mathrm{tr}(\mathbf{W}^T\mathbf{X}_i) - 1]. \qquad (22)\end{aligned}$$

Since $\mu_i \geq 0$, β_i should satisfy the following condition

$$0 \leq \beta_i \leq C. \qquad (23)$$

By KKT condition (Borwein & Lewis, 2010), we also have **0** should be in the subgradient of Q when the optimal solution is taken:

$$\mathbf{0} \in \partial Q(\tilde{\mathbf{W}}, \tilde{\boldsymbol{\beta}})\Big|_{\mathbf{W}=\tilde{\mathbf{W}}}. \qquad (24)$$

Based on Eqn. (1), we have

$$\tilde{\mathbf{W}} = \mathcal{D}_\tau\Big(\sum_{i=1}^{n}\tilde{\beta}_i y_i \mathbf{X}_i\Big),$$

where $\tilde{\boldsymbol{\beta}}$ is the corresponding value of Lagrangian multiplier when $\tilde{\mathbf{W}}$ is the optimal solution. □

$$S^* = \operatorname*{argmax}_S E(S) \qquad (25)$$

References

Borwein, Jonathan M and Lewis, Adrian S. *Convex analysis and nonlinear optimization: theory and examples*, volume 3. Springer Science & Business Media, 2010.