The first and the fourth equation are identical to the Equations $(*_1)$ and $(*_2)$ that we obtained in Example 50.10. Since $\lambda, \mu, \alpha, \beta \geq 0$, the second and the third equation are equivalent to the box constraints

$$0 \leq \lambda_i, \mu_j \leq K, \quad i = 1, \ldots, p, \; j = 1, \ldots, q.$$

Using the equations that we just derived, after simplifications we get

$$G(\lambda, \mu, \alpha, \beta) = -\frac{1}{2} \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} X^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} + \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} \mathbf{1}_{p+q},$$

which is independent of α and β and is identical to the dual function obtained in $(*_4)$ of Example 50.10. To be perfectly rigorous,

$$G(\lambda, \mu) = \begin{cases} -\frac{1}{2} \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} X^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} + \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} \mathbf{1}_{p+q} & \text{if } \begin{cases} \sum_{i=1}^p \lambda_i = \sum_{j=1}^q \mu_j \\ 0 \leq \lambda_i \leq K, \; i = 1, \ldots, p \\ 0 \leq \mu_j \leq K, \; j = 1, \ldots, q \end{cases} \\ -\infty & \text{otherwise.} \end{cases}$$

As in Example 50.10, the the dual program can be formulated as

$$\begin{aligned} &\text{maximize} \quad -\frac{1}{2} \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} X^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} + \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} \mathbf{1}_{p+q} \\ &\text{subject to} \\ &\qquad \sum_{i=1}^p \lambda_i - \sum_{j=1}^q \mu_j = 0 \\ &\qquad 0 \leq \lambda_i \leq K, \quad i = 1, \ldots, p \\ &\qquad 0 \leq \mu_j \leq K, \quad j = 1, \ldots, q, \end{aligned}$$

or equivalently

Dual of Soft margin SVM (SVM_{s2}):

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2} \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} X^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} - \begin{pmatrix} \lambda^\top & \mu^\top \end{pmatrix} \mathbf{1}_{p+q} \\ &\text{subject to} \\ &\qquad \sum_{i=1}^p \lambda_i - \sum_{j=1}^q \mu_j = 0 \\ &\qquad 0 \leq \lambda_i \leq K, \quad i = 1, \ldots, p \\ &\qquad 0 \leq \mu_j \leq K, \quad j = 1, \ldots, q. \end{aligned}$$

If (w, ϵ, ξ, b) is an optimal solution of Problem (SVM_{s2}), then the complementary slackness conditions yield a classification of the points u_i and v_j in terms of the values of λ and μ.