(1) Since the constraints are affine and the objective function is convex, by Theorem 50.19(2) the duality gap is zero, so for any minimum w of $J(w,b) = (1/2)w^\top w$ and any maximum (λ,μ) of G, we have

$$J(w,b) = \frac{1}{2}w^\top w = G(\lambda,\mu).$$

But by $(*_1)$

$$w = -X\begin{pmatrix}\lambda\\ \mu\end{pmatrix} = \sum_{i=1}^p \lambda_i u_i - \sum_{j=1}^q \mu_j v_j,$$

so

$$\begin{pmatrix}\lambda^\top & \mu^\top\end{pmatrix} X^\top X \begin{pmatrix}\lambda\\ \mu\end{pmatrix} = w^\top w,$$

and we get

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

so

$$w^\top w = \begin{pmatrix}\lambda^\top & \mu^\top\end{pmatrix}\mathbf{1}_{p+q} = \sum_{i=1}^p \lambda_i + \sum_{j=1}^q \mu_j,$$

which yields

$$G(\lambda,\mu) = \frac{1}{2}\left(\sum_{i=1}^p \lambda_i + \sum_{j=1}^q \mu_j\right).$$

The above formulae are stated in Vapnik [182] (Chapter 10, Section 1).

(2) It is instructive to compute the Lagrangian of the dual program and to derive the KKT conditions for this Lagrangian.

The conditions $\lambda \geq 0$ being equivalent to $-\lambda \leq 0$, and the conditions $\mu \geq 0$ being equivalent to $-\mu \leq 0$, we introduce Lagrange multipliers $\alpha \in \mathbb{R}^p_+$ and $\beta \in \mathbb{R}^q_+$ as well as a multiplier $\rho \in \mathbb{R}$ for the equational constraint, and we form the Lagrangian

$$\begin{aligned} L(\lambda,\mu,\alpha,\beta,\rho) = {} & \frac{1}{2}\begin{pmatrix}\lambda^\top & \mu^\top\end{pmatrix} X^\top X \begin{pmatrix}\lambda\\ \mu\end{pmatrix} - \begin{pmatrix}\lambda^\top & \mu^\top\end{pmatrix}\mathbf{1}_{p+q} \\ & - \sum_{i=1}^p \alpha_i\lambda_i - \sum_{j=1}^q \beta_j\mu_j + \rho\Big(\sum_{j=1}^q \mu_j - \sum_{i=1}^p \lambda_i\Big). \end{aligned}$$

It follows that the KKT conditions are

$$X^\top X\begin{pmatrix}\lambda\\ \mu\end{pmatrix} - \mathbf{1}_{p+q} - \begin{pmatrix}\alpha\\ \beta\end{pmatrix} + \rho\begin{pmatrix}-\mathbf{1}_p\\ \mathbf{1}_q\end{pmatrix} = 0_{p+q}, \qquad (*_4)$$