

Since

$$(w^\top \quad \epsilon^\top \quad \xi^\top \quad b) C^\top \begin{pmatrix} \lambda \\ \mu \\ \alpha \\ \beta \end{pmatrix} = (w^\top \quad \epsilon^\top \quad \xi^\top \quad b) \begin{pmatrix} X & 0_{n,p+q} \\ -I_{p+q} & -I_{p+q} \\ \mathbf{1}_p^\top & -\mathbf{1}_q^\top \\ 0_{p+q}^\top & 0_{p+q}^\top \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \\ \alpha \\ \beta \end{pmatrix},$$

we get

$$\begin{aligned} (w^\top \quad \epsilon^\top \quad \xi^\top \quad b) C^\top \begin{pmatrix} \lambda \\ \mu \\ \alpha \\ \beta \end{pmatrix} &= (w^\top \quad \epsilon^\top \quad \xi^\top \quad b) \begin{pmatrix} X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \\ -\begin{pmatrix} \lambda + \alpha \\ \mu + \beta \end{pmatrix} \\ \mathbf{1}_p^\top \lambda - \mathbf{1}_q^\top \mu \end{pmatrix} \\ &= w^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} - \epsilon^\top (\lambda + \alpha) - \xi^\top (\mu + \beta) + b(\mathbf{1}_p^\top \lambda - \mathbf{1}_q^\top \mu), \end{aligned}$$

and since

$$(\mathbf{1}_{p+q}^\top \quad 0_{p+q}^\top) \begin{pmatrix} \lambda \\ \mu \\ \alpha \\ \beta \end{pmatrix} = \mathbf{1}_{p+q}^\top \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = (\lambda^\top \quad \mu^\top) \mathbf{1}_{p+q},$$

the Lagrangian can be rewritten as

$$\begin{aligned} L(w, \epsilon, \xi, b, \lambda, \mu, \alpha, \beta) &= \frac{1}{2} w^\top w + w^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} + \epsilon^\top (K \mathbf{1}_p - (\lambda + \alpha)) + \xi^\top (K \mathbf{1}_q - (\mu + \beta)) \\ &\quad + b(\mathbf{1}_p^\top \lambda - \mathbf{1}_q^\top \mu) + (\lambda^\top \quad \mu^\top) \mathbf{1}_{p+q}. \end{aligned}$$

To find the dual function $G(\lambda, \mu, \alpha, \beta)$ we minimize $L(w, \epsilon, \xi, b, \lambda, \mu, \alpha, \beta)$ with respect to w, ϵ, ξ and b . Since the Lagrangian is convex and $(w, \epsilon, \xi, b) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}$, a convex open set, by Theorem 40.13, the Lagrangian has a minimum in (w, ϵ, ξ, b) iff $\nabla L_{w, \epsilon, \xi, b} = 0$, so we compute its gradient with respect to w, ϵ, ξ and b , and we get

$$\nabla L_{w, \epsilon, \xi, b} = \begin{pmatrix} w + X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \\ K \mathbf{1}_p - (\lambda + \alpha) \\ K \mathbf{1}_q - (\mu + \beta) \\ \mathbf{1}_p^\top \lambda - \mathbf{1}_q^\top \mu \end{pmatrix}.$$

By setting $\nabla L_{w, \epsilon, \xi, b} = 0$ we get the equations

$$\begin{aligned} w &= -X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} & (*_w) \\ \lambda + \alpha &= K \mathbf{1}_p \\ \mu + \beta &= K \mathbf{1}_q \\ \mathbf{1}_p^\top \lambda &= \mathbf{1}_q^\top \mu. \end{aligned}$$