

The Lagrangian $L(w, \epsilon, \xi, b, \eta, \lambda, \mu, \alpha, \beta, \gamma)$ with $\lambda, \alpha \in \mathbb{R}_+^p$, $\mu, \beta \in \mathbb{R}_+^q$, and $\gamma \in \mathbb{R}_+$ is given by

$$L(w, \epsilon, \xi, b, \eta, \lambda, \mu, \alpha, \beta, \gamma) = \frac{1}{2}w^\top w - K_m\eta + K_s(\epsilon^\top \quad \xi^\top) \mathbf{1}_{p+q} \\ + (w^\top \quad (\epsilon^\top \quad \xi^\top) \quad b \quad \eta) C^\top \begin{pmatrix} \lambda \\ \mu \\ \alpha \\ \beta \\ \gamma \end{pmatrix}.$$

Since

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

the Lagrangian can be written as

$$L(w, \epsilon, \xi, b, \eta, \lambda, \mu, \alpha, \beta, \gamma) = \frac{1}{2}w^\top w - K_m\eta + K_s(\epsilon^\top \mathbf{1}_p + \xi^\top \mathbf{1}_q) + w^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} - \epsilon^\top(\lambda + \alpha) \\ - \xi^\top(\mu + \beta) + b(\mathbf{1}_p^\top \lambda - \mathbf{1}_q^\top \mu) + \eta(\mathbf{1}_p^\top \lambda + \mathbf{1}_q^\top \mu) - \gamma\eta \\ = \frac{1}{2}w^\top w + w^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} + (\mathbf{1}_p^\top \lambda + \mathbf{1}_q^\top \mu - K_m - \gamma)\eta \\ + \epsilon^\top(K_s \mathbf{1}_p - (\lambda + \alpha)) + \xi^\top(K_s \mathbf{1}_q - (\mu + \beta)) + b(\mathbf{1}_p^\top \lambda - \mathbf{1}_q^\top \mu).$$

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

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

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

$$w = -X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \tag{*w}$$