and $\alpha_i \lambda_i = 0$ for $i = 1, \ldots, p$ and $\beta_j \mu_j = 0$ for $j = 1, \ldots, q$.

But $(*_4)$ is equivalent to

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

which is precisely the result of adding $\alpha \geq 0$ and $\beta \geq 0$ as slack variables to the inequalities $(*_3)$ of Example 50.6, namely

$$-X^\top X \begin{pmatrix} \lambda \\ \mu \end{pmatrix} + b \begin{pmatrix} \mathbf{1}_p \\ -\mathbf{1}_q \end{pmatrix} + \mathbf{1}_{p+q} \leq 0_{p+q},$$

to make them equalities, where ρ plays the role of b.

When the constraints are *affine*, the dual function $G(\lambda, \nu)$ can be expressed in terms of the conjugate of the objective function J.

50.11 Conjugate Function and Legendre Dual Function

The notion of conjugate function goes back to Legendre and plays an important role in classical mechanics for converting a Lagrangian to a Hamiltonian; see Arnold [5] (Chapter 3, Sections 14 and 15).

Definition 50.11. Let $f\colon A \to \mathbb{R}$ be a function defined on some subset A of $\mathbb{R}^n$. The *conjugate* f^* of the function f is the partial function $f^*\colon \mathbb{R}^n \to \mathbb{R}$ defined by

$$f^*(y) = \sup_{x \in A}(\langle y, x \rangle - f(x)) = \sup_{x \in A}(y^\top x - f(x)), \quad y \in \mathbb{R}^n.$$

The conjugate of a function is also called the *Fenchel conjugate*, or *Legendre transform* when f is differentiable.

As the pointwise supremum of a family of affine functions in y, the conjugate function f^* is *convex*, even if the original function f is not convex.

By definition of f^* we have

$$f(x) + f^*(y) \geq \langle x, y \rangle = x^\top y,$$

whenever the left-hand side is defined. The above is known as *Fenchel's inequality* (or *Young's inequality* if f is differentiable).

If $f\colon A \to \mathbb{R}$ is convex (so A is convex) and if $\mathbf{epi}(f)$ is closed, then it can be shown that $f^{**} = f$. In particular, this is true if $A = \mathbb{R}^n$.