Given problem,

$$\begin{cases} & \min_x f(x) \\ \text{s.t.} & h_i(x) \leq 0 \\ & l_j(x) = 0 \end{cases} \tag{7.1}$$

The KKT conditions are

1. $0 \in \partial(f(x) + u^T h(x) + v^T l(x))$
2. $u_i h_i = 0, \forall i$
3. $h_i(x) \leq 0, l_i(x) = 0$
4. $v_i \geq 0, \forall i$

- Necessary (conditions) for optimality, under strong duality.
- Always sufficient for optimality.
- Duality gap
  for $x(u, v)$ feasible, $f(x) - f(x^*) \leq f(x) - g(u, v)$
- Strong duality
  for some feasible (u^*, v^*), any primal solution x^* minimizes $L(x, u^*, v^*)$

7.1 Conjugate Function

Given $f : \mathbb{R}^n \to \mathbb{R}$, $f^* : \mathbb{R}^n \to \mathbb{R}$

$$f^*(y) = \max_x y^T x - f(x) \tag{7.2}$$

Note: The conjugate function is always convex.

"Largest gap between the line $y^T x$ and $f(x)$".

If f is differentiable, then this is called "Legendre transform" of f.

Properties

- Fenchel's inequality $\forall x, y$

$$f(x) + f^*(y) \geq x^T y \tag{7.3}$$

- $\implies f^{**} \leq f$
- if f is closed, convex, then $f^{**} = f$

- if f is closed, convex, then $\forall x, y$

$$x \in \partial f^*(y) \Longleftrightarrow y \in \partial f(x) \tag{7.4}$$
$$\Longleftrightarrow f(x) + f^*(y) = x^T y \tag{7.5}$$

- If $f(u, v) = f_1(u) + f_2(v)$, then

$$f^*(w, z) = f_1^*(w) + f_2^*(z) \tag{7.6}$$

- If $f(x) = ag(x) + b$, then

$$f^*(y) = ag^*(\frac{y}{a}) - b \tag{7.7}$$

- If $A \in \mathbb{R}^{n\times n}$, non-singular, $f(x) = g(Ax + b)$, then

$$f^*(y) = g^*(A^{-T}y) - b^T A^{-T} y \tag{7.8}$$

Examples

1. $f(x) = x^T Qx$, $Q \succ 0$

$$\max_x y^T x - \frac{1}{2} x^T Q x$$

Take partial derivative towards x and set to 0,

$$y - Qx = 0 \Longrightarrow Q^{-1}y = x^*$$
$$f^*(y) = \frac{1}{2} y^T Q^{-1} y$$

2. If $f(x) = I_C(x) = \begin{cases} 0 & x \in C, \\ \infty & \text{elsewhere.} \end{cases}$

$$f^*(y) = \max_{x\in C} y^T x \qquad\qquad \text{support function}$$

3. Dual norms
Given a norm $||\cdot||$,

$$|y^T x| \leq \|x\| \|y\|_*$$

For Euclidean norm, the conjugate is still the Euclidean norm, and the inequality is called Cauchy-Schwarz inequality.
For ℓ_p-norm,

$$(||x||_p)_* = ||y||_q$$
$$\frac{1}{q} + \frac{1}{p} = 1$$

Trace norm

$$(\|X\|_{tr})_* = \|X\|_{op} = \sigma_1(x)$$
$$\|X\|_{**} = \|X\|$$

Conjugate $f(x) = \|x\|$

$$f^*(y) = I_{\{z:\|z\|_* \leq 1\}}(y)$$

4. Squared norms $f(x) = \frac{1}{2}\|x\|^2$,
$$f^*(y) = \frac{1}{2}\|y\|_*^2$$
5. Affine function $f(x) = a^T x + b$
$$f^*(x) = \max_x y^T x - a^T x - b$$
bounded if and only if $y = a$.
$$f^*(y) = \begin{cases} -b & y = a \\ \infty & \text{else} \end{cases}$$
6. $f(x) = -\log x \qquad \text{dom}(f) : x > 0$
$$f^*(y) = \max_x y^T x + \log x$$
unbounded if $y \geq 0$
$x^* = -\frac{1}{y}$
$$f^*(y) = \begin{cases} -\log(-y) - 1 & y < 0 \\ \infty & \text{else} \end{cases}$$
Why?
$$-f^*(u) = \min_x f(x) - u^T x$$
Try $\min_x f(x) + g(x)$
$$\begin{aligned} \iff \min_{x,z} & f(x) + g(z) \\ \text{s.t. } & x = z \end{aligned}$$
$$\begin{aligned} g(u) &= \min_{x,z} f(x) + g(z) + u^T(z - x) \\ &= \min_x\{f(x) - u^T x\} + \min_z\{g(z) - (-u)^T z\} \\ &= -\max_x\{u^T x - f(x)\} - \max_z\{(-u)^T z - g(z)\} \\ &= -f^*(u) - g^*(-u) \end{aligned}$$
7. Dual of $\min_x f(x) + \|x\|$ is
$$\max_u -f^*(u) - I_{\{z:\|z\|_* \leq 1\}}(u)$$
Ex. Lasso
$$\begin{aligned} &\min_\beta \frac{1}{2}\|y - X\beta\|_2^2 + \lambda\|\beta\|_1 \\ \iff &\min_{\beta,z} \frac{1}{2}\|y - z\|_2^2 + \lambda\|\beta\|_1 \\ &\text{s.t. } z = X\beta \end{aligned}$$
Dual of $\frac{1}{2}\|v\|_2^2$ is $\frac{1}{2}\|v\|_2^2$
$$g(u) = -\frac{1}{2}\|u\|_2^2 + y^T u - I_{\{v:\|v\| \leq 1\}}(\frac{X^T u}{\lambda})$$

References