CSE 592: Convex Optimization
HW 2

Bhushan Sonawane
SBU ID: 111511679

1 Duality

1.1 Use the general form for the dual of a problem with linear equalities and an arbitrary objective to write down the dual of (2) in terms of the conjugate of the f_0. If necessary, simplify the dual by eliminating unnecessary variables

Following is the lagrangian of the given function.

$$L(w, z, \nu) = \sum_{i=1}^{m} g(z_i) + \sum_{i=1}^{m} \nu(-z_i + y_i - w^T x_i) \tag{1}$$

which is equivalent to,

$$L(w, z, \nu) = \sum_{i=1}^{m} g(z_i) - \sum_{i=1}^{m} \nu_i z_i - \sum_{i=1}^{m} \nu_i w^T x_i + \sum_{i=1}^{m} \nu_i y_i$$

Dual of Lagrangian can be computed as follows:

$$G(\nu) = \inf_{w,z}(\sum_{i=1}^{m} g(z_i) - \sum_{i=1}^{m} \nu_i z_i - \sum_{i=1}^{m} \nu_i w^T x_i + \sum_{i=1}^{m} \nu_i y_i)$$

Which is equivalent to,

$$G(\nu) = \inf_{z}(\sum_{i=1}^{m} g(z_i) - \sum_{i=1}^{m} \nu_i z_i) - \inf_{w}(\sum_{i=1}^{m} \nu_i w^T x_i) + \sum_{i=1}^{m} \nu_i y_i$$

Converting infimum to supremum, we get

$$G(\nu) = -\sup_{z}(\sum_{i=1}^{m} -g(z_i) + \nu^T Z) + \sup_{w}(\sum_{i=1}^{m} -\nu_i w^T x_i) + \sum_{i=1}^{m} \nu_i y_i$$

$$G(\nu) = -\sum_{i=1}^{m} g^*(\nu) + \sup_{w}(\sum_{i=1}^{m} -\nu_i w^T x_i) + \sum_{i=1}^{m} \nu_i y_i$$

Which equals,

$$G(\nu) = -\sum_{i=1}^{m} g^*(\nu) + \sum_{i=1}^{m} \nu_i y_i + \sup_{w}(\sum_{i=1}^{m} -\nu_i w^T x_i) \tag{2}$$

But,

$$\sup_{w} \sum_{i=1}^{m} -\nu_i w^T x_i = \begin{cases} 0 & \nu_i == 0 \text{ or } x_i == 0 \\ \text{-}\infty & \text{Otherwise} \end{cases}$$

Hence, equation (2) reduces to,

$$G(\nu) = -\sum_{i=1}^{m} g^*(\nu) + \sum_{i=1}^{m} \nu_i y_i + \begin{cases} 0 & \nu_i == 0 \text{ or } x_i == 0 \\ \text{-}\infty & \text{Otherwise} \end{cases} \tag{3}$$

1.2 Write down the KKT conditions for a pair of primal and dual optimal solutions of (2). Explain how to use the KKT conditions to easily obtain a primal optimal solution if you are given a dual optimal solution.

From equation (1), we know that, Lagrangian is as follows:

$$L(w, z, \nu) = \sum_{i=1}^{m} g(z_i) + \sum_{i=1}^{m} \nu(-z_i + y_i - w^T x_i)$$

As we don't have in-equality contraint, KKT condition have fewer terms as follows:

$$KKT\ conditions : \begin{cases} -z_i^* + y_i - w^T x_i = 0 & i = 1, ..., m \\ \nabla g_0(z^*) - \sum_{i=1}^{m} \nu_i^* = 0 \\ \sum_{i=1}^{m} \nu_i^* x_i^* = 0 \end{cases}$$

How to use the KKT conditions to easily obtain a primal optimal solution if you are given a dual optimal solution [1]

When the primal problem is convex, the KKT conditions are also sufficient for the points to be primal and dual optimal. f_i are convex and h_i are affine, and $x^{'}$, $\lambda^{'}$ and $\nu^{'}$ are any points that satisfies the KKT conditions,

$$KKT\ conditions : \begin{cases} f_i(x^{'}) \leq 0 & i = 1....m \\ h_i(x^{'}) = 0 & i = 1....m \\ \lambda_i^{'} \geq 0 & i = 1....m \\ \lambda_i^{'} f_i(x^{'}) = 0 & i = 1....m \\ \nabla f_0(x^{'}) + \sum_{i=1}^{m} \lambda_i^{'} \nabla f_i(x^{'}) + \sum_{i=1}^{p} \nu_i^{'} \nabla h_i(x\prime) = 0 \end{cases}$$

then, $x^{'}$ and $(^{'}, \nu^{'})$ are primal and dual optimal with zero duality gap.

First two conditions state that $x^{'}$ is primal feasible. and $\lambda_i^{'} \geq 0$ implies that L(x, $^{'}$, $\nu^{'}$) is convex in x.

Last KKT condition implies that it's gradient vanishes at x = $x^{'}$. Hence, we can conclude that,

$$g(\lambda^{'}, \nu^{'}) = L(x^{'}, \lambda^{'}, \nu^{'})$$

$$= f_0(x^{'}) + \sum_{i=1}^{m} \lambda_i^{'} f_i(x^{'}) + \sum_{i=1}^{p} \nu_i^{'} h_i(x^{'})$$

but, $h_i(x^{'}) = 0$ and $\lambda_i^{'} f_i(x^{'}) = 0$

$$= f_0(x^{'})$$

Any point that satisfies KKT conditions are primal and dual optimal and have zero duality gap.

Therefore, we can easily obtain primal optimal solution if dual optimal solutionis provided with use of KKT conditions.

1.3 Derive the Fenchel conjugate of g(z)

1.3.1 $g(z) = z^2/2$

From Fenchel Conjugate definition,

$$g^*(y) = \sum_{i=1}^{m} g^*(y_i)$$

$$g^*(y_i) = \sup_{z} z_i y_i - z_i^2/2$$

taking derivation w.r.t. z and putting to zero

$$\frac{d}{dz} g^*(y_i)) = y_i - z_i = 0$$

hence,

$$z_i = y_i$$

putting value of y back into equation (4) we get,

$$g^*(y_i) = \sup_z y_i y_i - y_i^2/2$$

$$g^*(y_i) = \sup_z y_i^2/2$$

$$g^*(y_i) = y_i^2/2$$

1.3.2 g(z) = max($|z|$ 1, 0)

From Fenchel Conjugate definition,

$$g^*(y) = \sup_z \sum_{i=1}^m z_i y_i - \sum_{i=1}^m max(|z_i| - 1, 0)$$

We know that,

$$g^*(y) = \sum_{i=1}^m g^*(y_i)$$

Considering i^{th} term

$$g^*(y_i) = \sup_z z_i y_i - max(|z_i| - 1, 0) \tag{4}$$

Case 1: $|z_i| - 1 \leq 0$

$$g^*(y_i) = \sup_z z_i y_i - 0$$

Taking derivative w.r.t. z_i

$$\frac{d}{dz} g^*(y_i) = y_i - 0$$

Equating derivative to 0, we get,

$$y_i = 0$$

substituting value of y_i back into equation (5), we get

$$g^*(y_i) = 0$$

Hence,

$$g^*(y_i) = \begin{cases} 0 & y_i == 0 \\ max(y_i, -y_i) & \text{Otherwise} \end{cases} \tag{5}$$

Case 2: $|Z_i| - 1 > 0$ **and** $Z_i \geq \mathbf{0}$

$$g^*(y_i) = \sup_z z_i y_i - z_i \tag{6}$$

Taking derivative w.r.t. z_i

$$\frac{d}{dz} g^*(y_i) = y_i - 1$$

$$y_i = 1$$

substituting value of y_i back into equation (5), we get

$$g^*(y_i) = \sup_z z_i - (z_i - 1)$$

$$g^*(y_i) = \sup_z z_i - z_i + 1$$

$$g^*(y_i) = 1$$

Hence,

$$g^*(y_i) = \begin{cases} 1 & y_i = 1 \\ \infty & \text{Otherwise} \end{cases} \tag{7}$$

Case 3: $|Z_i| - 1 > 0$ **and** $Z_i <$ **0**

$$g^*(y_i) = \sup_z z_i y_i + z_i \tag{8}$$

Taking derivative w.r.t. z_i

$$\frac{d}{dz} g^*(y_i) = -y_i + 1$$

$$y_i = 1$$

substituting value of y_i back into equation (5), we get

$$g^*(y_i) = \sup_z z_i - (z_i - 1)$$

$$g^*(y_i) = \sup_z 1$$

$$g^*(y_i) = 1$$

Hence,

$$g^*(y_i) = \begin{cases} 1 & y_i == 1 \\ \infty & \text{Otherwise} \end{cases}$$

Hence,

$$g^*(y_i) = \sum_{i=1}^{m} g^*(y_i)$$

where every

$$g^*(y_i) = \begin{cases} 0 & |z_i| - 1 \leq 0 \text{ and } y_i == 0 \\ max(y_i, -y_i) & |z_i| - 1 \leq 0 \text{ and } y_i \neq 0 \\ 1 & |z_i| - 1 > 0 \text{ and } (z_i < 0 \text{ or } y_i == 1) \\ \infty & \text{Otherwise} \end{cases}$$

1.4 Write down the Fenchel conjugate of the objective $f_0(z, w)$ of (2) (Hint: express the objective as an independent sum of functions of each of the optimization variables)

1.4.1 $g(z) = z^2/2$

Function,

$$\min_{w \in R^d, z \in R^m} \sum_{i=1}^{m} g(z_i)$$

s.t.

$$z_i = y_i - w^T x_i \qquad i = 1, ..., m$$

can be re-written as follows:

$$f(w,z) = \sum_{i=1}^{m} g(z_i) + < 0, w >$$

Hence,

$$f^*(\alpha,\beta) = \sum_{i=1}^{m} g^*(\beta_i) + \begin{cases} 0 & \alpha == 0 \\ \infty & \text{Otherwise} \end{cases} \tag{9}$$

Hence,

1.4.2 $g(z) = z^2/2$

$$g^*(\beta) = \sum_{i=1}^{m} \beta_i^2/2$$

from equation (10)

$$f^*(\alpha,\beta) = \sum_{i=1}^{m} \beta_i^2/2 + \begin{cases} 0 & \alpha == 0 \\ \infty & \text{Otherwise} \end{cases}$$

1.4.3 g(z) = max($|z|$ 1, 0)

$$g^*(\beta) = \sum_{i=1}^{m} g^*(\beta_i)$$

Where for each β_i is -

$$g^*(\beta_i) = \begin{cases} 0 & |z_i| - 1 \leq 0 \text{ and } \beta_i == 0 \\ max(\beta_i, -\beta_i) & |z_i| - 1 \leq 0 \text{ and } \beta_i \neq 0 \\ 1 & |z_i| - 1 > 0 \text{ and } (z_i < 0 \text{ or } \beta_i == 1) \\ \infty & \text{Otherwise} \end{cases}$$

from equation (10)

$$\sum_{i=1}^{m} g^*(\beta) + \begin{cases} 0 & \alpha == 0 \\ \infty & \text{Otherwise} \end{cases}$$

1.5 Write down the dual

From equation (3), we know that,

$$G(\nu) = -\sum_{i=1}^{m} g^*(\nu) + \sum_{i=1}^{m} \nu_i y_i + \begin{cases} 0 & \nu_i == 0 \text{ or } x_i == 0 \\ \infty & \text{Otherwise} \end{cases}$$

1.5.1 a.

$$G(\nu) = -\sum_{i=1}^{m} \nu^2/2 + \sum_{i=1}^{m} \nu_i y_i + \begin{cases} 0 & \nu_i == 0 \text{ or } x_i == 0 \\ \infty & \text{Otherwise} \end{cases}$$

1.5.2 b.

$$g^*(\nu_i) = \begin{cases} 0 & |z_i| - 1 \leq 0 \text{ and } \nu_i == 0 \\ max(\nu_i, -\nu_i) & |z_i| - 1 \leq 0 \text{ and } \nu_i \neq 0 \\ 1 & |z_i| - 1 > 0 \text{ and } (z_i < 0 \text{ or } \nu_i == 1) \\ \infty & \text{Otherwise} \end{cases}$$

$$G(\nu) = -\sum_{i=1}^{m} g^*(\nu_i) + \sum_{i=1}^{m} \nu_i y_i + \begin{cases} 0 & \nu_i == 0 \text{ or } x_i == 0 \\ \infty & \text{Otherwise} \end{cases}$$

References

[1] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* Cambridge University Press, New York, NY, USA, 2004. ISBN 0521833787.