4301 - HW2

September 2020

Problem 1: Duals of Duals

Consider the following optimization problem.

$$\min_{x,y\in\mathbb{R}} x^2 + y^2$$

such that

$$2x + y \leq 1$$
$$x \geq 0$$
$$y \geq 0$$

1. Is Slater's condition satisfied for this optimization problem?

   Yes. The objective function and all of the constraints are convex. So the Slater Condition holds and the duality gap is zero.

2. Use the method of Lagrange multipliers to construct a dual of this optimization problem.

$$\begin{aligned} L(x, y, \lambda) &= x^2 + y^2 + \lambda_1(2x + y - 1) - \lambda_2 x - \lambda_3 y \\ &= x^2 + (2\lambda_1 - \lambda_2)x + y^2 + (\lambda_1 - \lambda_3)y - \lambda_1 \\ &\overset{(a)}{=} (x + \frac{2\lambda_1 - \lambda_2}{2})^2 + (y + \frac{\lambda_1 - \lambda_3}{2})^2 - (\frac{2\lambda_1 - \lambda_2}{2})^2 - (\frac{\lambda_1 - \lambda_3}{2})^2 - \lambda_1 \end{aligned}$$

   This little bit of massage in (a) makes everything easier in the following. The dual function would be

$$g(\lambda) = \inf_{x,y} L(x, y, \lambda)$$
$$\lambda_1 \geq 0$$
$$\lambda_2 \geq 0$$
$$\lambda_3 \geq 0$$

$$\frac{\partial L(x, y, \lambda)}{\partial x} = 2x + 2\lambda_1 - \lambda_2 = 0 \to x = -\frac{2\lambda_1 - \lambda_2}{2}$$
$$\frac{\partial L(x, y, \lambda)}{\partial y} = 2y + \lambda_1 - \lambda_3 = 0 \to y = -\frac{\lambda_1 - \lambda_3}{2}$$

$$g(\lambda) = -(\frac{2\lambda_1 - \lambda_2}{2})^2 - (\frac{\lambda_1 - \lambda_3}{2})^2 - \lambda_1$$
$$\text{s.t. } \lambda \succeq 0$$

3. Use the method of Lagrange multipliers to construct a dual of your dual from (2) of this optimization problem. Did you recover the primal problem?

Dual of dual is not always equivalent to the primal. However, in this particular case, based on Fenchel-Moreau Theorem, we expect to retrieve the primal problem after taking the dual of dual.

The dual function $g(\lambda)$ is concave. However, we only know how to construct the dual function of convex functions. There are a few easy approaches to solve this issue. The easiest approach is to multiply -1 into the concave function $g(y)$ and build a convex function out of it. We will take this -1 back later. So, with the convex function $-g(\lambda)$ build the Lagrangian function as

$$L^*(\lambda, \nu) = (\frac{2\lambda_1 - \lambda_2}{2})^2 + (\frac{\lambda_1 - \lambda_3}{2})^2 + \lambda_1 - \lambda_1\nu_1 - \lambda_2\nu_2 - \lambda_3\nu_3 \tag{1}$$

The dual of $-g(\lambda)$ would be the concave function

$$p(\nu) = \inf_{\lambda} L^*(\lambda, \nu).$$

$$\nu_1 \geq 0$$

$$\nu_2 \geq 0$$

$$\nu_3 \geq 0$$

$$\frac{\partial L(\lambda\nu)}{\partial\lambda_1} = \frac{2(2\lambda_1 - \lambda_2)}{2} + \frac{(\lambda_1 - \lambda_3)}{2} + 1 - \nu_1 = 0 \tag{2}$$

$$\frac{\partial L(\lambda, \nu)}{\partial\lambda_2} = -\frac{2\lambda_1 - \lambda_2}{2} - \nu_2 = 0 \rightarrow \nu_2 = -\frac{2\lambda_1 - \lambda_2}{2} \tag{3}$$

$$\frac{\partial L(\lambda\nu)}{\partial\lambda_3} = -\frac{\lambda_1 - \lambda_3}{2} - \nu_3 = 0 \rightarrow \nu_3 = -\frac{\lambda_1 - \lambda_3}{2} \tag{4}$$

Add ν_2 and ν_3 from (3) and (4) into (2). The result is $\nu_1 = 1 - 2\nu_2 - \nu_3$. Now substitute ν_1 in (1) to get

$$\begin{aligned} p(\nu) &= (\frac{2\lambda_1 - \lambda_2}{2})^2 + (\frac{\lambda_1 - \lambda_3}{2})^2 + \lambda_1(2\nu_2 + \nu_3) - \lambda_2\nu_2 - \lambda_3\nu_3 \\ &= (\frac{2\lambda_1 - \lambda_2}{2})^2 + (\frac{\lambda_1 - \lambda_3}{2})^2 + \nu_2(2\lambda_1 - \lambda_2) + \nu_3(\lambda_1 - \lambda_3) \end{aligned} \tag{5}$$

Don't forget that since $\nu_1 \geq 0$, it is necessary that $2\nu_2 + \nu_3 \leq 1$. One more time, add ν_2 and ν_3 from (3) and (4) but this time into (5) and get the concave function

$$\begin{aligned} p(\nu) &= (\nu_2)^2 + (\nu_3)^2 + \nu_2(-2\nu_2) + \nu_3(-2\nu_3) \\ &= -\nu_2^2 - \nu_3^2 \end{aligned}$$

It is time to take that -1 back. So the dual of dual is the following minimization problem

$$\min_{\nu_2, \nu_3 \in \mathbb{R}} \nu_2^2 + \nu_3^2$$

such that

$$2\nu_2 + \nu_3 \leq 1$$

$$\nu_2 \geq 0$$

$$\nu_3 \geq 0$$

Problem 2: Projections onto Convex Hulls

1. Colab Q2.1

2. Use the method of Lagrange multipliers to construct a dual of this optimization problem.

$$\min_{\lambda \in \mathbb{R}^M} \frac{1}{2} \|q - \sum_{m=1}^{M} \lambda_m x^{(m)}\|_2^2$$

such that

$$\sum_{m=1}^{M} \lambda_m = 1,$$
$$\lambda \succeq 0.$$

Define matrix $X \in \mathbb{R}^{n \times M}$ such that $x^{(m)}$ is the m^{th} column of X for $m \in \{1, \ldots, M\}$. In addition, $\lambda \in \mathbb{R}^{M \times 1}$,$q \in \mathbb{R}^{n \times 1}$ and $x^{(m)} \in \mathbb{R}^{n \times 1}$ for all m. The Lagrangian function is

$$\begin{aligned} L(\lambda, \nu, \mu) &= \frac{1}{2}\|q - X\lambda\|_2^2 - \lambda^T \nu + \mu(\lambda^T \mathbf{1} - 1) \qquad (6) \\ &= \frac{1}{2} q^T q + \frac{1}{2} \lambda^T X^T X \lambda - \lambda^T X^T q - \lambda^T \nu + \mu(\lambda^T \mathbf{1} - 1) \end{aligned}$$

with $\nu \in \mathbb{R}^{M \times 1}$, $\mu \in \mathbb{R}$ and $\mathbf{1} = [1]^M \in \mathbb{R}^{M \times 1}$. So the dual function would be

$$g(\nu, \mu) = \inf_{\lambda} L(\lambda, \nu, \mu)$$

such that

$$\nu \succeq 0.$$

$$\frac{\partial L(\lambda, \nu, \mu)}{\partial \lambda} = X^T X \lambda - X^T q - \nu + \mu \mathbf{1} = 0 \qquad (7)$$

$$\lambda^* = (X^T X)^{-1}(\nu - \mu \mathbf{1} + X^T q)$$

Since $X^T X$ is not guaranteed to be invertible, we cannot substitute λ^* in (6). On the other hand, this is a convex optimization, and we know that the problem has an answer. So, **this attempt fails**.

An alternative would be **pseudoinverse which will be discussed in the class in the near future.**

For now, lets change the main variable and define a new optimization problem.

$$\min_{p \in \mathbb{R}^n} \frac{1}{2} \|q - p\|_2^2$$

such that

$$\sum_{m=1}^{M} \lambda_m x^{(m)} - p = \mathbf{0},$$
$$\sum_{m=1}^{M} \lambda_m - 1 = 0,$$
$$-\lambda \preceq 0.$$

With $\nu \in \mathbb{R}^M$, $\mu_1 \in \mathbb{R}^n$ and $\mu_2 \in \mathbb{R}$, the Lagrangian function is

$$\begin{aligned} L(p, \nu, \mu) &= \frac{1}{2}\|q - p\|_2^2 - \lambda^T \nu + (X\lambda - p)^T \mu_1 + \mu_2(\lambda^T \mathbf{1} - 1) \\ &= \frac{1}{2}\|q - p\|_2^2 + \lambda^T(-\nu + X^T \mu_1 + \mu_2 \mathbf{1}) - p^T \mu_1 - \mu_2. \end{aligned}$$

The dual function is

$$g(\nu, \mu) = \inf_p L(p, \nu, \mu)$$

such that

$$\nu \succeq 0.$$

$$\begin{aligned} \frac{\partial L(p, \nu, \mu)}{\partial p} &= p - q - \mu_1 = 0 \\ \frac{\partial L(p, \nu, \mu)}{\partial \lambda} &= -\nu + X^T \mu_1 + \mu_2 \mathbf{1} = \mathbf{0} \end{aligned}$$

Adding these two to the Lagrangian results in

$$\begin{aligned} g(\nu, \mu) &= \frac{1}{2}\|q - p\|_2^2 + \lambda^T(-\nu + X^T \mu_1 + \mu_2 \mathbf{1}) - p^T \mu_1 - \mu_2 \\ &= \frac{1}{2}\|\mu_1\|_2^2 - (q + \mu_1)^T \mu_1 - \mu_2 \\ &= \frac{1}{2}\|\mu_1\|_2^2 - q^T \mu_1 - \mu_1^T \mu_1 - \mu_2 \\ &= -\frac{1}{2}\|\mu_1\|_2^2 - q^T \mu_1 - \mu_2 \end{aligned}$$

And since $\nu \succeq 0$, this problem has the constraint of

$$X^T \mu_1 + \mu_2 \mathbf{1} \succeq 0.$$

This dual problem is a constrained and needs to be solved with Projected Gradient Decent. Before getting into solving this dual problem, lets first ease the notation of the constraint by defining $A = [X^T, \mathbf{1}] \in \mathbb{R}^{M,n+1}$ and $\mu = [\mu_1; \mu_2] \in \mathbb{R}^{n+1}$. From now, for any arbitrary point $k \in \mathbb{R}^{n+1}$, if $Ak \succeq 0$, then its projection is itself. But if $Ak \not\succeq 0$, then k needs to be projected on the hyperplane with the vector A. For now we do not solve this projection analytically, because it requires A to be full rank. And A being full rank is equivalent to $(X^T X)$ being invertable which we avoided before. So, we have to find the projection through a numerical optimization problem.

Problem 3: Convex Envelopes

1. For a finite point set, is the convex envelope differentiable? Explain.

   No. For one counter example, if the function f is convex, then $f_{\text{env}} = f$. In that case, if f is not differntiable, then f_{env} is not diffrentiable either.

2. Recall that, for any convex function $f : \mathbb{R}^n \to \mathbb{R}$, and any $x, x' \in \mathbb{R}^n$, $f(x) \geq f(x') + w^T(x - x')$, where w is a subgradient of f at x'. Using this, we can formulate the problem of evaluating the convex envelope of our collection of points at a query point $x \in \mathbb{R}^n$ as a convex optimization problem:

   $$f_{\text{env}}(x) = \sup_{w \in \mathbb{R}^n, y \in \mathbb{R}} y$$

   such that

   $$y^{(m)} \geq y + w^T(x^{(m)} - x), \text{ for all } m \in \{1, \ldots, M\}.$$

(a) Explain why this optimization problem is unbounded for certain choices of $x \in \mathbb{R}^n$.

For any query point outside the convex hull of xs, this problem is unbounded. It is because, there will be a direction for w in which $w^T(x^{(m)} - x) \leq 0$ for all $m \in \{1, \dots, M\}$. And if we do not bound $\|w\|_2^2$, then $w^T(x^{(m)} - x)$ will go to $-\infty$ for all m, and hence y will increase unboundedly.

(b) Colab Q3.b

(c) Construct a dual of the optimization problem in (b) using the method of Lagrange multipliers.

We write the dual of the both original problem and the modified version. There should be no difference in the optimization process.

The original:

$$f_{\text{env}} = \inf_{w \in \mathbb{R}^n, y \in \mathbb{R}} -y$$

such that

$$y + w^T(x^{(m)} - q) - y^{(m)} \leq 0$$
$$\|w\|_2^2 - \gamma^2 \leq 0.$$

The Lagrangian function is

$$\begin{aligned} L(w, y, \lambda) &= -y + \sum_{m=1}^{M} \lambda_m (y + w^T(x^{(m)} - q) - y^{(m)}) + \lambda_{M+1}(\|w\|_2^2 - \gamma^2) \\ &= -y + \sum_{m=1}^{M} \lambda_m y + \sum_{m=1}^{M} \lambda_m (w^T(x^{(m)} - q) - y^{(m)}) + \lambda_{M+1}(\|w\|_2^2 - \gamma^2) \end{aligned}$$

The dual function

$$g(\lambda) = \inf_{w} L(w, y, \lambda)$$

such that

$$\lambda \succeq \mathbf{0}.$$

$$\begin{aligned} \frac{\partial L(w, y, \lambda)}{\partial w} &= \sum_{m=1}^{M} \lambda_m (x^{(m)} - q) + 2\lambda_{M+1} w = 0 \rightarrow w_0 = \frac{\sum_{m=1}^{M} \lambda_m (q - x^{(m)})}{2\lambda_{M+1}}, \\ \frac{\partial L(w, y, \lambda)}{\partial y} &= -1 + \sum_{m=1}^{M} \lambda_m = 0 \end{aligned}$$

So

$$g(\lambda) = \sum_{m=1}^{M} \lambda_m (w_0^T(x^{(m)} - q) - y^{(m)}) + \lambda_{M+1}(\|w_0\|_2^2 - \gamma^2)$$

such that

$$\lambda \succeq \mathbf{0}$$
$$\sum_{m=1}^{M} \lambda_m = 1$$

The modified version:

$$h(w) = \min_{m \in \{1, \dots, M\}} \left[y^{(m)} - w^T(x^{(m)} - q) \right]$$

$$f_{\text{env}} = \inf_{w \in \mathbb{R}^n} -h(w)$$

such that

$$\|w\|_2^2 - \gamma^2 \leq 0.$$

The Lagrangian is

$$L(w, \lambda) = -h(w) + \lambda(\|w\|_2^2 - \gamma^2)$$

So

$$g(\lambda) = \inf_{w \in \mathbb{R}^n} L(w, \lambda).$$

With

$$m^* = \text{argmin}_{m \in \{1,\ldots,M\}} \left[y^{(m)} - w^T(x^{(m)} - q)\right]$$

we have

$$\frac{\partial L(w, \lambda)}{\partial w} = x^{(m^*)} - q + 2\lambda w = 0 \to w_0 = \frac{q - x^{(m^*)}}{2\lambda}.$$

Finally

$$g(\lambda) = -h(w_0) + \lambda(\|w_0\|_2^2 - \gamma^2)$$

such that

$$\lambda \geq 0$$

(d) Colab Q3.d