APPM 5360, Spring 2023 - Written Homework 9

Eappen Nelluvelil; Collaborators: Jack, Logan

April 7, 2023

1. Introduce a slack variable z such that $Ax + z = b$, and Lagrange multipliers λ (for the $\|x\|_1 \leq \tau$ constraint) and ν (for the $Ax + z = b$ constraint), and write down the Lagrangian and dual function for LS_τ (really, for the modified version of LS_τ that has the slack variables). Alternatively, use Fenchel-Rockafellar duality (in which case, you do not need to introduce z). Either way, the dual function should be found explicitly – it should be something that you could give to CVX/CVXPY to solve.

   After introducing the slack variable z, we have the following equivalent problem:

$$\begin{aligned} \min_{x \in \mathbb{R}^n, z \in \mathbb{R}^m} \quad & \|z\|_2 \\ \text{s.t.} \quad & Ax - b = -z \\ & \|x\|_1 \leq \tau. \end{aligned}$$

   The Lagrangian of this problem is given by

$$\mathcal{L}(x, z, \lambda, \nu) = \|z\|_2 + \lambda \left(\|x\|_1 - \tau\right) + \nu^T \left(Ax + z - b\right).$$

   The Lagrangian dual is then given by

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x,z} \mathcal{L}(x, z, \lambda, \nu) \\ &= \inf_{x,z} \|z\|_2 + \lambda \left(\|x\|_1 - \tau\right) + \nu^T \left(Ax + z - b\right) \\ &= \inf_{x} \left(\lambda \|x\|_1 + \nu^T (Ax)\right) + \inf_{z} \left(\|z\|_2 + \nu^T z\right) - \lambda\tau - \nu^T b \\ &= -\sup_{x} -\left(\lambda\|x\|_1 + \langle A^T \nu, x\rangle\right) - \sup_{z} -\left(\|z\|_2 + \langle \nu, z\rangle\right) - \lambda\tau - \nu^T b, \end{aligned}$$

   where we take the infimum separately over x and z since they are independent primal variables. Furthermore, we convert the infimums to supremums and find the supremums separately.

   (a) $-\sup_x -\left(\lambda\|x\|_1 + \langle A^T\nu, x\rangle\right)$
   Note that for the function $f(x) = \lambda\|x\|_1$, the conjugate function is given by

$$\begin{aligned} f^*(y) &= \sup_{x} \left(\langle y, x\rangle - \lambda\|x\|_1\right) \\ &= \begin{cases} 0 & \text{if } \|y\|_\infty \leq \lambda, \\ \infty & \text{else.} \end{cases} \end{aligned}$$

   Thus,

$$-\sup_{x} -\left(\lambda\|x\|_1 + \langle A^T\nu, x\rangle\right) = -f^*\left(-A^T\nu\right),$$

   and the supremum of 0 is obtained when $\|A^T\nu\|_\infty \leq \infty$.

(b) $-\sup_z -\left(\|z\|_2 + \langle \nu, z\rangle\right)$

Note that for the function $f(z) = \|z\|_2$, the conjugate function is given by

$$\begin{aligned} f^*(y) &= \sup_z \left(\langle y, z\rangle - \|z\|_2\right) \\ &= \begin{cases} 0 & \text{if } \|y\|_2 \leq 1, \\ \infty & \text{else.} \end{cases} \end{aligned}$$

Thus,

$$-\sup_z -\left(\|z\|_2 + \langle \nu, z\rangle\right) = -f^*(-\nu),$$

and the supremum of 0 is obtained when $\|\nu\|_2 \leq 1$.

Thus, the dual problem is given by

$$\begin{aligned} \max_{\lambda,\nu} \quad & -\lambda\tau - \nu^T b \\ \text{s.t.} \quad & \|A^T\nu\|_\infty \leq \lambda, \\ & \|\nu\|_2 \leq 1. \end{aligned}$$

2. If $\tau > 0$, can we guarantee strong duality for (LS$_\tau$)?

Yes. Our primal problem is of the form

$$\begin{aligned} \min_{x\in\mathcal{D}} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad i = 1, 2, \ldots, m, \\ & Ax = b, \end{aligned}$$

with $f_0, \ldots, f_m$ convex. For this type of problem, to guarantee strong duality of the primal problem, the only condition that needs to be satisfied is Slater's condition: There exists an $x \in \operatorname{relint}(\mathcal{D})$ such that

$$\begin{aligned} f_i(x) < 0, & \quad i = 1, 2, \ldots, m, \\ Ax = b. & \end{aligned}$$

Slater's theorem states that strong duality holds if Slater's condition holds and the primal problem is convex.

Concretely, our primal problem is

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \|\Psi D^T x - b\|_2 \\ \text{s.t.} \quad & \|x\|_1 - \tau \leq 0. \end{aligned}$$

The objective function is convex, and the inequality constraint is also convex. Thus, (LS$_\tau$) is convex, and the point $x = 0$ is strictly feasible as $\tau > 0$. Thus, Slater's condition is satisfied, and we have strong duality.

3. Let x_τ be the solution to (LS$_\tau$), and let

$$\sigma(\tau) = \|Ax_\tau - b\|_2$$

be the corresponding value of the objective. Suppose λ_τ and ν_τ are dual optimal variables. What is σ' (that is, $\frac{\mathrm{d}\sigma}{\mathrm{d}\tau}$)?

Hint: see section 5.6 in Boyd and Vandenberghe.

We assume that $\tau > 0$ per problem 2 to ensure that we have strong duality. We assume also that we have the optimal dual variables λ_τ and ν_τ. Per section 5.6 in Boyd and Vandenberghe, we have that $p^* = p^*(0,0) = g(\lambda_\tau, \nu_\tau)$, where $g(\lambda_\tau, \nu_\tau) = -\lambda_\tau \tau - v_\tau^T b$. We then have that

$$\begin{aligned} \sigma(\tau) &= \|Ax_\tau - b\|_2 \\ &= -\lambda_\tau \tau - v_\tau^T b. \end{aligned}$$

Taking the derivative of σ with respect to τ, we have that $\frac{\mathrm{d}\sigma}{\mathrm{d}\tau} = -\lambda_\tau$.

4. (a) What are the KKT conditions for (LS$_\tau$)? Simplify them as much as possible.
The KKT conditions for (LS$_\tau$) at the stationary point $(x_\tau, z_\tau, \lambda_\tau, \nu_\tau)$ are given by

$$\begin{aligned} \|x_\tau\|_1 - \tau &\leq 0, \\ Ax_\tau + z_\tau - b &= 0, \\ \lambda_\tau &\geq 0, \\ \lambda_\tau (\|x_\tau\|_1 - \tau) &= 0, \\ \nabla_x \mathcal{L}(x_\tau, z_\tau, \lambda_\tau, \nu_\tau) &= 0, \\ \nabla_z \mathcal{L}(x_\tau, z_\tau, \lambda_\tau, \nu_\tau) &= 0. \end{aligned}$$

The last two conditions are the stationary conditions. The first stationary condition simplifies to

$$\begin{aligned} \nabla_x \mathcal{L}(x_\tau, z_\tau, \lambda_\tau, \nu_\tau) = 0 &\implies 0 \in \lambda_\tau \partial (\|x_\tau\|_1) + \partial \left(\langle A^T \nu_\tau, x_\tau \rangle\right) \\ &\implies 0 \in \lambda_\tau \partial (\|x_\tau\|_1) + A^T \nu_\tau. \end{aligned}$$

Here, the ∂ refers to the sub-differential operator since the one-norm is not differentiable, but is sub-differentiable.
The second stationary condition simplifies to

$$\begin{aligned} \nabla_z \mathcal{L}(x_\tau, z_\tau, \lambda_\tau, \nu_\tau) = 0 &\implies 0 = \frac{z_\tau}{\|z_\tau\|_2} + \nu_\tau \\ &\implies \nu_\tau = -\frac{z_\tau}{\|z_\tau\|_2}. \end{aligned}$$

After simplifying the above KKT conditions, we obtain the following:

$$\begin{aligned} \|x_\tau\|_1 &\leq \tau, \\ z_\tau &= b - Ax_\tau, \\ \lambda_\tau &\geq 0, \\ \lambda_\tau (\|x_\tau\|_1 - \tau) &= 0, \\ \lambda_\tau \partial (\|x_\tau\|_1) + A^T \nu_\tau &\ni 0, \\ \nu_\tau &= -\frac{z_\tau}{\|z_\tau\|_2}. \end{aligned}$$

(b) Supposing we are given the primal solutions x_τ such that $f_0(x_\tau) > 0$, then use the stationary KKT condition to find the optimal Lagrange multiplier ν. Then use this to find the optimal Lagrange multiplier λ.
From question 4a, we have solved for the optimal ν_τ in terms of z_τ. By strong duality being available due to Slater's condition being satisfied, we have that λ_τ is necessarily positive. We then use the first stationary condition to solve for λ_τ, i.e.,

$$0 \in \lambda_\tau (\|x_\tau\|_1) + A^T \nu_\tau \implies -\frac{A^T \nu_\tau}{\lambda_\tau} \in \partial (\|x_\tau\|_1).$$

The subdifferential of $\|x_\tau\|_1$ is the set $[-1, 1]$, which implies that

$$\left| \left(\frac{A^T \nu_\tau}{\lambda_\tau} \right)_i \right| \leq 1, \quad i = 1, 2, \ldots, \text{length} \left(A^T \nu_\tau \right).$$

Since we have strong duality, the above inequality must be met for at least one index i, which occurs if $\lambda_\tau = \|A^T \nu_\tau\|_\infty$.