Continuous Optimization Methods: Homework 10

Due March 30, 2017 at 11:59 PM

Professor Marina Epelman

Byron Tasseff

tasseff@umich.edu

Problem 1

9

Part (a)

Let us first rewrite the function to be minimized in (P_c), i.e., $f_c(x) = \frac{1}{2}(x-c)^T(x-c)$. The objective of ($\mathrm{P}_d$) may similarly be rewritten as $f_d(x) = \frac{1}{2}(x-d)^T(x-d)$. Substituting the relation for d, we find

$$f_d(x) = \frac{1}{2}\Bigg[x^Tx - x^Tc - \left(\frac{1-e^Tc}{n}\right)x^Te - c^Tx + c^Tc \\ + \left(\frac{1-e^Tc}{n}\right)c^Te - \left(\frac{1-e^Tc}{n}\right)e^Tx + \left(\frac{1-e^Tc}{n}\right)e^Tc + \left(\frac{1-e^Tc}{n}\right)^2 e^Te\Bigg],$$

which, since $e^Tx = 1$ for feasible x, may be rearranged and simplified to form

$$f_d(x) = \frac{1}{2}\left[x^Tx - x^Tc - c^Tx + c^Tc + 2\left(\frac{1-e^Tc}{n}\right)e^Tc - 2\left(\frac{1-e^Tc}{n}\right)e^Tx + \left(\frac{1-e^Tc}{n}\right)^2 e^Te\right].$$

Calling the (constant) terms on the right 2α, this may be rewritten as $f_d(x) = \frac{1}{2}(x-c)^T(x-c) + \alpha$. Of course, minimization over this function, which only differs by a constant, will result in the same solution when compared to (P_c). Thus, (P_c) and (P_d) indeed have the same optimal solutions.

Part (b)

Assuming $X = \{x \in \mathbb{R}^n : x \geq \mathbf{0}\}$, we construct the Lagrangian associated with (P_d) as $L_d(d, w) = \frac{1}{2}(x-d)^T(x-d) + w(e^Tx - 1)$. Thus, $L_d^*(w) = -w + \inf_{x\in X} \frac{1}{2}(x-d)^T(x-d) + we^Tx$. The function inside the infinum is convex quadratic; thus, we may find its derivative and obtain the roots of the resulting expression, which gives $\bar{x} = d - we$. Substituting this result into $L_d^*(w)$ results in the expression $L_d^*(w) = -w + \frac{1}{2}(-we)^T(-we) + we^T(d - we) = -w + \frac{1}{2}w^2n + we^Td - w^2n = -w + we^Td - \frac{1}{2}w^2n$. Since we were operating within the domain $x \geq \mathbf{0}$, this implies $x = d - we \geq \mathbf{0}$, or $we \leq d$, giving us the dual formulation

$$\begin{aligned} v_d^* = \ \underset{w}{\text{supremum}} \quad & -w + we^Td - \frac{1}{2}w^2n \\ \text{subject to} \quad & we \leq d. \end{aligned}$$

Since $e^Td = 1$, this may be rewritten as

$$\begin{aligned} v_d^* = \ \underset{w}{\text{supremum}} \quad & -\frac{1}{2}w^2n \\ \text{subject to} \quad & we \leq d. \end{aligned}$$

If $d \geq \mathbf{0}$, as the problem suggests, then clearly $v_d^* = 0$ at $w^* = 0$. Of course, since $x = d - we$ and $w^* = 0$, we see $x^* = d$ is indeed an optimal solution of (P_d) since strong duality holds between the primal and dual.

Part (c)

shouldn't assume this

If there exists some $d_j < 0$ (where all other $d_{i \neq j} \geq 0$), for the inequality constraint of the dual in (b) to be satisfied, it must be that $w < 0$, as well. As the dual's objective function is clearly concave, it follows that any decrease in w will result in a decrease of the objective function. Thus, the dual will be maximized when $w^* = d_j$, and by the relation $x^* = d - w^* e$, we see $x_j^* = d_j - d_j = 0$, i.e., $x_j^* = 0$.

Part (d)

The problem can be solved using the following algorithm: given some vector $c = c^n$, compute $d^n = c^n + \left(\frac{1 - e^T c^n}{n}\right) e$. Next, inspect the components of d^n: if $d^n \geq \mathbf{0}$, we may assert $x^* = d^n$. Otherwise, if there exists some $d_j^n < 0$, set the corresponding x_j^* equal to zero and form a new vector c^{n-1} containing all elements of d^n except d_j^n. Using this new vector c^{n-1}, solve the reduced-dimension problem $(\mathrm{P}_{c^{n-1}})$. Perform these steps iteratively, reducing the dimension by one each time, until reaching the required termination criteria (i.e., when $d^k \geq \mathbf{0}$). Clearly, the algorithm presented solves (P_c) via variable elimination (eliminating components of d sequentially), in at most n steps (as the worst case is when all entries of of d are initially negative).

Problem 2

Clearly, if C and D are equal, then $S_C(y) = S_D(y)$. We are thus only left with proving that if $S_C(y) = S_D(y)$, then $C = D$. To do so, we begin by proving $C \subseteq D$ via contradiction. Suppose $\bar{x} \in C$ and $\bar{x} \notin D$. Since $\bar{x} \notin D$, by Theorem B.3.1 of Freund and Vera, there must exist a nonzero vector p and scalar α such that $p^T\bar{x} > \alpha$ and $p^T x < \alpha$ for all $x \in D$. This implies $\sup\{p^T x : x \in D\} \leq \alpha < p^T\bar{x} \leq \sup\{p^T x : x \in C\}$. Of course, this contradicts our assumption that $S_C(p) = S_D(p)$ for any $p \in \mathbb{R}^n$, and so it must be that $C \subseteq D$.

To prove $D \subseteq C$, we proceed in the other direction. That is, suppose $\bar{x} \in D$ and $\bar{x} \notin C$. Since $\bar{x} \notin C$, by Theorem B.3.1 of Freund and Vera, there must exist a nonzero vector p and scalar α such that $p^T\bar{x} > \alpha$ and $p^T x < \alpha$ for all $x \in C$. This implies $\sup\{p^T x : x \in C\} \leq \alpha < p^T\bar{x} \leq \sup\{p^T x : x \in D\}$. This again contradicts our assumption that $S_C(p) = S_D(p)$, and so $D \subseteq C$. This completes our proof that $C = D$. We may thus assert $C = D$ if and only if their support functions are equal, i.e., for any $y \in \mathbb{R}^n$, $S_C(y) = S_D(y)$.

Problem 3

We begin by writing the Lagrangians corresponding to (P) and $(\bar{P})$ as $L(x,u) = f(x) + \sum_{i=1}^{m} u_i g_i(x)$ and $\bar{L}(x,u) = f(x) + \sum_{i=1}^{r} u_i g_i(x)$. Note that we may write $L(x,u)$ alternatively as $L(x,u) = \bar{L}(x,u) + \sum_{i=r+1}^{m} u_i g_i(x)$. Thus, $L^*(u) = \inf_{x\in X} f(x) + \sum_{i=1}^{m} u_i g_i(x) = \inf_{x\in X} \bar{L}(x,u) + \sum_{i=r+1}^{m} u_i g_i(x)$ and $\bar{L}^*(u) = \inf_{x\in \bar{X}} f(x) + \sum_{i=1}^{r} u_i g_i(x)$. We may now write the first Lagrangian dual, (D), as

$$\begin{aligned} v^* = \ & \underset{u}{\text{supremum}} \quad \inf_{x\in X} f(x) + \sum_{i=1}^{m} u_i g_i(x) \\ & \text{subject to} \quad u_i \geq 0, \ i = 1, 2, \ldots, m, \end{aligned}$$

and the second Lagrangian dual, $(\bar{D})$, as

$$\begin{aligned} \bar{v}^* = \ & \underset{u}{\text{supremum}} \quad \inf_{x\in \bar{X}} f(x) + \sum_{i=1}^{r} u_i g_i(x) \\ & \text{subject to} \quad u_i \geq 0, \ i = 1, 2, \ldots, r. \end{aligned}$$

First, it is important to observe that $L(x,u) = \bar{L}(x,u) + \sum_{i=r+1}^{m} u_i g_i(x) \leq \bar{L}(x,u)$ for any selection of x and u, as the term $\sum_{i=r+1}^{m} u_i g_i(x)$ is a sum of nonpositive components. Thus, we may assert $\inf_{x\in X} L(x,u) \leq \inf_{x\in X} \bar{L}(x,u)$. Furthermore, since $\bar{X} \subseteq X$, surely $\inf_{x\in X} \bar{L}(x,u) \leq \inf_{x\in \bar{X}} \bar{L}(x,u)$. Using this result, we ascertain $\inf_{x\in X} L(x,u) = L^*(u) \leq \inf_{x\in \bar{X}} \bar{L}(x,u) = \bar{L}^*(u)$, leading to the conclusion that $v^* \leq \bar{v}^*$ when v^* is finite. Otherwise, if $v^* = -\infty$, it is trivial to see $v^* \leq \bar{v}^*$. Finally, by weak duality, we are assured $\bar{v}^* \leq f^*$. Summarizing the preceding steps, we have thus shown that $v^* \leq \bar{v}^* \leq f^*$, completing our proof.

Problem 4

Part (a)

We denote the nonlinear optimization problem associated with Approach 1 as (P_1), i.e.,

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_1(x) \\ \text{subject to} \quad & f_j(x) \le b_j,\ j = 2, 3, \ldots, s \\ & g_i(x) \le 0,\ i = 1, 2, \ldots, m, \end{aligned}$$

where $b_j,\ j = 2, 3, \ldots, s$ are fixed parameters. Similarly, we may denote the nonlinear optimization problem associated with Approach 2 as (P_2), i.e.,

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \sum_{j=1}^{s} w_j f_j(x) \\ \text{subject to} \quad & g_i(x) \le 0,\ i = 1, 2, \ldots, m, \end{aligned}$$

where $w_j \ge 0,\ j = 1, \ldots, s$ are fixed parameters, with $w_1 > 0$.

Part (b)

Suppose x^* is an optimal solution to (P_2) for some selection of weights $w_j,\ j = 1, \ldots, s$, with $w_1 > 0$. Clearly, selections of $b_j = f_j(x^*)$ for $j = 2, \ldots, s$ would imply that x^* is a feasible solution of (P_1), as well. Next, note that we may reformulate (P_1) as ($\tilde{\mathrm{P}}_1$), i.e.,

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & w_1 f_1(x) \\ \text{subject to} \quad & f_j(x) \le b_j,\ j = 2, 3, \ldots, s \\ & g_i(x) \le 0,\ i = 1, 2, \ldots, m, \end{aligned}$$

with $w_1 > 0$, as this modification would not change the feasible region nor optimal solutions $\bar{x}$ when compared to those of (P_1), only the objective values. Next, suppose that x^*, an optimal solution to (P_2), is not an optimal solution to ($\tilde{\mathrm{P}}_1$) nor (P_1). Let us denote an optimal solution to these problems, instead, as $\bar{x}$. This would imply $w_1 f_1(\bar{x}) < w_1 f_1(x^*)$. Since $\bar{x}$ is a feasible solution to (P_1), $f_j(\bar{x}) \le b_j,\ j = 2, 3, \ldots, s$, or using our previous assumption, $f_j(\bar{x}) \le f_j(x^*),\ j = 2, 3, \ldots, s$. Thus, summing the constraints $w_1 f_1(\bar{x}) < w_1 f_1(x^*)$ and $f_j(\bar{x}) \le f_j(x^*),\ j = 2, 3, \ldots, s$ (each scaled by a nonnegative scalar w_j), we observe

$$w_1 f_1(\bar{x}) + \sum_{j=2}^{s} w_j f_j(\bar{x}) = \sum_{j=1}^{s} w_j f_j(\bar{x}) < w_1 f_1(x^*) + \sum_{j=2}^{s} w_j f_j(x^*) = \sum_{j=1}^{s} w_j f_j(x^*),$$

i.e., that $\bar{x}$ is a better solution to (P_2) than x^*. However, this is a contradiction to our initial assumption that x^* is a globally optimal solution to (P_2). As we have now proven by contradiction, we can indeed find target levels $b_j,\ j = 2, 3, \ldots, s$, such that the same solution of (P_2) is optimal for (P_1).

Part (c)

Assuming a constraint qualification is satisfied for both problems, we write the KKT conditions for (P_1) as

$$\nabla f_1(x) + \sum_{j=2}^{s} v_j \nabla f_j(x) + \sum_{i=1}^{m} u_i \nabla g_i(x) = 0,$$

with $v \geq \mathbf{0}$, $u \geq \mathbf{0}$, $v_j\left(f_j(x) - b_j\right) = 0,\ j = 2, 3, \ldots, s$, and $u_i g_i(x) = 0,\ i = 1, 2, \ldots, m$. For ($P_2$),

$$\sum_{j=1}^{s} w_j \nabla f_j(x) + \sum_{i=1}^{m} \tilde{u}_i \nabla g_i(x) = 0,$$

with $\tilde{u} \geq \mathbf{0}$ and $\tilde{u}_i g_i(x) = 0,\ i = 1, 2, \ldots, m$.

Part (d)

If all functions are convex, we see both problems may be labeled as convex. Combining Theorems 3.1.6 and 3.2.1 of Freund and Vera, then, if (P_1) has a globally optimal solution x^*, as the problem implies, it should have associated KKT multipliers satisfying the first set of conditions in (c) above. Suppose x^* is an optimal solution to (P_1) for some selection of target levels $b_j,\ j = 2, \ldots, s$, and x^* satisfies the corresponding KKT conditions with KKT multipliers v^* and u^*. Thus, we can easily find weights $w_j,\ j = 1, \ldots, s$ with $w_1 > 0$ for (P_2), namely $w_j = v_j^*,\ j = 2, \ldots, s$, and $w_1 = 1$. Using this selection of weights, the KKT conditions of (P_2) become equivalent to (P_1), with $\tilde{u} = u$ trivially satisfying the remainder. This shows us that, for this selection of weights, the same solution obtained by Approach 1 is indeed optimal in Approach 2.

Problem 5

Note that for the problem labeled (3), the leading terms of the objective function, $\frac{1}{2}R_{ij}x_{ij}^2$, are strictly convex since we are given $R_{ij} > 0,\ \forall (i,j) \in A$. Thus, the objective function is strictly convex, and the problem (3) is convex. Invoking Theorem 3.2.1 of Freund and Vera, since the problem is convex, satisfaction of KKT conditions for some feasible solution $\bar{x}$ and associated KKT multipliers is sufficient for asserting global optimality of $\bar{x}$. We begin by constructing the Lagrangian of the problem described, i.e.,

$$L(x, v) = \sum_{(i,j)\in A} \left(\frac{1}{2} R_{ij} x_{ij}^2 - t_{ij} x_{ij} \right) + \sum_{i \in N} v_i \left(\sum_{j:(j,i)\in A} x_{ji} - \sum_{j:(i,j)\in A} x_{ij} \right).$$

Using the Lagrangian, we can rewrite the KKT conditions associated with the problem as $\nabla_x L(\bar{x}, v) = \mathbf{0}$. Each entry of $\nabla_x L(\bar{x}, v) = \mathbf{0}$ must thus satisfy

$$\frac{\partial L(\bar{x}, v)}{\partial x_{ij}} = R_{ij}\bar{x}_{ij} - t_{ij} + v_j - v_i = 0,$$

giving rise to Ohm's law, $v_i - v_j = R_{ij}\bar{x}_{ij} - t_{ij},\ \forall (i,j) \in A$. The solution $\bar{x}$ to the stated problem must thus satisfy Ohm's law. Moreover, it must also satisfy Kirchhoff's current law since the associated constraints appear in the original problem. Finally, by Corollary 2.1.2 of Freund and Vera, since the objective function of (3) is strictly convex, we are guaranteed that $\bar{x}$ is the *unique* global optimal solution to the problem.