MATH 408 FINAL EXAM SOLUTIONS March 16, 2020

(1) (20 points) Compute and classify all critical points of the function $f(x_1, x_2) := (x_1+x_2)^2-8(x_1+x_2)$.

Solution:
(15 points) Compute critical points: $\nabla f(x) = (2(x_1+x_2)-8)\begin{pmatrix}1 & 1\end{pmatrix}$. Hence $\nabla f(x) = 0$ if and only if $x_1 + x_2 = 4$, so the set of critical points is given by

$$\mathcal{C} := \{(x_1, x_2)\ |x_1 + x_2 = 4\}.$$

(5 points) Classification of critical points: The key is to show that $\mathcal{C}$ is the set for **global minimizers for** f. There are many ways to do this. Here are 5.

(a) $f(x) = ((x_1 + x_2) - 4)^2 - 16$. Hence the set of global minima of f is $\mathcal{C}$.

(b) $g(z) = z^2 - 8z$ has $g'(z) = 2z - 8$ and $g''(z) = 2$ so g is strictly convex with unique global minimizer $z = 4$. Consequently, $\mathcal{C}$ is the set for global minimizers for f since $f(x) = g(x_1+x_2)$.

(c) As above, $g(z) = z^2 - 8z$ is a strictly convex function and $f(x) = g\left([1\ 1]\begin{pmatrix}x_1\\x_1\end{pmatrix}\right)$, that is, f is of the form $h(Ax + b)$ where h is convex. Therefore, Part (2) of Theorem 5.16 on page 65 of the course notes, f is convex so that the set of global minimizers coincides with its critical points $\mathcal{C}$.

(d) $\nabla^2 f(x) = 2\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}$ with principal minors 1 and 0 so that $\nabla^2 f(x)$ is positive semidefinite. At this point we can either use our knowledge of quadratic functions or convex functions to assert that the set of global minimizers coincides with its critical points $\mathcal{C}$.

(e) $\frac{1}{2}f(x) + 8 = \frac{1}{2}\left\|[1\ 1]\begin{pmatrix}x_1\\x_1\end{pmatrix} - 4\right\|_2^2$. Hence minimizing f is the same as solving a linear least-squares problem whose solution set is given by $\mathcal{C}$. Therefore, the set of global minimizers coincides with its critical points $\mathcal{C}$.

---

(2) (30 points) Show that the function $f : \mathbb{R}^n \to \mathbb{R}$ given by $f(x) = \exp(\frac{1}{2}\|x\|_2^2)$ is convex.

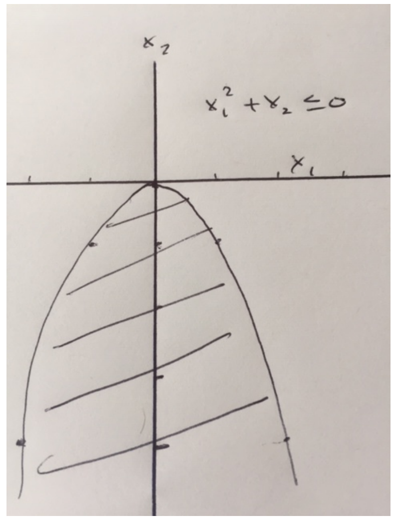
Solution: Again, there are many ways to do this. Here are two.

(a) Let $h(z) = \exp(z)$ and $g(x) = \frac{1}{2}\|x\|_2^2$. Since $h(z) = h'(z) = h''(z)$, h is a nondecreasing convex function of z. We also have $\nabla^2 g(x) = I$ so that g is a strictly convex function that maps into the domain of g. Hence, by Part (1) of Theorem 5.16 on page 65 of the course notes, f is convex.

(b) $\nabla f(x) = xf(x)$ and $\nabla^2 f(x) = f(x)[I + xx^T]$, and $u^T\nabla^2 f(x)u = f(x)[\|u\|_2^2 + (\langle x, u\rangle)^2] > 0$ for $u \neq 0$. Hence f is a strictly convex function.

(3) Consider the problem $\min\{x_1^2 - x_2 \,|x_1^2 + x_2 \le 0\}$.

(a) (5 points) Graph the constraint region $\{(x_1,\ x_2)\,|x_1^2 + x_2 \le 0\}$.



(b) (20 points) Compute a KKT pair for this problem.

KKT conditions:

(i) $x_1^2 + x_2 \le 0$

(ii) $0 \le \lambda$

(iii) $\lambda(x_1^2 + x_2) = 0$

(iv) $0 = \nabla_x L(x,\lambda) = \begin{pmatrix} 2x_1(1+\lambda) \\ -1+\lambda \end{pmatrix}$, where $L(x,\lambda) = x_1^2 - x_2 + \lambda(x_1^2 + x_2)$. (10 poiints)

Hence, by (iv), $\lambda = 1$ and $x_1 = 0$. By (iii), $x_2 = 0$. Therefore, since (i) and (ii) are also satisfied, $(x_1,\ x_2) = (0,\ 0)$ and $\lambda = 1$ is a KKT pair for this problem. (10 points)

(c) (5 points) Compute the tangent cone to the constraint region at the solution to this problem.

Set $c_1(x_1,\ x_2) = x_1^2 + x_2$ and $\Omega := \{x\,|c(x) \le 0\}$. Since $\nabla^2 c(x) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ is positive semidefinite, c is convex. Also, since $c(0,-1) = -1 < 0$, the Slater constraint qualification is satisfied. Hence, by the Theorem 5.18 page 66, the MFCQ is satisfied and so Ω is regular at $(0,0)$ giving

$$T((0,0)|\Omega) = \{d\,|\nabla c(0,0)^T d \le 0\} = \left\{ \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \middle| 0 \ge \begin{pmatrix} 0 \\ 1 \end{pmatrix}^T \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = d_2 \right\}.$$

Now return to the graph and see that the tangent cone at the origin is everything on or below the x_1 axis.

(d) (10 points) Show that the second-order sufficiency condition for this problem is satisfied at the KKT pair computed above.

Solution: First observe that $\nabla^2_{xx} L((0,0),1) = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$ and $\nabla f(0,0) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$, where $f(x_1,x_2) := x_1^2 - x_2$. Hence $d \in T((0,0)|\Omega) \setminus \{(0,0)\}$ satisfies $0 = \nabla f(0,0)^T d = -d_2$ if and only if $d_2 = 0$ (5 points). Consequently, for every such $d \ne 0$ (i.e. $d_1 \ne 0$), $d^T \nabla^2 L((0,0),1) d = 4d_1^2 > 0$, which shows that the second-order sufficiency condition in Theorem 4.9 on page 55 of the notes is satisfied (5 points).

(4) (30 points) Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable and let $\mathsf{e} \in \mathbb{R}^n$ denote the vector of all ones. Show that if $\bar{x}$ is a local solution to the problem $\min \{f(x) \mid 0 \le x \le \mathsf{e}\}$, then

$$\begin{aligned} \frac{\partial f}{\partial x_i}(\bar{x}) &\ge 0 \quad \text{if } \bar{x}_i = 0, \\ \frac{\partial f}{\partial x_i}(\bar{x}) &= 0 \quad \text{if } 0 < \bar{x}_i < 1, \\ \frac{\partial f}{\partial x_i}(\bar{x}) &\le 0 \quad \text{if } \bar{x}_i = 1. \end{aligned}$$

Hint: KKT conditions and $\{x \mid 0 \le x \le \mathsf{e}\} = \{x \mid 0 \le x_i \le 1,\ i = 1, 2, \ldots, n\}$.

Solution: Let $\Omega := \{x \in \mathbb{R}^n \mid 0 \le x \le \mathsf{e}\}$. Since Ω is convex and the Slater condition is satisfied (0 is in the interior of Ω), the MFCQ is satisfied at every point of Ω (Theorem 5.8 page 66). Therefore the local solution $\bar{x}$ is a KKT point by Theorem 4.6 page 53 and Theorem 4.22 page 50 (5 points). The KKT conditions tell us that that there exist $\bar{u}, \bar{v} \in \mathbb{R}^n$ such that
(i) $0 \le \bar{x}_i \le 1,\ i = 1, \ldots, n$
(ii) $0 \le \bar{u}_i,\ 0 \le \bar{v}_i,\ i = 1, \ldots, n$
(iii) $0 = \bar{u}_i \bar{x}_i,\ 0 = \bar{v}_i(\bar{x}_i - 1),\ i = 1, \ldots, n$
(iv) $0 = \nabla_x L(\bar{x}, \bar{u}, \bar{v})$, or equivalently, $\frac{\partial f(\bar{x})}{\partial x_i} = \bar{u}_i - \bar{v}_i, \quad i = 1, \ldots, n$, since $L(x, u, v) = f(x) - u^T x + v^T(x - \mathsf{e})$. (15 points)

Hence, we have the following:
(a) if $\bar{x}_i = 0$, then, by (iii), $\bar{v}_i = 0$, and so by (iv) and (ii), $\frac{\partial f(\bar{x})}{\partial x_i} = \bar{u}_i \ge 0$.
(b) if $0 < \bar{x}_i < 1$, then by (iii),$\bar{v}_i = 0 = \bar{u}_i$, and so by (iv), $\frac{\partial f(\bar{x})}{\partial x_i} = 0$.
(c) if $\bar{x}_i = 1$, then, by (iii), $\bar{u}_i = 0$, and so by (iv) and (ii), $\frac{\partial f(\bar{x})}{\partial x_i} = -\bar{v}_i \le 0$.
The condition (a)-(c) establish the result. (10 points)

---

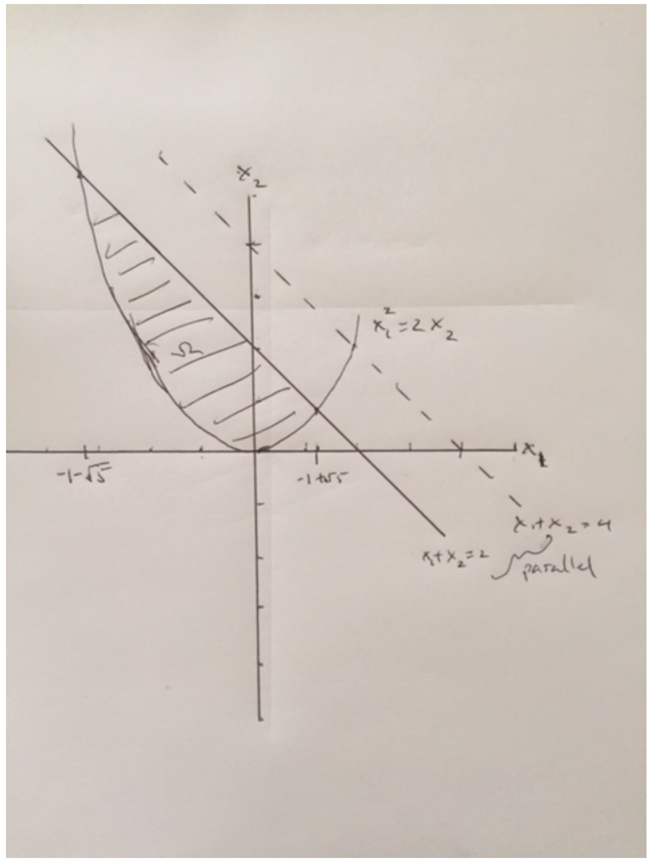
(5) Consider the problem

$$\begin{aligned} &\text{minimize } (x_1 + x_2)^2 - 8(x_1 + x_2) \\ &\text{subject to } x_1^2 \le 2x_2 \quad \text{and} \quad 2x_1 + 2x_2 \le 4, \end{aligned}$$

and note that this objective function occurs in problem 1.

(a) (10 points) Graph the constraint region and compare it to the graph of the set of critical points in problem 1. After thinking about the geometry of the setting, guess that one of the two dual variables takes the value zero.

Solution:



The graph of the objective is a parabolic valley whose bottom lies along the line $x_1 + x_2 = 4$. The function ascends uniformly away from this line. Hence (an educated guess) the solution set must be the line segment above the curve $x_1^2 = 2x_2$ and on the line $x_1 + x_2 = 2$ with $-1-\sqrt{5} \le x_1 \le -1+\sqrt{5}$ where the endpoints are the intersection points of the line $x_1+x_2 = 2$ and the curve $x_1^2 = 2x_2$. For the points with $-1-\sqrt{5} < x_1 < -1+\sqrt{5}$ the constraint $x_1^2 \le 2x_2$ is inactive, and so we guess that it's multiplier is zero at these points, i.e. $y_1 = 0$.

(b) (30 points) Describe the set of all KKT pairs for this problem.

Solution: The Lagrangian for this problem is

$$L((x_1, x_2), (y_1, y_2)) = (x_1 + x_2)^2 - 8(x_1 + x_2) + y_1(x_1^2 - 2x_2) + y_2(2x_1 + 2x_2 - 4),$$

and the KKT conditions are

(i) $x_1^2 \le 2x_2 \quad \text{and} \quad 2x_1 + 2x_2 \le 4$

(ii) $0 \le y_1, y_2$

(iii) $y_1(x_1^2 - 2x_2) = 0$ and $y_2(2x_1 + 2x_2 - 4) = 0$

(iv) $0 = \nabla_x L(x_1, x_2, y_1, y_2) = \begin{bmatrix} 2(x_1 + x_2) - 8 + 2y_1 x_1 + 2y_2 \\ 2(x_1 + x_2) - 8 - 2y_1 + 2y_2 \end{bmatrix}.$

From part (a), we guess that $y_1 = 0$ and $x_1 + x_2 = 2$. Then (iv) tells us that $\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{bmatrix} 4 - 8 + 2y_2 \\ 4 - 8 + 2y_2 \end{bmatrix}$, or $y_2 = 2$. Hence the set of KKT pairs is given by

$$\left\{ ((x_1, 2 - x_1), (0, 2) \;\middle|\; -1 - \sqrt{5} < x_1 < -1 + \sqrt{5} \right\}.$$

Since the problem is convex, Theorem 5.19 page 66 tells us that the set of optimal solutions is given by

$$\left\{ (x_1, 2 - x_1) \;\middle|\; -1 - \sqrt{5} < x_1 < -1 + \sqrt{5} \right\}.$$

(6) (40 points) Let $A \in \mathbb{R}^{m\times n},\ b \in \mathbb{R}^m,\ c \in \mathbb{R}^n \setminus \{0\}$, and $\gamma \in \mathbb{R}$. Show that the Lagrangian dual for the problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\|Ax-b\|_2^2 \\ \text{subject to} & c^Tx=\gamma \quad \text{and} \quad 0 \le x, \end{array}$$

is the problem

$$\begin{array}{ll} \text{maximize} & -\frac{1}{2}\|y+b\|_2^2 - \lambda\gamma + \frac{1}{2}\|b\|_2^2 \\ \text{subject to} & 0 \le A^Ty+\lambda c\,, \end{array}$$

where the maximization occurs over the dual variables $y \in \mathbb{R}^m$ and $\lambda \in \mathbb{R}$.

Step 1: Rewrite the problem by introducing a new variable w that simplifies the objective. Don't forget to write the definition of the new variable as one of the constraints.

Step 2: Write the Lagrangian.

Step 3: Write the condition $0 = \nabla_{(x,w)}L$.

Step 4: Use this condition to eliminate the primal variables from L and obtain the dual objective as a function of the dual variables only.

Step 5: Clean up the dual problem a bit so that it corresponds to the one give above.

Solution: Introduce the new variable $w = Ax - b$ and rewrite the problem as

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\|Ax-b\|_2^2 \\ \text{subject to} & Ax-b=w,\ c^Tx=\gamma \quad \text{and} \quad 0 \le x\ . \end{array}$$

(10 points)
The Lagrangian for this problem is

$$L((w,x),(y,u,v)) = \frac{1}{2}\|w\|_2^2 + y^T(Ax-b-w) + u(c^Tx-\gamma) - v^Tx\ ,$$

where $v \ge 0$. The dual objective is $\psi(y,u,v) := \min_{(w,x)} L((w,x),(y,u,v))$. By convexity, (w,x) attains the minimum in the definition of ψ if and only if

$$0 = \nabla_{(w,x)}L((w,x),(y,u,v)) = \begin{bmatrix} w-y \\ A^Ty+uc-v \end{bmatrix},$$

or equivalently, $w = y$ and $0 = A^Ty + uc - v$.
(10 points)
Using these identities, we eliminate the primal variables from L:

$$\begin{aligned} \psi(y,u,v) &= L((w,x),(y,u,v)) \\ &= \frac{1}{2}\|w\|_2^2 - y^Tw + (A^Ty+uc-v)^T - b^Ty - \gamma u \\ &= \frac{1}{2}\|y\|_2^2 - y^Ty - b^Ty - \gamma u \quad (\text{since } w=y \text{ and } 0 = A^Ty+uc-v) \\ &= -\frac{1}{2}\|y\|_2^2 - b^Ty - \gamma u \\ &= -\frac{1}{2}\|y+b\|_2^2 - \gamma u + \frac{1}{2}\|b\|_2\,. \end{aligned}$$

(10 points)
Hence we may write the dual problem as

$$\begin{array}{ll} \text{maximize} & -\frac{1}{2}\|y+b\|_2^2 - \gamma u + \frac{1}{2}\|b\|_2 \\ \text{subject to} & A^Ty+uc=v \quad \text{and} \quad 0 \le v\ , \end{array}$$

or alternatively,

$$\begin{aligned} &\text{maximize} \;\; -\frac{1}{2}\|y+b\|_2^2 - \gamma u + \frac{1}{2}\|b\|_2 \\ &\text{subject to } A^T y + uc \geq 0 \; . \end{aligned}$$

(10 points)

Note that the constant term $\frac{1}{2}\|b\|_2$ does not change the optimal solution to the dual. For this reason it is often dropped and the dual is written as

$$\begin{aligned} &\text{maximize} \;\; -\frac{1}{2}\|y+b\|_2^2 - \gamma u \\ &\text{subject to } 0 \leq A^T y + uc \; . \end{aligned}$$