Math 408 Homework Set 7 Solutions

(1) Find the global minimizers and maximizers, if they exist, for the following functions.

(a) $f(x) = x_1^2 - 4x_1 + 2x_2^2 + 7$

Solution: This function is fully separable, $f(x) = f_1(x_1) + f_2(x_2)$, where $f_1(x_1) = x_1^2 - 4x_1$ and $f_2(x_2) = 2x_2^2 + 7$. Hence we need only optimize f_1 and f_2 separately.

$$f_1'(x_1) = 2x_1 - 4,\ f"_1(x_1) = 2,\ f_2'(x_2) = 4x_2,\ f"_2(x_2) = 4\ .$$

Hence the unique critical point of f_1 is $x_1 = 2$ which is a global minimizer since f_1 is a parabola with positive curvature. Similarly, the unique global minimizer of f_2 is $x_2 = 0$. Therefore, the unique global minimizer of f is $(x_1, x_2) = (2, 1)$, and f has no other critical points.

(b) $f(x) = \mathsf{e}^{-\|x\|^2}$

Solution: Write f as

$$f(x) = \mathsf{e}^{-x_1^2}\mathsf{e}^{-x_2^2}\cdots\mathsf{e}^{-x_n^2} = \prod_{j=1}^{n} \mathsf{e}^{-x_j^2},$$

then it is easily seen that

$$\frac{\partial f}{\partial x_k}(x) = -2x_k f(x), \quad j = 1, 2, \ldots, n.$$

Hence

$$\nabla f(x) = -2f(x)x \quad \text{and} \quad \nabla^2 f(x) = 2f(x)\left[2xx^T - I\right]\ .$$

An expression of the form xy^T $(x, y \in \mathbb{R}^n)$, as appears above, is called the *outer product* of x and y. It is an $n \times n$ matrix whose ijth entry is $x_i y_j$. In particular, we have $x^T y = \text{trace}(xy^T)$.

Clearly, $x = 0$ is the unique critical point of f and this critical point is a local maximum of f since the Hessian of f at the origin is $-2I$ which is negative definite. Indeed, the origin is a global maximizer since $f(0) = 1$ which is the largest possible value of e^{ξ} for $\xi < 0$.

(c) $f(x) = x_1^2 - 2x_1x_2 + \frac{1}{3}x_2^3 - 8x_2$

Solution:

$$\nabla f(x) = \begin{bmatrix} 2(x_1 - x_2) \\ x_2^2 - 2x_1 - 8 \end{bmatrix} \quad \text{and} \quad \nabla^2 f(x) = \begin{bmatrix} 2 & -2 \\ -2 & 2x_2 \end{bmatrix}.$$

Compute the critical points by setting $\nabla f(x) = 0$. Setting $\partial f(x)/\partial x_1 = 0$ gives $x_1 = x_2$. Plug this into the equation $\partial f(x)/\partial x_2 = 0$ to get $0 = x_2^2 - 2x_2 - 8 = (x_2 - 4)(x_2 + 2)$. This gives 2 critical points

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}, \begin{pmatrix} -2 \\ -2 \end{pmatrix},$$

with

$$\nabla^2 f(4,4) = \begin{bmatrix} 2 & -2 \\ -2 & 8 \end{bmatrix} \quad \text{and} \quad \nabla^2 f(-2,-2) = \begin{bmatrix} 2 & -2 \\ -2 & -4 \end{bmatrix}.$$

It is easily shown that $\nabla^2 f(4,4)$ is positive definite and that $\nabla^2 f(-2,-2)$ has one positive and one negative eigenvalue. Hence $(4,4)$ is a local minimizer and $(-2,-2)$ is a saddle point. There are no global maximizers or minimizers since $f(0, x_2) = \frac{1}{3}x_2^3 - 8x_2$ which goes to $+\infty$ as $x_2 \uparrow +\infty$ and goes to $-\infty$ as $x_2 \downarrow -\infty$.

(d) $f(x) = (2x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - 1)^2$

Solution: Since f is a sum of squares, the smallest value f can take is zero. Hence any point $\bar{x}$ at which $f(\bar{x}) = 0$ is necessarily a global minimizer. To make $f(x) = 0$ each of the three squared terms in f must be zero. From the third term we get $x_3 = 1$. The second term gives $x_2 = x_3 = 1$, and the first term gives $2x_1 = x_2 = 1$ so $x_2 = 1/2$. Consequently, $(x_1, x_2, x_3) = (1/2, 1, 1)$ is the unique global minimizer of f.

Note that the function f is a convex function since we can write it in the form of a linear least squares objective:

$$f(x) = \|Ax - b\|_2^2, \quad \text{where } b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ and } A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Hence $(x_1, x_2, x_3) = (1/2, 1, 1)$ is also the unique critical point.

(e) $f(x) = x_1^4 + 16x_1x_2 + x_2^4$

Solution:

$$\nabla f(x) = \begin{pmatrix} 4x_1^3 + 16x_2 \\ 4x_2^3 + 16x_1 \end{pmatrix} \qquad \nabla^2 f(x) = \begin{bmatrix} 12x_1^2 & 16 \\ 16 & 12x_2^2 \end{bmatrix}$$

Hence x is a critical point if

$$\begin{aligned} 0 &= 4x_1^3 + 16x_2 \\ 0 &= 4x_2^3 + 16x_1 \ . \end{aligned}$$

Multiply the first equation by x_1 and the second by x_2 and subtract to get the equation

$$0 = 4[x_1^4 - x_2^4] = 4(x_1^2 + x_2^2)(x_1 + x_2)(x_1 - x_2).$$

This implies that either $x_1 = x_2$ or $x_1 = -x_2$. Plug this information into the first equation above to get

$$0 = 4x_1^3 \pm 16x_1 = 4x_1(x_1^2 \pm 4).$$

Therefore, the only possible critical points are

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ -2 \end{pmatrix}.$$

Since the gradient must be zero, x_1 and x_2 must have opposite sign. Plugging these vectors into the gradient, we see that only

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \end{pmatrix},$$

are critical points. Plugging these into the Hessian, we see that $x = 0$ is a saddle point (the eigenvalues of $\nabla^2 f(0)$ are ± 16), and the other two critical points are local minimizers. Furthermore, f is coercive since

$$\begin{aligned} f(x) &= x_1^4 + 16x_1x_2 + x_2^4 \\ &\geq x_1^4 - 16|x_1|\,|x_2| + x_2^4 \\ &\geq \begin{cases} x_1^4 - 16|x_1|^2 + x_2^4 & \text{if } |x_1| \geq |x_2|, \\ x_1^4 - 16|x_2|^2 + x_2^4 & \text{if } |x_2| \geq |x_1|, \end{cases} \end{aligned}$$

f is coercive (i.e. the right hand side of this inequality necessarily diverges to $+\infty$ as $\|x\|$ goes to infinity). Hence the critical points $(x_1, x_2) = (2, -2),\ (-2, 2)$ are global minimizers.

(f) $f(x) = (1 - x_1)^2 + \sum_{j=1}^{n-1} 10^j (x_j - x_{j+1}^2)^2$ (The Rosenbrock function)

Solution: Again, f is a sum of squares so any point at which the function takes the value zero is necessarily a global minimizer. The first term indicates that we should take $x_1 = 1$. The second term requires $x_2 = \pm x_1 = \pm 1$. The third term has $x_2 = x_3^2$, so $x_2 \geq 0$ implying that $x_2 = 1$. Moreover, $x_3 = \pm 1$. Continuing in this way we get $1 = x_1 = x_2 = \cdots = x_{n-1}$ and $x_n = \pm 1$, so that there are two global minimizers. There is only one other critical point for this function: $x_1 = 1/11, 0 = x_2 = x_2, \ldots, x_n$. One can show that it is a saddle point.

(2) Locate all of the KKT points for the following problems. Can you show that these points are local solutions? Global solutions?

(a)

$$\begin{array}{ll} \text{minimize} & e^{(x_1 - x_2)} \\ \text{subject to} & e^{x_1} + e^{x_2} \leq 20 \\ & 0 \leq x_1 \end{array}$$

Solution: This is convex problem so any local solution is a global solution. Obviously, we wish to make x_1 as small as possible and x_2 as big as possible. Hence, we must have $x_1 = 0$ which gives the solution $(x_1, x_2) = (0, \ln(20 - 1))$. By plugging this solution into the KKT conditions, we obtain the multipliers $(y_1, y_2) = (1, 2)/19$.

(b)

$$\begin{array}{ll} \text{minimize} & e^{(-x_1 + x_2)} \\ \text{subject to} & e^{x_1} + e^{x_2} \leq 20 \\ & 0 \leq x_1 \end{array}$$

Solution: Here we want to make x_1 as big as possible and x_2 as small as possible. By fixing x_1 at zero and sending x_2 to $-\infty$, the constraints are satisfied and the objective goes to zero. Hence, no solution exists and the optimal value is 0.

(c)

$$\begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 - 4x_1 - 4x_2 \\ \text{subject to} & x_1^2 \le x_2 \\ & x_1 + x_2 \le 2 \end{array}$$

Solution: This is a convex optimization problem and so any KKT point will give global optimality. Check that $(x_1, x_2) = (1, 1)$ and $(y_1, y_2) = (0, 2)$ is a KKT pair for this problem.

(d)

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\|x\|_2^2 \\ \text{subject to} & Ax = b \end{array}$$

where $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$ satisfies $\mathrm{Nul}\,(A^T) = \{0\}$.

Solution: This is a convex problem so $\bar{x}$ is a solution if and only if there is a $\bar{y}$ such that $(\bar{x}, \bar{y})$ is a KKT pair for this problem. The Lagrangian is $L(x, y) = \frac{1}{2}\|x\|_2^2 + y^T(b - Ax)$. The KKT conditions are $A\bar{x} = b$ and $\bar{x} = A^T\bar{y}$. Hence $b = A\bar{x} = AA^T\bar{y}$. Since $\mathrm{Nul}\,(A^T) = \{0\}$, $\mathrm{Nul}\,(AA^T) = \{0\}$ so that the matrix AA^T is invertible. Consequently, $\bar{y} = (AA^T)^{-1}b$ and $\bar{x} = A^T\bar{y} = A^T(AA^T)^{-1}b$.

(3) Show that the set

$$\Omega := \{x \in \mathbb{R}^2 | -x_1^3 \le x_2 \le x_1^3\}$$

is not regular at the origin. Graph the set Ω.

Solution: $T_\Omega(0, 0) = \{0\} \times \mathbb{R}_+$ while

$$\left\{d \,\middle|\nabla f_1(0,0)^T d \le 0 \text{ and } \nabla f_1(0,0)^T d \le 0\right\} = \{0\} \times \mathbb{R} \ ,$$

where $f_1(x_1, x_2) = -(x_1^3 + x_2)$ and $f_2(x_1, x_2) = x_2 - x_1^3$.

(4) Construct an example of a constraint region of the form

$$\{x \mid f_1(x) \le 0 \ i = 1, \dots, s, \ f_i(x) = 0 \ i = s+1, \dots, m\}$$

at which the MFCQ is satisfied, but the LI condition is not satisfied.

Solution: $\left\{(x_1, x_2) \,\middle| x_1^3 \le x_2 \text{ and } 0 \le x_2\right\}$.

(5) Suppose $\Omega = \{x\,;\, Ax \le b, \ Ex = h\}$ where $A \in \mathbb{R}^{m\times}$, $E \in \mathbb{R}^{k\times n}$, $b \in \mathbb{R}^m$, and $h \in \mathbb{R}^k$.

(a) Given $x \in \Omega$, show that

$$T_\Omega(x) = \{d \,:\, A_{i\cdot}d \le 0 \text{ for } i \in I(x), \ Ed = 0\},$$

where $A_{i\cdot}$ denotes the ith row of the matrix A and $I(x) = \{i \ A_{i\cdot}x = b_i\}$.

Solution: It was shown in class (see page 3 of the course notes *Optimality Conditions for Constrained Problems*) that

(♠) $$T_\Omega(x) \subset \{d \,:\, A_{i\cdot}d \le 0 \text{ for } i \in I(x), \ Ed = 0\},$$

so we need only show the reverse inclusion. Let $x \in \Omega$ and d be and element of the set of the right hand side of (♠). Note that by continuity there is a $\bar{t} > 0$ such that $x + td \in \Omega$ for all $0 \le t \le \bar{t}$. Hence $d \in T_\Omega(x)$.

(b) Given $x \in \Omega$, show that every $d \in T_\Omega(x)$ is a feasible direction for Ω at x.

Solution: This is what we showed in the answer to the previous question.

(c) Note that parts (a) and (b) above show that

$$T_\Omega(x) = \bigcup_{\lambda>0} \lambda(\Omega - x)$$

whenever Ω is a convex polyhedral set. Why?

Solution: Because every polyhedral convex set canb be given a representation

$$\Omega = \{x\,;\, Ax \le b, \ Ex = h\} \ .$$

(6) Show that each of the following functions is convex or strictly convex.

(a) $f(x, y) = 5x^2 + 2xy + y^2 - x + 2y + 3$

(b) $f(x, y) = \begin{cases} (x + 2y + 1)^8 - \log((xy)^2), & \text{if } 0 < x, \ 0 < y, \\ +\infty, & \text{otherwise.} \end{cases}$

(c) $f(x, y) = 4e^{3x-y} + 5e^{x^2+y^2}$

(d) $f(x,y) = \begin{cases} x + \frac{2}{x} + 2y + \frac{4}{y}, & \text{if } 0 < x,\ 0 < y, \\ +\infty, & \text{otherwise.} \end{cases}$

Solution

(a) f is quadratic with $Q = \begin{bmatrix} 10 & 2 \\ 2 & 2 \end{bmatrix}$, where Q is positive definite. Hence f is strictly convex.

(b) For $(x,y) \in \mathbb{R}^2_{++}$,

$$f(x,y) = (x + 2y + 1)^8 - 2\log x \ - \ 2\log y.$$

Since $\beta \mapsto \beta^8$, the composition $(x+2y+1)^8$ is convex Also, the mapping $\mu \mapsto -\log \mu$ for $\mu > 0$ is strictly convex. Hence the result follows since the non-negative linear combination of convex functions is convex. It is strictly convex due the log terms in x and y.

(c) The result follows since the non-negative linear combination of convex functions is convex. It is strictly convex since the mapping $(x,y) \mapsto e^{x^2+y^2}$ is strictly convex.

(d) For $(x,y) \in \mathbb{R}^2_{++}$, $\nabla^2 f(x,y) = \begin{bmatrix} 2x^{-2} & 0 \\ 0 & 4y^{-2} \end{bmatrix}$ which is positive definite, so f is strictly convex.

(7) Consider the global minimizers of the functions given in the previous problem if they exist.

(a) Compute the unique global minimizer.

Solution: $\bar{x} = -Q^{-1}b$ where $b = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

(b) Show that that global minimizer is obtained by solving the equation $4x(x + \sqrt{2x} + 1)^7 = 1$ for $x > 0$, then setting $y = \sqrt{x/2}$.

Solution: Setting the gradients equal to zero tells us that

$$\begin{aligned} \frac{1}{4} &= x(x + 2y + 1)^7 \\ \frac{1}{8} &= y^2(x + 2y + 1)^7, \end{aligned}$$

yielding $y = \sqrt{x/2}$ since x and y are positive. Plugging this expression for y into the second expression above gives $4x(x + \sqrt{2x} + 1)^7 = 1$ for $x > 0$.

(c) Show that the unique global solution is given by numerically solving the equation $5y\exp(10(y^2+1)) = 2$ for y then set $x = -3y$.

Solution: Setting the gradients equal to zero tells us that

$$\begin{aligned} 0 &= 12\exp(3x - y) + 10x\exp(x^2 + y^2) \\ 0 &= -4\exp(3x - y) + 10y\exp(x^2 + y^2). \end{aligned}$$

Multiplying the first expression by y and the second by $-x$ and adding yields $0 = (12y + 4x)\exp(3x - y)$ so that $x = -3y$. Plugging this expression for x into the first equation above gives $5y\exp(10(y^2+1)) = 2$.

(d) Compute the unique global minimizer.

Solution: Just set the gradient to zero to get $x = y = \sqrt{2}$.

(8) Let $Q \in \mathcal{S}^n_{++}$ and $c \in \mathbb{R}^n$. By making explicit use of Q^{-1}, compute the Lagrangian dual to the convex quadratic program

$$\mathcal{Q} \quad \begin{array}{ll} \text{minimize} & \frac{1}{2}x^TQx + c^Tx \\ \text{subject to} & Ax \le b,\ 0 \le x\ . \end{array}$$

Solution: The Lagrangian is $L(x,u,v) = \frac{1}{2}x^TQx + c^Tx + u^T(Ax - b) - v^Tx$ with $0 \le u,\ 0 \le v$. Stationarity of the Lagrangian gives

$$0 = \nabla_x L(x,u,v) = Qx + c + A^Tu - v,$$

so that $\bar{x} = -Q^{-1}(c + A^T u - v)$. Plugging this into L gives

$$\begin{aligned} L(\bar{x}, u, v) &= \frac{1}{2}(c + A^T u - v)^T Q^{-1}(c + A^T u - v) - (c + A^T u - v)^T Q^{-1}(c + A^T u - v) - b^T u \\ &= -\left[\frac{1}{2}(c + A^T u - v)^T Q^{-1}(c + A^T u - v) + b^T u\right]. \end{aligned}$$

Hence the dual is

$$\begin{array}{ll} \text{maximize} & -\left[\frac{1}{2}(c + A^T u - v)^T Q^{-1}(c + A^T u - v) + b^T u\right] \\ \text{subject to} & 0 \le u,\ 0 \le v. \end{array}$$

Moreover, if $(\bar{u}, \bar{v})$ solve the dual, then $\bar{x} = -Q^{-1}(c + A^T\bar{u} - \bar{v})$ solves the primal.

(9) Consider the functions

$$f(x) = \frac{1}{2}x^T Q x - c^T x$$

and

$$f_t(x) = \frac{1}{2}x^T Q x - c^T x + t\phi(x),$$

where $t > 0$, $Q \in \mathbb{S}^n_+$, $c \in \mathbb{R}^n$, and $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is given by

$$\phi(x) = \begin{cases} -\sum_{i=1}^n \ln x_i & \text{, if } x_i > 0,\ i = 1, 2, \ldots, n, \\ +\infty & \text{, otherwise.} \end{cases}$$

(a) Show that ϕ is a convex function.
(b) Show that both f and f_t are convex functions.
(c) Show that the solution to the problem $\min f_t(x)$ always exists and is unique.
(d) Let $\{t_i\}$ be a decreasing sequence of positive real scalars with $t_i \downarrow 0$, and let x^i be the solution to the problem $\min f_{t_i}(x)$. Show that if the sequence $\{x^i\}$ has a cluster point $\bar{x}$, then $\bar{x}$ must be a solution to the problem $\min\{f(x) : 0 \le x\}$.
Hint: Use the KKT conditions for the QP $\min\{f(x) : 0 \le x\}$.

Solution

(a) $\nabla^2\phi(x) = \text{diag}(x)^{-2}$ which is positive definite on $\mathbb{R}^n_{++}$.
(b) $\nabla^2 f(x) = Q$ which is positive definite. The result follows from (a) and the fact that the sum of two convex functions is convex.
(c) The sum of a symmetric positive definite and symmetric positive semi-definite matrix is symmetric and positive definite. So ∇f_t is everywhere positive definite which implies that f_t is strictly convex. Hence if a solution exists, it must be unique.
(d) For all $i = 1, 2, \ldots,$ we have $0 = \nabla f_{t_i}(x^i) = Qx^i - c - t_i\text{diag}(x^i)^{-1}\mathbf{1}$, where $\mathbf{1}$ is the vector of all ones. Without loss of generality we can assume that $x^i \to \bar{x}$. Set $\bar{v} := Q\bar{x} - c$ so that $t_i\text{diag}(x^i)^{-1}\mathbf{1} = Qx^i - c \to \bar{v} \ge 0$. Observe that $\bar{v}^T\bar{x} = \lim_i t_i\mathbf{1}^T\text{diag}(x^i)^{-1}x^i = \lim_i t_i n = 0$. Hence $(\bar{x}, \bar{v})$ is a KKT pair for the convex optimization problem $\min\{f(x)\ |0 \le x\}$ which implies that $\bar{x}$ must be a solution to this problem.