# Answer Key to Homework #6

Raymond Deneckere

Fall 2017

1. Let $f : \mathbb{R}^n \to \mathbb{R}$ is a $C^2$-function. Define $S = \{u \in \mathbb{R}^n : \| u \| = 1\}$, and let $x^* \in \mathbb{R}^n$. Suppose that for every $u \in S$, the function $g(\lambda) = f(x^* + \lambda u)$ satisfies $g'(0) = 0$ and $g''(0) < 0$, so that $g(\cdot)$ has a strict local maximum at $\lambda = 0$.

   (a) Interpret $g'(0)$.

   We may compute

   $$g'(\lambda) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x^* + \lambda u)u_i$$

   and so

   $$g'(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x^*)u_i = \nabla f(x^*)'u = D_u f(x)$$

   Thus $g'(0)$ is the directional derivative of $f$ in the direction of $u$ at the point $x^* \in \mathbb{R}^n$.

   (b) Prove that $x^*$ is a strict local maximum of $f$.

   Differentiating $g'(\lambda)$ with respect to $\lambda$, and evaluating the resulting expression at $\lambda = 0$ yields

   $$g'(\lambda) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x^*)u_i u_j = u'D^2 f(x^*)u < 0$$

   for all $u \in S$. Now let $z \in \mathbb{R}^n$ be such that $z \neq 0$, and let $u = \frac{1}{||z||}z$. Then we have

   $$z'D^2 f(x^*)z = ||z||^2 u'D^2 f(x^*)u < 0\,,$$

   since $||z|| \neq 0$. Thus $D^2 f(x^*)$ is negative definite. From part (a), upon setting $u = e_i$, we see that $\frac{\partial f}{\partial x_i}(x^*) = 0$ for all $i = 1, .., n$, and so we also have $Df(x^*) = 0$. This implies that $x^*$ is a strict local maximizer of $f(\cdot)$.

2. Sundaram, #4, parts (a), (b) and (c), p. 110.

   (a) We may compute

$$Df(x,y) = \left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\right) = \left(6x^2 + y^2 + 10x, 2xy + 2y\right)$$

A critical point of $f$ is a point for which $Df(x,y) = 0$. Thus critical points are the solutions to the equation system

$$\begin{aligned} 6x^2 + y^2 + 10x &= 0 \\ 2xy + 2y &= 0 \end{aligned}$$

From the second equation, we have $xy + y = 0$, so we have two possibilities. Either $y = 0$, in which case the first equation yields either $x = 0$ or $x = -\frac{5}{3}$, or $y \neq 0$, in which case $x = -1$. The first equation then becomes $y^2 = 4$, from which we deduce either $x = 2$ or $x = -2$. Thus the critical points are $(0,0), (-\frac{5}{3}, 0), (-1, 2)$ and $(-1, -2)$.

Let us now compute the second derivative of $f$:

$$D^2 f(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 12x + 10 & 2y \\ 2y & 2x + 2 \end{bmatrix}$$

Evaluated at each of the four critical points, this is respectively equal to

$$\begin{bmatrix} 10 & 0 \\ 0 & 2 \end{bmatrix} \quad \begin{bmatrix} -10 & 0 \\ 0 & -\frac{4}{3} \end{bmatrix} \quad \begin{bmatrix} -2 & 4 \\ 4 & 0 \end{bmatrix} \quad \begin{bmatrix} -2 & -4 \\ -4 & 0 \end{bmatrix}$$

Since the first matrix is positive definite, $f$ has a strict local minimum at $(0,0)$. The second matrix is negative definite, so $f$ has a strict local maximum at $(-\frac{5}{3}, 0)$. However, the last two matrices are neither positive semi-definite nor negative semi-definite. Hence the points $(-1,2)$ and $(-1,-2)$ are neither local maxima nor local minima.

(b) Since

$$Df(x,y) = \left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\right) = \left(2e^{2x}(x+y^2+2y)+e^{2x}, e^{2x}(2y+2)\right)$$

$(x,y)$ is a critical point when

$$\begin{aligned} 2e^{2x}(x+y^2+2y)+e^{2x} &= 0 \\ e^{2x}(2y+2) &= 0 \end{aligned}$$

Since $e^{2x}>0$ for all $x\in\mathbb{R}$ these equations admit only one critical point, namely $(\frac{1}{2},-1)$. On the other hand,

$$D^2f(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 4e^{2x}(x+y^2+2y)+4e^{2x} & 2e^{2x}(2y+2) \\ 2e^{2x}(2y+2) & 2e^{2x} \end{bmatrix}$$

we have

$$D^2f(\frac{1}{2},-1) = \begin{bmatrix} 2e & 0 \\ 0 & 2e \end{bmatrix}$$

which is positive definite. Hence $(\frac{1}{2},-1)$ is a strict local minimum.

(c) We have

$$Df(x,y) = \left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\right) = \left(y(a-x-y)-xy, x(a-x-y)-xy\right),$$

$(x,y)$ is a critical point when

$$\begin{aligned} ay-2xy-y^2 &= 0 \\ ax-2xy-x^2 &= 0 \end{aligned}$$

Subtracting the second equation from the first yields

$$x^2 - y^2 - a(x-y) = (x-y)(x+y-a) = 0$$

So either $x = y$ or $x + y = a$. Substituting these relations into the first equation, we obtain the four critical points $(0,0), (\frac{a}{3}, \frac{a}{3}), (a, 0)$ and $(0, a)$.

Now we may also compute

$$D^2 f(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} -2y & a - 2x - 2y \\ a - 2x - 2y & -2x \end{bmatrix}$$

Evaluated at each of the above four critical points, this respectively becomes

$$\begin{bmatrix} 0 & a \\ a & 0 \end{bmatrix} \quad \begin{bmatrix} -\frac{2a}{3} & -\frac{a}{3} \\ -\frac{a}{3} & -\frac{2a}{3} \end{bmatrix} \quad \begin{bmatrix} 0 & -a \\ -a & -2a \end{bmatrix} \quad \begin{bmatrix} -2a & -a \\ -a & 0 \end{bmatrix}$$

When $a = 0$, we have only one critical point $(0,0)$ and

$$D^2 f(x,y) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

which is both positive semi-definite and negative semi-definite. In general, $(0,0)$ could therefore be a local maximum or local minimum. But since $f(x,y) = -xy(x+y+1)$ when $a = 0$, it is neither a local maximum nor a local minimum. This is because we can make $f$ strictly positive or strictly negative in a neighborhood of the origin.

When $a \neq 0$, the second matrix is negative (positive) definite whenever $a > 0$ $(a < 0)$. Hence the critical point $(\frac{a}{3}, \frac{a}{3})$ is then a strict local maximum (minimum). However, it is not a global maximum (minimum) because $f(x,y)$ is unbounded when $x \to \pm\infty$. However, since the other three matrices are neither positive semi-definite nor negative semi-definite, neither of the three critical points $(0,0), (a,0)$ and $(0,a)$ are local maxima nor local minima.

3. Sundaram, #1, p.142.

Substituting $y^2 = 1 - x^2$ into $f(x, y) = x^2 - y^2$ we obtain $g(x) = 2x^2 - 1$, a single variable unconstrained problem. Solving the first order condition $g'(x) = 0$, we have $x = 0$. Since $g$ is a strictly convex function of $x$, $x = 0$ is a global minimum. Substituting $x = 0$ into the constraint yields $y = \pm 1$. Hence $(0, 1)$ and $(0, -1)$ are the global minima of $f$ on the constraint set.

On the other hand, by substituting $x^2 = 1 - y^2$ into $f(x, y) = x^2 - y^2$ we obtain $h(y) = 1 - 2y^2$, a single variable unconstrained problem. Solving the first order condition $h'(y) = 0$, we have $y = 0$. Since $h$ is a strictly concave function of $y$, $y = 0$ is a global maximum. Substituting $y = 0$ into the constraint yields $x = \pm 1$. Hence $(1, 0)$ and $(-1, 0)$ are the global maxima of $f$ on the constraint set.

We can also solve the problem by the method of Lagrange. First, let us make sure the method is applicable. The objective $f(x, y)$ is continuous, since it is a polynomial in $x$ and $y$. The constraint set $S = \{(x, y) : x^2 + y^2 = 1\}$ is compact, since it is closed and bounded. To see that it is closed, note that if $(x_n, y_n) \in S$ and $(x_n, y_n) \to (x, y)$, it follows from continuity of the function $x^2 + y^2$ that $(x, y) \in S$. Boundedness follows because $S \subset B(0, 2)$. Thus it follows from the Weierstrass Theorem that $f$ attains its global maximum and minimum on $S$. Furthermore, denoting the constraint by the function $k(x, y) = x^2 + y^2 - 1$, we have $Dk(x, y) = (2x, 2y)$, the constraint qualification holds whenever $(x, y) \neq (0, 0)$, which holds everywhere on the constraint set $S$. Furthermore, since both the objective function and the constraint function are polynomials, they are both $C^1$ functions. (This can also be shown by computing the derivatives, and showing that they are continuous functions). Thus the Theorem of Lagrange applies.

Let $L = x^2 - y^2 + \lambda(x^2 + y^2 - 1)$, where $\lambda$ is the Lagrange multiplier of the constraint $x^2 + y^2 = 1$. Taking the partial derivatives of $L$ w.r.t. $x, y$ and $\lambda$, we obtain:

$$\frac{\partial L}{\partial x} = 2x(1 + \lambda) = 0$$
$$\frac{\partial L}{\partial y} = -2y(1 - \lambda) = 0$$
$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 - 1 = 0$$

From the first equation, either $x = 0$ or $\lambda = -1$. First, let $x = 0$; then from the third equation we have $y = \pm 1$. Then from the second equation, we must have $\lambda = 1$. Hence we get two solutions, $(0, 1)$ and $(0, -1)$. Next, let $x \neq 0$, so that $\lambda = -1$. Then from the second equation we must have $y = 0$. Substituting $y = 0$ into the third equation then yields $x = \pm 1$. Hence we obtain the two solutions $(1, 0)$ and $(-1, 0)$. Thus the method of Lagrange yields the same set of solutions as the previous substitution method.

4. Sundaram, #2, p.142.

Substituting $y = 1 - x$ into the objective $f(x, y) = x^3 + y^3$ we obtain an unconstrained problem $h(x) = x^3 + (1 - x)^3 = 1 - 3x + 3x^2$. Since $h(x)$ is unbounded when $x$ approaches $\pm\infty$, this problem has no global maximizer.

Here is what the method of Lagrange would yield if we failed to check for existence first. Let $L = x^3 + y^3 + \lambda(x + y - 1)$ be the Lagrangean associated with the problem. Taking the partial derivatives w.r.t. $x, y$ and $\lambda$ yields

$$\begin{aligned} \frac{\partial L}{\partial x} &= 3x^2 + \lambda = 0 \\ \frac{\partial L}{\partial y} &= 3y^2 + \lambda = 0 \\ \frac{\partial L}{\partial y} &= x + y - 1 = 0 \end{aligned}$$

From the first two equations, if $\lambda = 0$, then we must have $x = y = 0$. This contradicts the third equation, so we have $\lambda \neq 0$. So we have $\lambda = -3x^2$ and $\lambda = -3y^2$. This implies $x = \pm y$. Substituting into the third equation we obtain a contradiction when $x = -y$. When $x = y$, we obtain $x = y = \frac{1}{2}$. This point $(\frac{1}{2}, \frac{1}{2})$ is a global minimizer. This can be seen from the unconstrained problem: $h(x)$ is a strictly convex function and attains a unique global minimum a $x = \frac{1}{2}$.

5. Sundaram, #3 part (a), p. 142.

Note that the objective is continuous ($f$ is a polynomial), and that the constraint set is the circumference of a circle and hence is compact (for details, see the answer to problem 3, above). It then follows from the Weierstrass Theorem that the optimization problem has a global maximizer and minimizer. Now both the objective $f$ and the constraint function $g$, being polynomials in $x$ and $y$ are $C^1$-functions. (This can also be verified by computing all partial derivatives of $f$ and $g$, and showing that they are continuous functions). Since $Dg(x,y) = (2x, 2y)$, the constraint qualification holds as long as $(x,y) \neq (0,0)$. However, $x = y = 0$ would require $a = 0$, in which case there is only one feasible point. Thus we may without loss of generality assume that $a \neq 0$, and the constraint qualification then holds for all points in the feasible set. Thus we can apply the Theorem of Lagrange.

Let $L = xy + \lambda(x^2 + y^2 - 2a^2)$ be the Lagrangean associated with our optimization problem. Taking the partial derivatives of $L$ w.r.t. $(x, y, \lambda)$, we obtain:

$$\frac{\partial L}{\partial x} = y + 2\lambda x = 0 \tag{1}$$

$$\frac{\partial L}{\partial y} = x + 2\lambda y = 0 \tag{2}$$

$$\frac{\partial L}{\partial y} = x^2 + y^2 - 2a^2 = 0 \tag{3}$$

First, we claim that any solution $(x, y, \lambda)$ to (1)-(3) must have $x \neq 0$. Indeed, if we had $x = 0$, then from (1) we must have $y = 0$. But this contradicts (3), so $x \neq 0$. Similar reasoning also establishes that $y \neq 0$. Hence from (1) and (2) we obtain $\lambda = -\frac{y}{2x}$ and $\lambda = -\frac{x}{2y}$. Thus we must have $x^2 = y^2$, implying $x = \pm y$ and $\lambda = \pm\frac{1}{2}$. Substituting this into (3) we obtain four critical points of the Lagrangean: $(a, a, -\frac{1}{2}), (a, -a, \frac{1}{2}), (-a, a, \frac{1}{2})$ and $(-a, -a, -\frac{1}{2})$. The points $(a, a)$ and $(-a, -a)$ are the global maximizers with $f(a, a) = f(-a, -a) = a^2$. The points $(-a, a)$ and $(a, -a)$ are the global minimizers, with $f(-a, a) = f(a, -a) = -a^2$.