Math 217 Fall 2013 Homework 7 Solutions

Due Thursday Nov. 7, 2013 5pm

- This homework consists of 6 problems of 5 points each. The total is 30.
- You need to fully justify your answer – prove that your function indeed has the specified property – for each problem.
- Please read this week's lecture notes before working on the problems.

Question 1. *Let $f(x,y) = x^3 + y^3 + x\,y^2$. Calculate its Taylor expansion to degree 2 with remainder (that is $n = 2$, the remainder involves 3rd order derivatives) at $(1,0)$.*

Solution. We have

$$\frac{\partial f}{\partial x} = 3\,x^2 + y^2, \qquad \frac{\partial f}{\partial y} = 3\,y^2 + 2\,x\,y, \tag{1}$$

$$\frac{\partial^2 f}{\partial x^2} = 6\,x, \quad \frac{\partial^2 f}{\partial x \partial y} = 2\,y, \quad \frac{\partial^2 f}{\partial y^2} = 6\,y + 2\,x \tag{2}$$

$$\frac{\partial^3 f}{\partial x^3} = 6, \qquad \frac{\partial^3 f}{\partial x^2 \partial y} = 0, \quad \frac{\partial^3 f}{\partial x \partial y^2} = 2, \quad \frac{\partial^3 f}{\partial y^3} = 6. \tag{3}$$

Therefore the Taylor expansion with remainder is

$$f(x,y) = 1 + 3\,(x-1) + \frac{1}{2}\,[6\,(x-1)^2 + 2\,y^2] + [(x-1)^3 + (x-1)\,y^2 + y^3]. \tag{4}$$

Question 2. *Let $f(x,y) = \frac{x^2}{y}$. Calculate its Taylor polynomial of degree 3 (that is P_3) at $(1,1)$.*

Solution. We have

$$\frac{\partial f}{\partial x} = 2\,x\,y^{-1}, \qquad \frac{\partial f}{\partial y} = -x^2\,y^{-2}; \tag{5}$$

$$\frac{\partial^2 f}{\partial x^2} = 2\,y^{-1}, \qquad \frac{\partial^2 f}{\partial x \partial y} = -2\,x\,y^{-2}, \qquad \frac{\partial^2 f}{\partial y^2} = 2\,x^2\,y^{-3}; \tag{6}$$

$$\frac{\partial^3 f}{\partial x^3} = 0, \qquad \frac{\partial^3 f}{\partial x^2 \partial y} = -2\,y^{-2}, \qquad \frac{\partial^3 f}{\partial x \partial y^2} = 4\,x\,y^{-3}, \qquad \frac{\partial^3 f}{\partial y^3} = -6\,x^2\,y^{-4}. \tag{7}$$

Thus P_3 at $(1,1)$ is

$$1 + [2\,(x-1) - (y-1)] + [(x-1)^2 - 2\,(x-1)\,(y-1) + (y-1)^2] + [-(x-1)^2\,(y-1) + 2\,(x-1)\,(y-1)^2 - (y-1)^3]. \tag{8}$$

Question 3. *Let $f(x,y,z) = \frac{\cos x \cos y}{\cos z}$. Calculate its Hessian matrix at $(0,0,0)$.*

Solution. We have

$$\frac{\partial f}{\partial x} = -\frac{\sin x \cos y}{\cos z}, \quad \frac{\partial f}{\partial y} = -\frac{\cos x \sin y}{\cos z}, \quad \frac{\partial f}{\partial z} = \frac{\cos x \cos y \sin z}{(\cos z)^2}. \tag{9}$$

Next

$$\frac{\partial^2 f}{\partial x^2} = -\frac{\cos x \cos y}{\cos z}, \quad \frac{\partial^2 f}{\partial y^2} = -\frac{\cos x \cos y}{\cos z}, \quad \frac{\partial^2 f}{\partial z^2} = \frac{\cos x \cos y \cos z}{(\cos z)^2} + 2\,\frac{\cos x \cos y\,(\sin z)^2}{(\cos z)^3}; \tag{10}$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\sin x \sin y}{\cos z}, \quad \frac{\partial^2 f}{\partial y \partial z} = -\frac{\cos x \sin y \sin z}{(\cos z)^2}, \quad \frac{\partial^2 f}{\partial z \partial x} = -\frac{\sin x \cos y \sin z}{(\cos z)^2}. \tag{11}$$

So its Hessian matrix at $(0,0,0)$ is

$$H_f(0,0,0) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{12}$$

Question 4. *Let $f\colon \mathbb{R}^N \mapsto \mathbb{R}$ belong to C^2, that is all of its second order partial derivatives exist and are continuous. Let $\boldsymbol{x}_0 \in \mathbb{R}^N$. Assume*

$$\forall \boldsymbol{v} \in \mathbb{R}^N, \boldsymbol{v} \neq \boldsymbol{0} \quad \boldsymbol{v}^T H(\boldsymbol{x}_0)\, \boldsymbol{v} > 0 \tag{13}$$

where $H(\boldsymbol{x}_0)$ is the Hessian matrix of f at $\boldsymbol{x}_0$. Prove that there is $r > 0$ such that for all $\boldsymbol{x} \in B(\boldsymbol{x}_0, r)$, there holds

$$\forall \boldsymbol{v} \in \mathbb{R}^N, \boldsymbol{v} \neq \boldsymbol{0} \quad \boldsymbol{v}^T H(\boldsymbol{x})\, \boldsymbol{v} > 0. \tag{14}$$

Solution. Since $f \in C^2$, each entry $h_{ij}(\boldsymbol{x})$ of the Hessian matrix $H(\boldsymbol{x})$ is continuous. Now define

$$g(\boldsymbol{x}, \boldsymbol{v}) \colon \mathbb{R}^{2N} \mapsto \mathbb{R}^{N \times N} \tag{15}$$

as

$$g(\boldsymbol{x}, \boldsymbol{v}) = \boldsymbol{v}^T H(\boldsymbol{x})\, \boldsymbol{v} = \sum_{i,j=1}^N v_i\, h_{ij}(\boldsymbol{x})\, v_j. \tag{16}$$

Since g is the sum of products $\sum_{i,j=1}^N v_i\, h_{ij}(\boldsymbol{x})\, v_j$ and $v_i, v_j, h_{ij}(\boldsymbol{x})$ are all continuous as functions of $(\boldsymbol{x}, \boldsymbol{v})$, g is a continuous function of $(\boldsymbol{x}, \boldsymbol{v})$.

Now consider the bounded closed set $A := \{(\boldsymbol{x}, \boldsymbol{v}) |\, \boldsymbol{x} = \boldsymbol{x}_0, \|\boldsymbol{v}\| = 1\}$. By assumption we have

$$g(\boldsymbol{x}, \boldsymbol{v}) > 0 \tag{17}$$

for all $(\boldsymbol{x}, \boldsymbol{v}) \in A$. By continuity for each point $(\boldsymbol{x}_0, \boldsymbol{v}_0)$ in A there is $r_{\boldsymbol{x}_0, \boldsymbol{v}_0} > 0$ such that

$$\forall (\boldsymbol{x}, \boldsymbol{v}) \in B((\boldsymbol{x}_0, \boldsymbol{v}_0), r_{x_0, v_0}), \qquad g(\boldsymbol{x}, \boldsymbol{v}) > 0 \tag{18}$$

By Heine-Borel A is compact, so there are finitely many such balls covering A. Now take r to be the smallest of their radius. We have, in particular,

$$\forall (\boldsymbol{x}, \boldsymbol{v}) \text{ with } \boldsymbol{x} \in B(\boldsymbol{x}_0, r), \|\boldsymbol{v}\| = 1, \qquad g(\boldsymbol{x}, \boldsymbol{v}) > 0. \tag{19}$$

Now for any $\boldsymbol{u} \in \mathbb{R}^N, \boldsymbol{u} \neq \boldsymbol{0}$, we have $\left\| \frac{\boldsymbol{u}}{\|\boldsymbol{u}\|} \right\| = 1$ and therefore

$$\boldsymbol{u}^T H(\boldsymbol{x})\, \boldsymbol{u} = \|\boldsymbol{u}\|^2 \left[\left(\frac{\boldsymbol{u}}{\|\boldsymbol{u}\|} \right)^T H(\boldsymbol{x}) \left(\frac{\boldsymbol{u}}{\|\boldsymbol{u}\|} \right) \right] > 0. \tag{20}$$

Remark. Alternatively we can prove by contradiction. Assume that for every $r > 0$, there is $\boldsymbol{x}_r \in B(\boldsymbol{x}_0, r)$ and nonzero $\boldsymbol{v}_r \in \mathbb{R}^N$ such that $\boldsymbol{v}_r^T H(\boldsymbol{x}_r)\, \boldsymbol{v}_r \leqslant 0$. Then setting $\boldsymbol{u}_r := \frac{\boldsymbol{v}_r}{\|\boldsymbol{v}_r\|}$ we have

$$\boldsymbol{u}_r^T H(\boldsymbol{x}_r)\, \boldsymbol{u}_r \leqslant 0. \tag{21}$$

But $\boldsymbol{u}_r \in S := \{\|\boldsymbol{x}\| = 1\}$ which is bounded and closed and is therefore compact. Thus there is a cluster point $\boldsymbol{u}$ such that for some $r_n \longrightarrow 0$, $\boldsymbol{u}_{r_n} \longrightarrow \boldsymbol{u}$. Clearly $\boldsymbol{x}_r \longrightarrow \boldsymbol{x}_0$. Thus

$$\boldsymbol{u}_{r_n}^T H(\boldsymbol{x}_{r_n})\,\boldsymbol{u}_{r_n} \longrightarrow \boldsymbol{u}^T H(\boldsymbol{x}_0)\,\boldsymbol{u} \Longrightarrow \boldsymbol{u}^T H(\boldsymbol{x}_0)\,\boldsymbol{u} \leqslant 0. \tag{22}$$

Finally as $\|\boldsymbol{u}_{r_n}\| = 1$ for all n, $\boldsymbol{u} \neq 0$. Contradiction.

Question 5. *Prove*

$$a, b \geqslant 0, n \geqslant 1 \Longrightarrow \left(\frac{a+b}{2}\right)^n \leqslant \frac{a^n + b^n}{2} \tag{23}$$

through solving $\min f(x, y) = x^n + y^n$ *subject to the constraint* $x + y = l > 0$.

Solution. Form the Lagrange function

$$L(x, y, \lambda) = (x^n + y^n) - \lambda\,(x + y - l). \tag{24}$$

Then the necessary conditions are

$$n\,x^{n-1} - \lambda = \frac{\partial L}{\partial x} = 0 \tag{25}$$

$$n\,y^{n-1} - \lambda = \frac{\partial L}{\partial y} = 0 \tag{26}$$

$$x + y - l = \frac{\partial L}{\partial \lambda} = 0. \tag{27}$$

Solving this we have $x^{n-1} = y^{n-1}, x + y = l > 0$. The only solution is $x = y = l/2$. Now the Hessian matrix at $\left(\frac{l}{2}, \frac{l}{2}\right)$ is $n\,(n-1)\,(l/2)^{n-2} I$ where I is the identity matrix. It is easy to check that this matrix is positive definite. Therefore $\left(\frac{l}{2}, \frac{l}{2}\right)$ is the only stationary point and a strict local minimizer.

Now we show that it is the global minimizer. Assume otherwise, that is there is $x_1 + y_1 = l$ such that $f(x_1, y_1) < f\left(\frac{l}{2}, \frac{l}{2}\right)$. Since $\left(\frac{l}{2}, \frac{l}{2}\right)$ is a strict local minimizer, the supreme between $\left(\frac{l}{2}, \frac{l}{2}\right)$ and (x_1, y_1) is reached and has to be different from both (x_1, y_1) and $\left(\frac{l}{2}, \frac{l}{2}\right)$. This point must be a local maximum and is then a stationary point, contradiction.

So we have proved

$$f(x, y) \geqslant f\left(\frac{l}{2}, \frac{l}{2}\right). \tag{28}$$

This means

$$x^n + y^n \geqslant \left(\frac{x+y}{2}\right)^n + \left(\frac{x+y}{2}\right)^n \tag{29}$$

and the conclusion follows.

Question 6. *Let* $f: \mathbb{R}^N \mapsto \mathbb{R}$ *belong to* C^2. *Let* $\boldsymbol{x}_0 \in \mathbb{R}^N$ *be a local* ***maximizer*** *for* f. *Prove*

a) $(\mathrm{grad}\, f)(\boldsymbol{x}_0) = \mathbf{0}$;

b) $\forall \boldsymbol{v} \in \mathbb{R}^N$, $\boldsymbol{v}^T H(\boldsymbol{x}_0)\,\boldsymbol{v} \leqslant 0$ *where* $H(\boldsymbol{x}_0)$ *is the Hessian matrix of* f *at* $\boldsymbol{x}_0$.

Solution.

a) Assume grad $f \neq 0$ at $\boldsymbol{x}_0$. Denote $\boldsymbol{v} := (\mathrm{grad}\ f)(\boldsymbol{x}_0)$. Since $f \in C^2$ it is in particular differentiable at x_0 and therefore

$$\frac{\partial f}{\partial \boldsymbol{v}} = (\mathrm{grad}\, f)(\boldsymbol{x}_0) \cdot \boldsymbol{v} = \|(\mathrm{grad}\, f)(\boldsymbol{x}_0)\|^2 > 0. \tag{30}$$

By definition

$$\lim_{h \longrightarrow 0} \frac{f(\boldsymbol{x}_0 - h\,\boldsymbol{v}) - f(\boldsymbol{x}_0)}{h} > 0. \tag{31}$$

and consequently there is $\delta > 0$ such that

$$\forall |h| < \delta, \qquad \frac{f(\boldsymbol{x}_0 - h\,\boldsymbol{v}) - f(\boldsymbol{x}_0)}{h} > 0 \tag{32}$$

which gives

$$\forall h \in (0, \delta), \qquad f(\boldsymbol{x}_0 - h\,\boldsymbol{v}) > f(\boldsymbol{x}_0). \tag{33}$$

Now for any $r > 0$, take $\boldsymbol{x} = \boldsymbol{x}_0 - h\,\boldsymbol{v}$ with $0 < h < \min\left(\frac{r}{\|\boldsymbol{v}\|}, \delta\right)$. Then

$$\|h\,\boldsymbol{v}\| < r \Longrightarrow \boldsymbol{x} \in B(\boldsymbol{x}_0, r) \tag{34}$$

but we have $f(\boldsymbol{x}) > f(\boldsymbol{x}_0)$. Contradiction.

b) Assume there is $\boldsymbol{v} \in \mathbb{R}^N$ such that $\boldsymbol{v}^T H(\boldsymbol{x}_0)\,\boldsymbol{v} > 0$. Then since $f \in C^2$, each $h_{ij}(\boldsymbol{x})$ of the Hessian matrix is continuous. Consequently the function

$$g(\boldsymbol{x}) := \boldsymbol{v}^T H(\boldsymbol{x})\,\boldsymbol{v} = \sum_{i,j=1}^{N} v_i\,h_{ij}(\boldsymbol{x})\,v_j \tag{35}$$

is continuous. Thus there is $\delta > 0$ such that

$$\forall \boldsymbol{x} \in B(\boldsymbol{x}_0, \delta), \qquad \boldsymbol{v}^T H(\boldsymbol{x})\,\boldsymbol{v} > 0. \tag{36}$$

Now for any $r > 0$ consider

$$\boldsymbol{x} := \boldsymbol{x}_0 + h\,\boldsymbol{v} \tag{37}$$

with

$$h = \frac{\min(r, \delta)}{2\,\|\boldsymbol{v}\|}. \tag{38}$$

Clearly $\boldsymbol{x} \in B(\boldsymbol{x}_0, r) \cap B(\boldsymbol{x}_0, \delta)$.

Taylor expansion gives

$$\begin{aligned} f(\boldsymbol{x}) &= f(\boldsymbol{x}_0) + (\mathrm{grad}\, f)(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2}\,(\boldsymbol{x} - \boldsymbol{x}_0)^T H(\boldsymbol{\xi})\,(\boldsymbol{x} - \boldsymbol{x}_0) \\ &= f(\boldsymbol{x}_0) + \frac{h^2}{2}\,\boldsymbol{v}^T H(\boldsymbol{\xi})\,\boldsymbol{v} \\ &> f(\boldsymbol{x}_0). \end{aligned} \tag{39}$$

Here the last inequality is because $\boldsymbol{x}, \boldsymbol{x}_0 \in B(\boldsymbol{x}_0, \delta) \Longrightarrow \boldsymbol{\xi} \in B(\boldsymbol{x}_0, \delta) \Longrightarrow \boldsymbol{v}^T H(\boldsymbol{\xi})\,\boldsymbol{v} > 0$. This contradicts $\boldsymbol{x}_0$ being a local maximizer.