Homework 3 for Math 1540

Zhen Yao

Problem 33. Prove that:

(a) There is a unique continuous function $f : [0,1] \to \mathbb{R}$ such that
$$f(x) = 1 + \int_0^x t^2 f(t)\, dt \quad \text{for all } x \in [0,1].$$

(b) The function from (a) is of class $f \in C^\infty(0,1)$.

Proof.

(a) Consider the mapping $T : C([0,1],\mathbb{R}) \to C([0,1],\mathbb{R})$ defined by
$$T(f)(x) = g(x) + \int_0^x t^2 f(t^2)\, dt.$$

Clearly, f is a solution to the problem if and only if $T(t) = f$. Since $C([0,1],\mathbb{R})$ is compact, then we need to prove that $T(f)$ is a contraction.

Given $f, h \in C([0,1],\mathbb{R})$, we have
$$\begin{aligned}
d_\infty(T(f), T(h)) &= \sup_{x\in[0,1]} \left| \int_0^x t^2 f(t^2)\, dt - \int_0^x t^2 h(t^2)\, dt \right| \\
&\le \sup_{x\in[0,1]} \int_0^x \left| f(t^2) - h(t^2) \right| t^2\, dt \\
&\le d_\infty(f,h) \int_0^1 t^2\, dt \\
&= \frac{1}{3} d_\infty(f,h)
\end{aligned}$$

which implies that T is a contraction. And the result follows.

(b) We have $f'(x) = x^2 f(x)$, then clearly $f \in C^\infty(0,1)$.

∎

Problem 34. Prove that if $f : \mathbb{R}^n \to \mathbb{R}$ satisfies $|f(x) - f(y)| \le M\|x-y\|^{3/2}$, then f is constant.

Proof. It suffice to show that $Df(x) = 0$ on $\mathbb{R}^n$. By definition, we have
$$\frac{\|f(x+h) - f(x) - 0\cdot h\|}{\|h\|} \le \frac{M\|h\|^{3/2}}{\|h\|} = M\|h\|^{\frac{1}{2}} \xrightarrow{h\to\infty} 0,$$
which implies $Df(x) = 0$. Hence, f is constant. ∎

Problem 35. Prove that if the partial derivatives $\partial f/\partial x_1$ and $\partial f/\partial x_2$ of a function $f : \mathbb{R}^2 \to \mathbb{R}$ exist at every point of $\mathbb{R}^2$, and the partial derivative $\partial f/\partial x_1$ is continuous on $\mathbb{R}^2$, then f is differentiable at every point of $\mathbb{R}^2$.

Proof. Let $y = (y_1, y_2), x = (x_1, x_2) \in \mathbb{R}^2$, then we have
$$\begin{aligned}
f(y_1,y_2) - f(x_1,x_2) &= f(y_1,y_2) - f(x_1,y_2) + f(x_1,y_2) - f(x_1,x_2) \\
&= \frac{\partial f}{\partial x_1}(\xi, y_2)(y_1 - x_1) + \frac{\partial f}{\partial x_2}(x_1,x_2)(y_2 - x_1) + \varphi(y_2 - x_2)|y_2 - x_2|.
\end{aligned}$$

By definition, we have

$$\begin{aligned}
&\frac{\left|f(y_1,y_2)-f(x_1,x_2)-\sum_{i=1}^{2}\frac{\partial f}{\partial x_i}(x_1,x_2)(y_i-x_i)\right|}{\|y-x\|}\\
\leq&\underbrace{\left|\frac{\partial f}{\partial x_1}(\xi,y_2)-\frac{\partial f}{\partial x_1}(x_1,y_2)\right|}_{\xrightarrow[y\to x]{}0}\cdot\underbrace{\frac{|y_1-x_1|}{\|y-x\|}}_{\leq 1}+\underbrace{|\varphi(y_2-x_2)|}_{\xrightarrow[y\to x]{}0}\cdot\underbrace{\frac{|y_2-x_2|}{\|y-x\|}}_{\leq 1}\to 0.
\end{aligned}$$

Hence, f is differentiable at every point of $\mathbb{R}^2$. $\square$

Problem 36. Prove that if $f\in C^1(\mathbb{R}^n)$ and ∇f is L-Lipschitz, $\|\nabla f(x)-\nabla f(y)\|\leq L\|x-y\|$, $x,y\in\mathbb{R}^n$, then

$$\frac{|f(y)-f(x)-\nabla f(x)(y-x)|}{\|y-x\|}\leq L\|y-x\|\quad\text{for all } x,y\in\mathbb{R}^n.$$

Proof. Since $f\in C^1(\mathbb{R}^n)$, then $f(y)-f(x)=\nabla f(\xi)(y-x)$, where $\xi\in\overline{xy}$. Then we have

$$\begin{aligned}
LHS&=\frac{|\nabla f(\xi)(y-x)-\nabla f(x)(y-x)|}{\|y-x\|}\\
&\leq\|\nabla f(x)-\nabla f(y)\|\\
&\leq L\|y-x\|.
\end{aligned}$$

$\square$

Problem 37. Let $f\in C^2(\mathbb{R}^2)$. Suppose that $\nabla f=0$ on a compact set $E\subset\mathbb{R}^2$. Prove that there is a constant $M>0$ such that $|f(x)-f(y)|\leq M|x-y|^2$ for all $x,y\in E$.
Hint: Cannot use mean-value theorem here, since $f(y)-f(x)=f'(\xi)(y-x)$ and maybe $\xi\notin E$.

Proof. Based on Problem 38, we have

$$f(y)-f(x)=\nabla f(x)(y-x)+\int_0^1(1-t)(y-x)^T f(x+t(y-x))(y-x)\,dt.$$

Then we have

$$|f(y)-f(x)|\leq|y-x|^2\int_0^1\left|D^2f(x+t(y-x))\right|\,dt.$$

Since E is compact, then $E\subset\overline{B}$, where $\overline{B}$ is a closed ball. And we can set $M=\sup_{\xi\in\overline{B}}|D^2f(\xi)|$, then we find the M that satisfies the condition. $\square$

Problem 38. Suppose that $f\in C^2(\mathbb{R}^n)$ has a local minimum at $x=0$ and $f(0)=0$. Prove that for any $x\in\mathbb{R}^n$

$$f(x)=\int_0^1(1-t)x^TD^2f(tx)x\,dt,$$

where $x\in\mathbb{R}^n$ is a column vector and x^T is the horizontal vector.

Proof. For any $g \in C^2(\mathbb{R}^n)$, we can derive second order Taylor expansion with integral remainder

$$\begin{aligned} g(1) - g(0) &= \int_0^1 g'(t)\,dt \\ &= -\int_0^1 g'(t)(1-t)'\,dt \\ &= -g'(t)(1-t)\Big|_0^1 + \int_0^1 g''(t)(1-t)\,dt \\ &= g'(0) + \int_0^1 g''(t)(1-t)\,dt, \end{aligned}$$

which implies

$$g(1) = g(0) + g'(0) + \int_0^1 g''(t)(1-t)\,dt.$$

Now we set $g(t) = f(x + t(y-x))$, and we have

$$\begin{aligned} g'(t) &= Df(x + t(y-x))(y-x) \\ g''(t) &= (y-x)^T D^2 f(x + t(y-x))(y-x) \end{aligned}$$

and then

$$f(y) - f(x) = g(1) - g(0) = g'(0) + \int_0^1 g''(t)(1-t)\,dt.$$

Hence, we have

$$f(y) = f(x) + Df(x)(y-x) + \int_0^1 (1-t)(y-x)^T D^2 f(x + t(y-x))(y-x)\,dt,$$

letting $y \to x$ and $x \to 0$, we have

$$f(x) = \int_0^1 (1-t)x^T D^2 f(tx)x\,dt.$$

$\square$

Problem 39. We know that $f \in C^2(\mathbb{R})$ is convex if and only if $f'' \geq 0$ on $\mathbb{R}$. Prove that $f \in C^2(\mathbb{R}^n)$ is convex if and only if $H_x(f) = D^2 f(x)$ is positive semidefinite for all $x \in \mathbb{R}^n$.

Proof. f is convex if and only if its restriction to any line in $\mathbb{R}^n$ is convex. This is equivalent to that $\forall x \in \mathbb{R}^n, \forall v \in \mathbb{R}^n, \|v\| = 1$, $g(t) = f(x + tv)$ is convex. Then we have $g''(t) \geq 0$. Since we have

$$\begin{aligned} g'(t) &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x + tv)v_i \\ g''(t) &= \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(x + tv)v_i v_j \\ &= v^T D^2 f(x + tv)v, \end{aligned}$$

then f is convex if and only if $\forall x \in \mathbb{R}^n, \forall v \in \mathbb{R}^n, \|v\| = 1$, $v^T D^2 f(x + tv)v \geq 0$. Hence, $D^2 f(x) \geq 0, \forall x \in \mathbb{R}^n$. $\square$

Problem 40. Prove that $x_0 \in \mathbb{R}^n$ is a critical point of $f \in C^2(\mathbb{R}^n)$, and $H_{x_0}(f)$ is positive definite, then there are $M > 0$ and $\varepsilon > 0$ such that

$$f(x) \geq f(x_0) + M\|x - x_0\|^2 \quad \text{whenever } \|x - x_0\| < \varepsilon.$$

Proof. With $Df(x_0) = 0$ and $H_{x_0}(f)$ is positive definite, we have f has local minimum at x_0. Then

$$\begin{aligned} f(x) &= f(x_0) + Df(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^T D^2 f(\xi)(x - x_0), \xi \in \overline{xy} \\ &= f(x_0) + \frac{1}{2}(x - x_0)^T D^2 f(\xi)(x - x_0), \end{aligned}$$

and we want to prove $\exists M, \exists \varepsilon, |\xi - x_0| < \varepsilon$ such that $v^T D^2 f(\xi) v \geq M\|v\|^2$.

Suppose not, then there exists $\xi_i \to x_0, \|v_i\| = 1$, such that $v_i^T D^2 f(\xi_i) v_i \leq 1/i$. Then we have $D^2 f(\xi_i) \to D^2 f(x_0)$ as $\xi_i \to x_0$ and hence $D^2 f(x_0) \leq 0$, which is a contradiction. $\square$

Problem 41. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a given function. Prove that if there are $M > 0$ and $\varepsilon > 0$ such that

$$f(x) \geq f(x_0) + M\|x - x_0\| \quad \text{whenever } \|x - x_0\| < \varepsilon,$$

then f is not differentiable at x_0.

Proof. Suppose f is differentiable at x_0. $f(x) = f(x_0) + Df(\xi)(x - x_0), \xi \in \overline{xx_0}$. Then let $\xi_i \to x_0$, and we have $Df(\xi_i) \to Df(x_0)$ which is finite. Then, as $x_i \to x_0$, we have

$$f(x_i) = f(x_0) + Df(\xi_i)(x_i - x_0) \to f(x_0),$$

which is a contraction. $\square$

Problem 42. Consider an open ball $B = B(a, r) \subset \mathbb{R}^n$. Prove that the function

$$\varphi(x) = \begin{cases} \exp\left(\frac{1}{|x-a|^2 - r^2}\right) & \text{if } x \in B\,, \\ 0 & \text{if } x \in \mathbb{R}^n \setminus B\,, \end{cases}$$

in infinitely differentiable on $\mathbb{R}^n$.

Proof. Consider the function

$$g(x) = \begin{cases} \exp\left(-\dfrac{1}{r^2 - |x|^2}\right), x \in B, \\ 0, x \in \mathbb{R}^n \setminus B, \end{cases}$$

And we can know that g is infinitely differentiable everywhere but when $|x| = r^2$. Indeed, for $|x| < r^2$, we have

$$Dg(x) = -2g(x)\frac{|x|^{1/2}}{(r^2 - |x|^2)^2}(x_1, \cdots, x_n)$$

with induction we can prove that $g^{(k)}(x) = P\left(|x|^{-1}\right) g(x)$, where $P\left(|x|^{-1}\right)$ is a polynomial of $|x|^{-1}$. Thus, g is infinitely differentiable when $|x| \neq r^2$.

Also, by definition, for any $a \in \mathbb{R}^n, |a| = r^2$, we have

$$Dg(a) = \lim_{h \to 0^-} \frac{g(a - h) - g(a)}{|h|}$$

Since $B = B(a, r) \subset \mathbb{R}^n$ is compact in $\mathbb{R}^n$, then with mean-value theorem, we have

$$\frac{g(a-h) - g(a)}{|h|} = Dg(\xi),$$

and as $h \to 0^-$, $\xi \to a$, and then

$$Dg(a) = \lim_{h \to 0^-} \frac{g(a-h) - g(a)}{|h|} = \lim_{h \to 0^-} Dg(\xi) = 0.$$

Also, by induction we can prove that $D^{(k)}g(a) = 0$, hence g is differentiable everywhere. Now let $f(x) = g(x - a)$, then the proof is complete. □

Problem 43. A C^2 function $f : \mathbb{R}^2 \to \mathbb{R}$ is called harmonic if

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0\,.$$

Prove that if a harmonic function has a local maximum at (x_0, y_0), then *all* second order partial derivatives of f vanish at (x_0, y_0).

Proof. Since f has a local maximum at (x_0, y_0), then its Hessian is negative semidefinite at this point, and when consider it on a vector $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, we have

$$v^T H_{x_0}(f) v = f_{xx}(x_0, y_0) + f_{yy}(x_0, y_0) + 2 f_{xy}(x_0, y_0) \le 0.$$

Also, with Sylvester's criterion, we have

$$f_{xx}(x_0, y_0) \le 0,$$

and

$$f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0) \ge 0.$$

Combining these implies that all second order partial derivatives of f vanish at (x_0, y_0). □

Problem 44. Suppose that a function $f : \mathbb{R}^2 \to \mathbb{R}$ of class C^2 satisfies the inequality $f_{xx} + f_{yy} \ge 0$ at every point of $\mathbb{R}^2$. Suppose also that all its critical points are non-degenerate, i.e. the matrix of second order derivatives at the critical point has non-zero determinant. Prove that f cannot have local maximum.

Proof. Suppose that f has local maximum at its critical point x_0, then we have $f_{xx} < 0$, and $f_{xx} f_{yy} - f_{xy}^2 > 0$, which implies $f_{xx} f_{yy} > 0$ and $f_{yy} < 0$. This is a contradiction with the fact that $f_{xx} + f_{yy} \ge 0$ at every point of $\mathbb{R}^2$. □

Problem 45. Let $f \in C^2(\Omega) \cap C^0(\overline{\Omega})$, where $\Omega \subset \mathbb{R}^n$ is open and bounded. Let $\Delta f = \sum_{i=1}^n \partial^2 f / \partial x_i^2$ be the Laplace operator.

(a) Show that if for some $\varepsilon > 0$ and $x_0 \in \Omega$ we have $\Delta f(x_0) \ge \varepsilon$, then f has no local maximum at x_0.

(b) Conclude that if $\Delta f(x) \ge \varepsilon$ for some $\varepsilon > 0$ and all $x \in \Omega$, then we have $\sup_\Omega f = \sup_{\partial\Omega} f$.

(c) Conclude that if $\Delta f(x) \ge 0$ for all $x \in \Omega$, then we have $\sup_\Omega f = \sup_{\partial\Omega} f$.

Hint for part (c): *Observe that $\Delta |x|^2 = 2n$. Use it to modify a function f in (c) so that you can apply part (b).*

Proof.

(a) Local maximum requires that $H_{x_0}(f)$ is negative definite, which means the trace $\partial^2 f/\partial x_i^2, i = 1, 2, \cdots, n$ of $H_{x_0}(f)$ are not positive. This is a contradiction with the fact $\Delta f(x_0) \geq \varepsilon$, then f has no local maximum at x_0.

(b) With (a), we can know that f has no local maximum in $\Omega \setminus \partial\Omega$. Thus, $\sup_\Omega f = \sup_{\partial\Omega} f$.

(c) Let $f_\varepsilon(x) = f(x) + \varepsilon|x|^2$, then $\Delta f_\varepsilon(x) = \Delta f(x) + 2\varepsilon n$. Then, we have

$$\sup_\Omega f(x) \leq \sup_\Omega f_\varepsilon(x) \leq \sup_{\partial\Omega} f_\varepsilon(x) \leq \sup_{\partial\Omega} f(x) + 2\varepsilon n \xrightarrow{\varepsilon\to 0} \sup_{\partial\Omega} f(x).$$

□

Problem 46. Let $A = (a_{ij})$ be an $n \times n$ matrix. The *cofactor* $A_{i,j}$ is the product of $(-1)^{i+j}$ with the $(n-1)\times(n-1)$ determinant of A obtained by deleting the ith row and jth column in A. Prove that

$$\frac{\partial}{\partial a_{ij}}(\det A) = A_{ij}.$$

Proof. The determinant of A is

$$\det A = \sum_{i=1}^n a_{ij} A_{i,j},$$

then it follows that

$$\frac{\partial}{\partial a_{ij}}(\det A) = A_{i,j}.$$

□

Problem 47. Let $A(t) = (x_{ij}(t)) : (a,b) \to \mathbb{R}^{n\times n}$ be a smooth matrix-valued curve. Prove that if $A(0) = I$, then

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}(\det A(t)) = \sum_{i=1}^n x'_{ii}(0) = \operatorname{tr} A'(0).$$

Proof. Denote the determinant function of a matrix by d as $\det A = d(r_1, r_2 \cdots, r_n)$, where r_n are rows of A. Then we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\det A(t) = d(r'_1, r_2, \cdots, r_n) + d(r_1, r'_2, \cdots, r_n) + \cdots + d(r_1, \cdots, r'_n).$$

When $t = 0$, the right hand side become the trace of $A'(0)$. Indeed, the first term $d(r'_1, r_2, \cdots, r_n)$ at $t = 0$ will be

$$\det\begin{pmatrix} x'_{11} & x'_{12} & \cdots & x'_{1n} \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = x'_{11}(0),$$

thus, we have $\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}(\det A(t)) = \sum_{i=1}^n x'_{ii}(0)$. □

Method II for Problem 47.

Proof. From last problem, we have for the determinant function $\det : \mathbb{R}^{n^2} \to \mathbb{R}$,

$$(\nabla \det)(A) = [A_{ij}]_{n\times n}.$$

Then we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} \det A(t) &= (\nabla \det(A)) \cdot A'(t) \\ &= \underbrace{[A_{ij}] \cdot \left[x'_{ij}\right]}_{\text{scalar product}} \\ &= \sum_{i,j=1}^{n} A_{ij}(t) x'_{ij}(t).\end{aligned}$$

Also, we have $[A_{ij}(0)] = \delta_{ij}$, hence

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\det A(t)) = \sum_{i,j=1}^{n} \delta_{ij} x'_{ij}(0) = \sum_{i=1}^{n} x'_{ii}(0).$$

□

Problem 48. Let $M_{n\times n}$ denote the vector space of real $n \times n$ matrices. Define a map $f : M_{n\times n} \to M_{n\times n}$ by $f(X) = X^2$. Find the derivative of f.

Proof. $Df(X) = X'X + XX'$. □

Problem 49. The class of invertible matrices $GL(n, \mathbb{R})$ forms an open subset in the space of all $n\times n$ matrices $M_{n\times n} = \mathbb{R}^{n^2}$. Let $F : GL(n, \mathbb{R}) \to GL(n, \mathbb{R})$, $F(A) = A^{-1}$. Prove that the function F is C^∞ smooth (as a mapping from $\mathbb{R}^{n^2}$ to $\mathbb{R}^{n^2}$) and that for any $A \in \mathbb{R}^{n\times n}$ we have

$$DF(A)B = -A^{-1} \circ B \circ A^{-1} \quad \text{for all } B \in \mathbb{R}^{n\times n}$$

Proof. Since $AA^{-1} = I$, then $A'A^{-1} + A\left(A^{-1}\right)' = 0$. Thus, $DF(A) = \left(A^{-1}\right)' = -A^{-1}A'A^{-1}$. We can prove by induction that $F(A)$ is infinitely differentiable.

For any $B \in \mathbb{R}^{n\times n}$, we can have

$$\begin{aligned}F(A + B) - F(A) &= (A + B)^{-1} - A^{-1} \\ &= \left(A\left(I + A^{-1}B\right)\right)^{-1} - A^{-1} \\ &= \left(\left(I + A^{-1}B\right) - I\right) A^{-1},\end{aligned}$$

Then, for B small enough, we have

$$\begin{aligned}F(A + B) - F(A) - \left(-A^{-1}BA^{-1}\right) &= \left(\left(I + A^{-1}B\right) - I + A^{-1}B\right) A^{-1} \\ &= \left(\sum_{k=0}^{\infty} (-1)^k \left(A^{-1}B\right)^k - I + A^{-1}B\right) A^{-1} \\ &= \sum_{k=2}^{\infty} (-1)^k \left(A^{-1}B\right)^k A^{-1} \\ &\to 0.\end{aligned}$$

Hence, $DF(A)B = -A^{-1}BA^{-1}$. □