THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
MATH 2058 Honours Mathematical Analysis I 2022-23
Test 2 solutions
2nd December 2022

- Please send an email to echlam@math.cuhk.edu.hk if you have any questions.

1. (25 points)

   Let $C[a,b]$ be the set of continuous real-valued functions on the closed and bounded interval $[a,b]$. Let $\mathcal{F} \subseteq C[a,b]$ be a non-empty subset that satisfies the following condition: for any $u, v \in \mathcal{F}$, $u \wedge v \in \mathcal{F}$, where $u \wedge v(x) := \min\{u(x), v(x)\}$ for any $x \in [a,b]$.

   (i) Let $g \in C[a,b]$, suppose that $g(x) = \inf\{h(x) : h \in \mathcal{F}, g \le h\}$ for any $x \in [a,b]$, prove that for any $\epsilon > 0$, there exists some $f \in \mathcal{F}$ such that $|f(x) - g(x)| < \epsilon$ for all $x \in [a,b]$.

   (ii) Does the result of Part (i) holds if g is only assumed to be bounded, instead of continuous?

   (iii) Does the result of Part (i) holds if the domain $[a,b]$ is replaced by the unbounded closed interval $[a,\infty)$?

   Solutions.

   (i) For each $y \in [a,b]$, by assumption, given any $\epsilon > 0$, there exists some $h_y \in \mathcal{F}$ with $g \le h$ so that $g(y) - \epsilon < g(y) \le h_y(y) < g(y) + \epsilon$. Here the subscript is to signify that that the dependence of h on y. Since both h_y and g are continuous function, the strict inequality signs are preserved in a small neighborhood of y. More specifically, if we consider the continuous function $\tilde{h}_y(x) = g(x) + \epsilon - h_y(x)$, since $\epsilon := \tilde{h}_y(y) > 0$, by continuity there is some $\delta'_y > 0$ so that for any $x \in (y - \delta'_y, y + \delta'_y) \cap [a,b]$, $|\tilde{h}_y(x) - \tilde{h}_y(y)| < \epsilon$. Hence we have $\tilde{h}_y(x) > \tilde{h}_y(y) - \epsilon = 0$. Likewise, $g(x) - \epsilon - h_y(x)$ is negative at $x = y$, so there is some small $\delta''_y > 0$ so that on $(y - \delta''_y, y + \delta''_y) \cap [a,b]$, we have $g(x) - \epsilon - h_y(x) < 0$.

   Now we take $\delta_y = \min\{\delta'_y, \delta''_y\}$, then over $I_y := (y - \delta_y, y + \delta_y)$, we have $g - \epsilon < h_y < g + \epsilon$. The collection of open intervals $\{I_y\}_{y \in [a,b]}$ forms an open cover of $[a,b]$, so by compactness there is a finite subcover $\{I_{y_i}\}_{i=1}^n$. Now we claim that $f := \min\{h_{y_i} : i = 1, ..., n\}$ is the desired function.

   Firstly, $f \in \mathcal{F}$ because each of $h_{y_i} \in \mathcal{F}$, and $\mathcal{F}$ is closed under taking minimum of functions. Next, suppose that $x \in [a,b]$, then $x \in I_{y_i}$ for some i, therefore $f(x) \le h_{y_i}(x) < g(x) + \epsilon$. And also the minimum $f(x) = h_{y_j}(x)$ for some j, so we have $g(x) - \epsilon < g(x) \le h_{y_j}(x) = f(x)$. Since x is arbitrary, we have shown that $g(x) - \epsilon < f(x) < g(x) + \epsilon$, i.e. $|f(x) - g(x)| < \epsilon$.

   (ii) No, a counter-example is given by $\mathcal{F} = \{h_n(x) := x^n | n \in \mathbb{N}\} \subset C[0,1]$. The function $g(x) = 0$ for $x \in [0,1)$ and $g(1) = 1$, is a bounded discontinuous function that can be realized as the infimum of $\mathcal{F}$. To see this, note that $\lim_{n\to\infty} x^n = 0$ for $1 > x \ge 0$ and equals 1 if $x = 1$.

Then for $\epsilon = \frac{1}{2}$, for any $h_n \in \mathcal{F}$, if we take $y \geq \sqrt[n]{\frac{1}{2}}$, then we have $y^n \geq \frac{1}{2}$. In other words, $|h_n(y) - g(y)| \geq \frac{1}{2}$. So the result of Part (i) does not hold for this example.

(iii) No, a counter-example is given by $\mathcal{F} = \{p_n(x) := \frac{x}{n} | n \in \mathbb{N}\} \subset C[0, \infty)$. The constant function $g(x) = 0$ is continuous and can be realized as the infimum of $\mathcal{F}$, since $\lim_{n\to\infty} \frac{x}{n} = 0$.

For $\epsilon = 1$, and for any $p_n \in \mathcal{F}$, if we take $y \geq n$, then $p_n(y) = \frac{y}{n} \geq 1$, so that $|p_n(y) - g(y)| \geq 1$. The result of Part (i) does not hold for this example.

2. (25 points)

For $x = (x_1, ..., x_m)$ and $y = (y_1, ..., y_m)$ in $\mathbb{R}^m$, let $||x|| := \sqrt{x_1^2 + ... + x_m^2}$ and $\langle x, y\rangle := \sum_{k=1}^m x_k y_k$. Let A be an $m \times m$ matrix and let $B := \{x \in \mathbb{R}^m : ||x|| \leq 1\}$. Define $q : B \to \mathbb{R}$ by

$$q(x) := \langle Ax, x\rangle, \quad x \in B.$$

(i) Show that $\{||Ax|| : x \in \mathbb{R}^m, ||x|| = 1\}$ is bounded.

(ii) Show that the function q is Lipschitz on B, i.e., there is some $C > 0$ such that $|q(x) - q(y)| \leq C||x - y||$ for any $x, y \in B$.

(iii) Show that

$$\sup\left\{\frac{|q(x) - q(y)|}{||x - y||} : x, y \in B, x \neq y\right\} = 2\sup\{|\langle Ax, x\rangle| : x \in \mathbb{R}^m, ||x|| = 1\}.$$

Solutions.

(i) Denote $A = (a_{ij})$, where a_{ij} is the entry at the i-th row and j-th column. Then $y = (y_1, ..., y_m) = Ax$ is a vector whose i-th component is given by $y_i = \sum_{j=1}^m a_{ij} x_j$. Write $M = \max\{a_{ij} : 1 \leq i, j \leq m\}$, then for $x \in \mathbb{R}^m$ with $||x|| = 1$, we have

$$\begin{aligned}
||Ax||^2 = ||y||^2 = \sum_{i=1}^m \left(\sum_{j=1}^m a_{ij} x_j\right)^2 &\leq \sum_{i=1}^m \left(\sum_{j=1}^m |a_{ij}| \cdot |x_j|\right)^2 \\
&\leq \sum_{i=1}^m \left(\sum_{j=1}^m M|x_j|\right)^2 \\
&= mM^2 \sum_{j=1}^m |x_j|^2 = mM^2
\end{aligned}$$

That is, $||Ax|| \leq M\sqrt{m}$ on $||x|| = 1$. So it is bounded.

(ii) Note that the above argument implies that $||Ax|| \leq M\sqrt{m}||x||$. So we have

$$\begin{aligned}
|q(x) - q(y)| &= |\langle Ax, x\rangle - \langle Ax, y\rangle + \langle Ax, y\rangle - \langle Ay, y\rangle| \\
&\leq |\langle Ax, x - y\rangle| + |\langle A(x - y), y\rangle| \\
&\leq ||Ax|| \cdot ||x - y|| + ||A(x - y)|| \cdot ||y|| \\
&\leq M\sqrt{m}||x|| \cdot ||x - y|| + M\sqrt{m}||x - y|| \cdot ||y|| \\
&\leq 2M\sqrt{m}||x - y||.
\end{aligned}$$

The first inequality sign is due to triangle inequality and linearity of inner product. The second inequality sign is the Cauchy-Schwarz inequality. The third inequality is the estimate we obtained from Part (i). The last inequality is from the domain B, where $||x|| \leq 1$.

(iii) We will first prove the result for $A = A^T$ a symmetric matrix. In that case, we will need the following lemma.

Lemma. Let A be a symmetric matrix, then $||A|| := \sup\{||Ax|| : x \in \mathbb{R}^m; ||x|| = 1\} = \sup\{|\langle Ax, x\rangle| : x \in \mathbb{R}^m; ||x|| = 1\}$. The number $||A||$ is usually called the operator norm of A.

Proof. The $(\geq)$ is always true from the Cauchy-Schwarz inequality, regardless of whether A is symmetric, as

$$|\langle Ax, x\rangle| \leq ||Ax|| \cdot ||x|| = ||Ax||.$$

So the LHS is an upper bound of the values $|\langle Ax, x\rangle|$.

For the $(\leq)$ direction. We note that a symmetric matrix over $\mathbb{R}$ is orthogonally diagonalizable, i.e. there exists some orthogonal matrix Q such that $Q^T AQ = D$ is a digaonal matrix. Notice that an orthogonal matrix preserves the standard inner product, i.e. $\langle Qx, Qy\rangle = \langle x, Q^T Qy\rangle = \langle x, y\rangle$. Therefore $Q : \partial B \to \partial B$ is well-defined and is a bijection, i.e. Q preserves the length one vectors. Now consider

$$q_D(x) := \langle Dx, x\rangle = \langle Q^T AQx, x\rangle = \langle AQx, Qx\rangle = q_A(Qx).$$

Since Q is a bijection on $\{x : ||x|| = 1\}$, in particular $\sup\{|\langle Dx, x\rangle| : ||x|| = 1\} = \sup\{|\langle Ay, y\rangle| : ||y|| = 1\}$; and likewise $\sup\{||Dx|| : ||x|| = 1\} = \sup\{||Ay|| : ||y|| = 1\}$ by considering $y = Qx$. Therefore it suffices to prove $(\leq)$ for the diagonal matrix D. Let λ_i be the eigenvalue with respect to the i-th vector in the eigenbasis, suppose $|\lambda_k| = \rho(A) = \max\{|\lambda_i| : i = 1, ..., m\}$. Then, over $||x|| = 1$, we have

$$||Dx|| = \sqrt{\sum_{i=1}^{n} \lambda_i^2 x_i^2} \leq \sqrt{\lambda_k^2 \sum_{i=1}^{n} x_i^2} = |\lambda_k| = |\langle De_k, e_k\rangle|,$$

where e_k is the k-th standard basis vector. This proves the $(\leq)$ direction for D, hence for A. Also note that this argument implies that both of these supremums are in fact equal to $|\lambda_k|$. ■

By the above lemma, it suffices to prove the following equality in the case when A is symmetric:

$$\sup\left\{\frac{|q(x) - q(y)|}{||x - y||} : x, y \in B, x \neq y\right\} = 2||A||.$$

Proof. The $(\leq)$ direction is obtained by the calculation in Part (ii), where we have

$$\begin{aligned} |q(x) - q(y)| &\leq ||Ax|| \cdot ||x - y|| + ||A(x - y)|| \cdot ||y|| \\ &\leq ||A|| \cdot ||x - y|| + ||A\left(\frac{x - y}{||x - y||}\right)|| \cdot ||x - y|| \\ &\leq 2||A|| \cdot ||x - y||. \end{aligned}$$

For the ($\geq$) direction, we take $x = x_k$ an unit length eigenvector for the eigenvalue λ_k, and consider $y = tx_k$ depending on a parameter $t \in (0, 1)$. Then since q is quadratic,

$$\frac{|q(x_k) - q(tx_k)|}{||(1-t)x_k||} = \frac{(1-t^2)|q(x_k)|}{(1-t)||x_k||} = (1+t)|\lambda_k| \to 2|\lambda_k| = 2||A|| \text{ as } t \to 1^-.$$

So the supremum of the values of $\frac{|q(x)-q(y)|}{||x-y||}$ must be at least $2||A||$. ■

This proves the equality of supremums in the case when A is symmetric. The general case follows almost immediately by noting that $q_{A+A^T}(x) = \langle (A + A^T)x, x\rangle = 2\langle Ax, x\rangle = 2q_A(x)$. So we may apply the result for the symmetric case to the symmetric matrix $A + A^T$, i.e. we have

$$\begin{aligned} 2\sup\left\{\frac{|q_A(x) - q_A(y)|}{||x-y||} : x, y \in B; x \neq y\right\} &= \sup\left\{\frac{|q_{A+A^T}(x) - q_{A+A^T}(y)|}{||x-y||} : x, y \in B; x \neq y\right\} \\ &= 2\sup\{|\langle (A + A^T)x, x\rangle| : ||x|| = 1\} \\ &= 4\sup\{|\langle Ax, x\rangle| : ||x|| = 1\}. \end{aligned}$$

Remark: This question is more linear algebra than analysis. Alternatively, you can prove the ($\leq$) direction by a higher dimensional version of the mean value theorem (although we haven't prove this rigorously yet, see MATH2060). The idea is that B is a convex domain, so for any two points $x, y \in B$ where $x \neq y$. We may connect them via the straight line $r(t) = tx + (1-t)y$ for $t \in [0, 1]$, which lie completely inside B. Then apply the mean value theorem on $f(t) = q(r(t))$, which says that $q(x) - q(y) = f(1) - f(0) = f'(t_0)(1-0) = \nabla q(r(t_0))(r'(t_0))$ for some $t_0 \in (0, 1)$. We have $r'(t) = x - y$ independent of t, and $\nabla q(x) = \langle (A + A^T)x, -\rangle$. So the above gives

$$\begin{aligned} |q(x) - q(y)| = |\langle (A + A^T)(r(t_0)), x - y\rangle| &\leq ||A + A^T|| \cdot ||x - y|| \\ &= 2||A|| \cdot ||x - y||. \end{aligned}$$

More generally, this argument can be generalized to the case for $f : K \to \mathbb{R}$ is any differentiable function over a convex compact domain $K \subset \mathbb{R}^m$, such that $\{||\nabla f(x)|| : x \in K\}$ is bounded, then f is Lipschitz with the minimal Lipschitz constant equals to the supremum of $||\nabla f(x)||$. The linear algebra we did for $q(x)$ is essentially trying to figure out what is this supremum of gradient.