Solution to Problem 1 (HW 9)

(with comments)

Kuba Skneczkowski

Let $I(u) = \int_0^1 |u(x)|^{3/2} \cos^2(x)\,dx$, $u \in L^2(0,1)$.

(A) I is continuous on $L^2(0,1)$. We need to prove:

$$u_n \to u \text{ in } L^2(0,1) \Rightarrow I(u_n) \to I(u) \text{ in } \mathbb{R}$$

It is sufficient to prove $I(u_n)^{2/3} \to I(u)^{2/3}$. Note that $I(u)^{2/3}$ defines a norm on $L^{3/2}((0,1), \lambda \cos^2)$ $(\supset L^2(0,1))$. Hence, by Δ inequality

$$| I(u_n)^{2/3} - I(u)^{2/3} | \leq | I(u_n - u) |^{2/3}$$

(as $| \|x\| - \|y\| | \leq \|x-y\|$). Then, by Hölder,

$$I(u_n - u) = \int |u_n - u|^{3/2} \cos^2(x) \leq \quad p = \tfrac{4}{3} \quad q = \text{whatever}$$

$$\leq \left(\int |u_n - u|^2\right)^{3/4} \left(\int_0^1 \cos^{2q}\right)^{1/q} \to 0 \ (n \to \infty).$$

COMMENT: Some of you used inequality

$$\big||a|^{3/2} - |b|^{3/2}\big| \leq |a-b|^{3/2} \qquad (*)$$

which is wrong (take $a = 2$ and $b = 1$). This inequality should be valid with some constant.

(B) $A := \{ (u,\lambda) \in L^2(0,1) \times \mathbb{R} : \ I(u) < \lambda \}$ is open and convex.

Open: Let $A^C := \{ (u,\lambda) \in L^2(0,1) \times \mathbb{R} : I(u) \geq \lambda \}$. We show that A^C is closed. Consider $(u_n, \lambda_n) \to (u,\lambda)$ in $L^2(0,1) \times \mathbb{R}$ s.t. $I(u_n) \geq \lambda_n$. We need $I(u) \geq \lambda$.

As I is continuous, $I(u_n) \to I(u)$. We also have $\lambda_n \to \lambda$ in $\mathbb{R}$. Passing to the limit in $I(u_n) \geq \lambda_n$ we obtain $I(u) \geq \lambda$.

Convex: Let $t \in (0,1)$. Take $(u, \lambda_u), (v, \lambda_v) \in A$. We need $I(tu + (1-t)v) < t\lambda_u + (1-t)\lambda_v$.

$$I(tu+(1-t)v) = \int_0^1 |tu+(1-t)v|^{3/2} \cos^2(x)\,dx$$

$$\leq \int_0^1 \big| t|u| + (1-t)|v| \big|^{3/2} \cos^2(x)\,dx$$

$$\leq t^{3/2} \int_0^1 |u|^{3/2} \cos^2(x)\,dx + (1-t)^{3/2} \int_0^1 |v|^{3/2} \cos^2(x)\,dx$$

(convexity of $x \mapsto |x|^{3/2}$ for $x \geq 0$)

$$\leq t \int_0^1 |u|^{3/2} \cos^2(x)\,dx + (1-t)\int_0^1 |v|^{3/2}\cos^2(x)\,dx$$

(as $t \in (0,1)$).

$$< t\lambda_u + (1-t)\lambda_v \qquad \text{(as } I(u) < \lambda_u,\ I(v) < \lambda_v\text{)}.$$

(c) Let $u \in L^2(0,1)$. Find $v_u \in L^2(0,1)$ s.t. for all $w \in L^2(0,1)$ we have

$$I(u+w) \geq I(u) + \langle v_u, w\rangle .$$

As suggested, we apply Hahn-Banach to the set $A \subset L^2(0,1) \times \mathbb{R}$ and $B = \{(u, I(u))\}$ (singleton). Both are nonempty, convex and disjoint. Moreover, A is open. Hence, there is $\varphi \in (L^2(0,1)\times\mathbb{R})^*$ and $\mu \in \mathbb{R}$ s.t.

$$\forall_{(w,\lambda)\in A} \qquad \varphi(w,\lambda) < \mu \leq \varphi(u, I(u)).$$

We know that $\varphi(w,\lambda) = \tilde{\varphi}(w) + a\cdot\lambda$ where $\tilde{\varphi}$ is is $(L^2(0,1))^*$ and $a \in \mathbb{R}$. Moreover, by RRT, we can write $\tilde{\varphi}(w) = \langle w, v_u\rangle$ for some $v_u \in L^2(0,1)$ (it depends on u because we fixed u in this construction)

Hence

$$\forall_{(\omega,\lambda)\in A} \quad \langle u, v_u \rangle + a\cdot\lambda < \mu \leq \langle u, v_u \rangle + a\cdot I(u)$$

Observe that $(u+\omega, (1+\varepsilon)\, I(u+\omega)) \in A$. Hence,

$$\langle u+\omega, v_u \rangle + a\,(1+\varepsilon)\, I(u+\omega) < \mu \leq \langle u, v_u \rangle + a\cdot I(u)$$

To conclude, we need to know that $a<0$.

Let $\omega = 0$. Then $a(1+\varepsilon)\, I(u) < a\cdot I(u) \Rightarrow a<0$

We divide by $a<0$ to get

$$(1+\varepsilon)\, I(u+\omega) > I(u) + \langle u, \overbrace{\tfrac{v_u}{a}}^{\text{new } v_u} \rangle .$$

As $\varepsilon>0$ is arbitrary, we send $\varepsilon \to 0$ and conclude the proof.

COMMENT (what is v_u) For a convex function $f: \mathbb{R} \to \mathbb{R}$ we have

$$f(y) \geq f(x) + f'(x)(y-x)$$

(second term in Taylor's expansion is ≥ 0 due to convexity). RHS as a function of y is the line supporting f at point x.

Similarly, if $f: \mathbb{R}^n \to \mathbb{R}$ is convex

$$f(y) \geq f(x) + Df(x) \cdot (y-x) \quad (Df \in \mathbb{R}^n).$$

Now, the (RHS) as a function of y is the supporting hyperplane.

So, in this exercise you've found gradient of $J: L^2(0,1) \to \mathbb{R}$.

Such gradients are then used for optimization, see GRADIENT DESCENT METHOD

https://en.wikipedia.org/wiki/Gradient_descent

(The article is concerned with functions on $\mathbb{R}^n$ but similar techniques are developed in Banach spaces).

Last comment concerns generality. One can easily generalize this problem to prove:

THEOREM. Let E be a Banach space, $I: E \to \mathbb{R}$ be continuous and convex. Then, for all $u \in E$, there is $\varphi_u \in E^*$ such that $I(u+w) \geq I(u) + \langle w, \varphi_u \rangle$.