Jensen' inequality

$f : \mathbb{R}^m \longrightarrow \mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$

is convex, i.e., $\forall\, u_0, u_1 \in \mathbb{R}^m,\ \theta \in [0,1]$

$f((1-\theta)u_0 + \theta u_1) \leq (1-\theta) f(u_0) + \theta f(u_1)$

a) Fix $u_1, \dots, u_K \in \mathbb{R}^m,\ \theta_1, \dots, \theta_K \geq 0$

such that $\sum_{k=1}^{K} \theta_k = 1$.

Proof via induction: For $K=2$, the claim follows from the definition of convexity.

Step: $K \to K+1$. W.l.o.g. we assume that $\theta_{K+1} \neq 1$ ($\theta_1 u_1 + .. + \theta_{K+1} u_{K+1} = u_{K+1}$ in that case $\longrightarrow$ trivial)

define $\tilde{u} := \frac{1}{1-\theta_{K+1}} \sum_{i=1}^{K} \theta_i u_i \in \mathbb{R}^m$

With this, we have $\sum_{j=1}^{K+1} \theta_j u_j = (1-\theta_{K+1})\tilde{u} + \theta_{K+1} u_{K+1}$

From the case $K=2$, we get

$$f\left(\sum_{j=1}^{k+1}\theta_j u_j\right) \leq (1-\theta_{k+1})\, f(\tilde{u}) + \theta_{k+1} f(u_{k+1})$$

Consider $\tilde{\theta}_i := \frac{1}{1-\theta_{k+1}}\theta_i \quad i=1,\dots,k$

we check that $\sum_{j=1}^{k}\tilde{\theta}_j = \frac{1}{1-\theta_{k+1}}\sum_{j=1}^{k}\theta_i = \frac{1-\theta_{k+1}}{1-\theta_{k+1}} = \underline{1}$

Thus, we continue our estimate above by using the inequality for k as

$$\begin{aligned} f\left(\sum_{j=1}^{k+1}\theta_j u_j\right) &\leq (1-\theta_{k+1})\sum_{j=1}^{k}\tilde{\theta}_j f(u_j) + \theta_{k+1} f(u_{k+1}) \\ &= \sum_{j=1}^{k}\theta_j f(u_j) + \theta_{k+1} f(u_{k+1}) \end{aligned}$$

$\square$

b) We assume that $\mathrm{vol}(\Omega) < \infty$ and consider $u : \Omega \to \mathbb{R}^m$ of the form

$$u(x) = \sum_{i=1}^{K} u_i \chi_{A_i} \quad \text{with } u_i \in \mathbb{R}^m$$

and $A_i \subset \Omega$ measurable and $A_i \cap A_j = \emptyset$

with $\theta_k := \frac{vol(A_k)}{vol(\Omega)}$ the claim follows from a) since $\frac{1}{vol(\Omega)} \int_\Omega u(x) dx = \sum_{k=1}^{K} \theta_k u_k$

and $\frac{1}{vol(\Omega)} \int_\Omega f(u) dx = \sum_{k=1}^{K} \theta_k f(u_k)$

$\square$

c) Assume that $f: \mathbb{R}^m \to \mathbb{R}$ is continuous.
Originally, the exercise was conceived to use an approximation argument via simple functions. However, the limit $\int_\Omega f(u_k) dx \to \int_\Omega f(u) dx$ is too complicated for the purpose of the exercise (no Lebesgue or Beppo-Levi thm.)

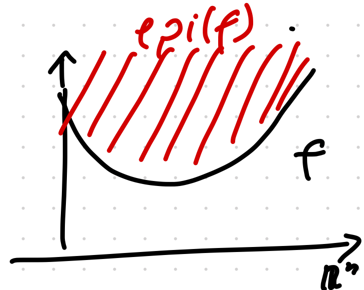
Instead, we use the following more general argument: We show that for every $a \in \mathbb{R}^m$ there exists $\xi_a \in \mathbb{R}^m$ such that

$$\forall v \in \mathbb{R}^m: \quad f(v) \geq f(a) + \xi_a \cdot (v - a) \quad (*)$$

For $v = u(x)$ and $a = \frac{1}{vol(\Omega)} \int_\Omega u(x) dx$ we get after integration over Ω

$$\int_\Omega f(u(x))\,dx \geq \mathrm{vol}(\Omega)\, f\Big(\frac{1}{\mathrm{vol}(\Omega)}\int_\Omega u(x)\,dx\Big) + \xi_a \cdot \underbrace{\Big(\int_\Omega u(x)\,dx - \int_\Omega u(x)\,dx\Big)}_{=0}$$

Thus, we show (*)



define <u>Epigraph</u> of f

$$\mathrm{epi}(f) = \{(u,\alpha) \in \mathbb{R}^m \times \mathbb{R} \mid f(u) \leq \alpha\}$$

check: epi(f) is convex?

It holds that $(u, f(u)+1) \in \mathrm{epi}(f)$

We use "Hyperplane seperation theorem"
(Hahn-Banach)

X Banach space, $A, B \subset X$ non-empty

(i) $\mathrm{int}(A) \neq \emptyset$, (ii) A, B convex

(iii) $B \cap \mathrm{int}(A) = \emptyset$

$$\boxed{\begin{aligned} &\Rightarrow \exists\, L \in X',\ \alpha \in \mathbb{R},\ \beta \in \mathbb{R} \\ &L(b) \lneq \alpha \lneq L(a) \quad \forall\, a \in \mathrm{int}(A),\ b \in B \\ &L(b) \leq \beta \leq L(a) \quad \forall\, a \in A,\ b \in B \end{aligned}}$$

Pick $X = \mathbb{R}^m$, $A = \mathrm{epi}(f)$, $B = \{(u, f(u))\}$

$\Rightarrow \exists\, L = (\tilde{\xi}, \gamma) \in \mathbb{R}^m \times \mathbb{R}$

$(\tilde{\xi} \cdot u, \gamma f(u)) \lneq (\tilde{\xi} \cdot v, \gamma \beta) \quad \forall (v, \beta) \in \mathrm{int}(\mathrm{epi}(f))$

$\quad$ " $\quad \leq \quad$ " $\quad \forall (v, \beta) \in \mathrm{epi}(f)$

choose $(v, \beta) = (u, f(u) + 1) \in \mathrm{int}\ \mathrm{epi}(f)$

it holds then that

$$\tilde{\xi} \cdot u + \gamma \cdot f(u) \lneq \tilde{\xi} \cdot u + \gamma (f(u) + 1)$$

$$\Rightarrow \boxed{\gamma > 1\ !}$$

It follows for every $(v, \beta) \in \mathbb{R}^m \times \mathbb{R}$ that

$f(u) \leq \beta + \frac{1}{\gamma}\tilde{\xi}(v-u)$ such that for

$\beta = I(v)$ we get

$$f(u) \leq f(v) + \underbrace{\frac{1}{\gamma}\tilde{\xi}}_{=:\xi} \cdot (v-u)$$

$\square$

Caution: If f is only lower semicontinuous, one has to show that $\mathrm{int}(\mathrm{epi}(f)) \neq \emptyset$.