Additional information to Problem Set 9

I. Jensen's inequality

(this is an extended version of pp. 29–30 in **Ch. 1.4**)

Jensen's inequality relates the value of a **convex** (resp. **concave**) function of an integral, say $f(\mathbb{E}[X])$, to the integral of this convex (concave) function, $\mathbb{E}[f(X)]$. Along with the Cauchy–Schwartz inequality (see Ch. 1.4 of Lecture Notes), this is one of mostly used inequalities in Analysis and Probability.

In the context of Probability Theory it can be explicitly stated as follows:

1. If $X : \Omega \to \mathbb{R}$ is a random variable and $f : \mathbb{R} \to \mathbb{R}$ is a **convex** (resp. **concave**) function, then

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)] \quad \text{(resp. } f(\mathbb{E}[X]) \geq \mathbb{E}[f(X)]\text{).} \tag{1}$$

2. If $X : \Omega \to \mathbb{R}$ is a **non-constant*** random variable and $f : \mathbb{R} \to \mathbb{R}$ is a **strictly convex** (resp. **strictly concave**) function, then we have the strict inequlaity

$$f(\mathbb{E}[X]) < \mathbb{E}[f(X)] \quad \text{(resp. } f(\mathbb{E}[X]) > \mathbb{E}[f(X)]\text{).} \tag{2}$$

Remarks:

$*$ Here we mean that there is no subset $\Omega_0 \subseteq \Omega$ with $\mathbb{P}(\Omega_0) = 1$ such that $X(\omega) \equiv c = const$ for all $\omega \in \Omega_0$.

$**$ Of course, we have to assume that both $\mathbb{E}[X]$ and $\mathbb{E}[f(X)]$ are well-defined, that is $\mathbb{E}|X| < +\infty$ and $\mathbb{E}|f(X)| < +\infty$.

Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is called:

(i) **convex** (resp. **concave**) if

$$f(tx_1+(1-t)x_2) \leq t\Phi(x_1)+(1-t)f(x_2) \quad \text{(resp. } f(tx_1+(1-t)x_2) \geq tf(x_1)+(1-t)f(x_2)\text{)}$$

for all values $x_1, x_2 \in \mathbb{R}$ and $0 \leq t \leq 1$ (note that the correponding point $y = tx_1 + (1-t)x_2$ belongs to the interval $[x_1, x_2]$);

(ii) **strictly convex** (resp. **strictly concave**) if

$$f(tx_1+(1-t)x_2) < tf(x_1)+(1-t)f(x_2) \quad \text{(resp. } f(tx_1+(1-t)x_2) > tf(x_1)+(1-t)f(x_2)\text{)}$$

for all values of $x_1, x_2 \in \mathbb{R}$ and $0 < t < 1$ (excluding $t = 0$ and $t = 1$, which correspond to the end points x_1 and x_2 of the interval $[x_1, x_2]$!!).

Similar definitions can be given for a function f defined on some (closed or open) interval (e.g. $[a, b]$ or (c, d)) in $\mathbb{R}$, or on the half-line $\mathbb{R}_+ = (0, +\infty)$.

Examples of

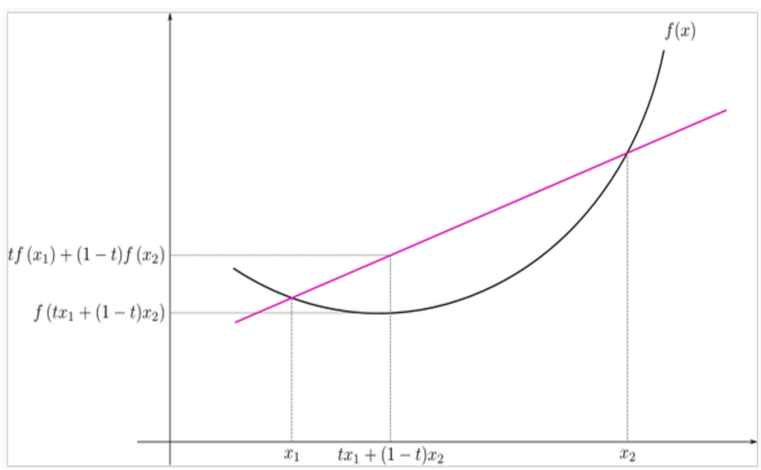
- **strictly convex** functions: $\mathbb{R} \ni x \longmapsto f(x) = x^2;\ x^4;\ x^{2n}$ (for any $n \in \mathbb{N}$); $\exp x$; $\exp(-x)$;

- **strictly concave** functions: $(0, +\infty) \ni x \to f(x) := \frac{1}{x}$; $\log x$; also any inverse $f(x) := \frac{1}{g(x)}$ to a positive, strictly convex functions $\mathbb{R}_+ \ni x \longmapsto g(x) \in \mathbb{R}_+$.

In these particular cases, Jensen's inequality reads as

$$\begin{gathered}
(\mathbb{E}[X])^2 \leq \mathbb{E}[X^2], \quad (\mathbb{E}[X])^{2n} \leq \mathbb{E}[X^{2n}], \\
e^{\mathbb{E}[X]} \leq \mathbb{E}\left[e^X\right], \quad \frac{1}{e^{\mathbb{E}[X]}} \leq \mathbb{E}\left[e^{-X}\right], \\
\log\left(\mathbb{E}[X]\right) \geq \mathbb{E}\left[\log\left(X\right)\right], \\
\frac{1}{\mathbb{E}(X)} \geq \mathbb{E}\left[\frac{1}{X}\right] \iff \mathbb{E}\left[\frac{1}{X}\right] \leq \frac{1}{\mathbb{E}(X)},
\end{gathered}$$

with the strict inequalities ("<" or ">") for non-constant $X : \Omega \to \mathbb{R}$.



Graphical illustration: Jensen's inequality generalizes the statement that a **secant** line (drawn in red) of a convex function lies above the graph.

The word "*convexity*" is one of the most important, natural, and fundamental notions in mathematics. Convex/concave functions play a significant role in mathematical economics, approximation theory, and optimization theory.