Exercise 3

Lorenzo Vigo

Optimization: 26^{th} September 2021

Proof, without using the above theorem, that for any $a \in R$, $f(x) = e^{ax}$ is a convex function.

We will use the definition of convexity that has been seen in class. A function $f : C \longrightarrow \mathbb{R}$ defined on a convex set C is called convex if for all $x_1, x_2 \in C$ and $0 \leq \lambda \leq 1$, it satisfies that:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \tag{1}$$

The given function $f(x) = e^{ax}$ is defined in the convex set $C = \mathbb{R}$. Translated to our function, the expression (1) is equivalent to:

$$e^{a(\lambda x_1+(1-\lambda)x_2)} \leq \lambda e^{ax_1} + (1-\lambda)e^{ax_2} \tag{2}$$

We have to show that expression (2) is true for all $a \in \mathbb{R}$ and all $x_1, x_2 \in \mathbb{R}$. We may assume that $x_1 > x_2$ without any loss of generalization. We will suppose that $a \geq 0$ from now on. The inequality in (2) may be transformed as it follows:

$$\begin{aligned} e^{a(\lambda x_1+(1-\lambda)x_2)} &\leq \lambda e^{ax_1} + (1-\lambda)e^{ax_2} \\ e^{a\lambda x_1}e^{ax_2}e^{-a\lambda x_2} &\leq \lambda e^{ax_1} + (1-\lambda)e^{ax_2} \end{aligned}$$

Now, we may divide both sides by e^{ax_2}, which always is a positive value:

$$e^{a\lambda x_1}e^{-a\lambda x_2} \leq \lambda e^{a(x_1-x_2)} + (1-\lambda)$$

At this point, it is convenient to define $t = (x_1 - x_2) > 0$.

$$e^{a\lambda t} \leq \lambda e^{at} + (1-\lambda) \tag{3}$$

Suppose that $a = 0$. Then:

$$1 \leq \lambda + (1-\lambda)$$

This last expression is trivially true. From now on, we will suppose that $a > 0$. We need to consider the Taylor Series expansions of both $e^{a\lambda x}$ and e^{ax}:

$$e^{a\lambda x} = 1 + \lambda ax + \frac{\lambda^2 a^2 x^2}{2} + \frac{\lambda^3 a^3 x^3}{6} + ... \tag{4}$$

$$e^{ax} = 1 + ax + \frac{a^2 x^2}{2} + \frac{a^3 x^3}{6} + ... \tag{5}$$

We will now express the right side of (3) using (5).

$$\lambda e^{ax} = \lambda + \lambda ax + \lambda\frac{a^2 x^2}{2} + \lambda\frac{a^3 x^3}{6} + ...$$

$$(1-\lambda) + \lambda e^{ax} = 1 + \lambda ax + \lambda\frac{a^2 x^2}{2} + \lambda\frac{a^3 x^3}{6} + ... \tag{6}$$

Evaluating $x = t$, we see that the left side of (3) is equivalent to (4) and the right side of (3) is equivalent to (6). Since $0 \leq \lambda \leq 1$, it is clear that (4) $\leq$ (6) because:

$$\frac{\lambda^2 a^2 x^2}{2} + \frac{\lambda^3 a^3 x^3}{6} + ... \leq \lambda\frac{a^2 x^2}{2} + \lambda\frac{a^3 x^3}{6} + ...$$

The case when $a < 0$ is analogous, but instead of diving both sides by e^{ax_2}, we divide them by $e^{-|a|x_1}$. The sign of t in this case is such that $t < 0$, and the sign of a and t compensate each other.

Therefore, (2) is true and in consequence, $f(x) = e^{ax}$ has been proven to be convex for any $a \in \mathbb{R}$.