VE414 Appendix 9

Jing Liu

UM-SJTU Joint Institute

June 27, 2019

Jensen's Inequality

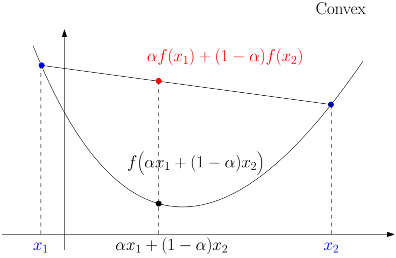
Definition

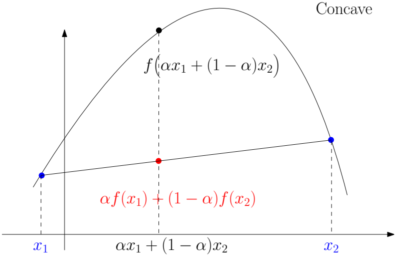
A function $f\colon \mathcal{S} \to \mathbb{R}$ is said to be convex if, for every $x_1, x_2 \in \mathcal{S}$, it satisfies

$$f\big(\alpha x_1 + (1-\alpha)x_2\big) \le \alpha f(x_1) + (1-\alpha)f(x_2) \qquad \text{for all} \quad \alpha \in [0,1]$$

It is said to be concave if, for every $x_1, x_2 \in \mathcal{S}$, it satisfies

$$f\big(\alpha x_1 + (1-\alpha)x_2\big) \ge \alpha f(x_1) + (1-\alpha)f(x_2) \qquad \text{for all} \quad \alpha \in [0,1]$$





Jensen's Inequality

Suppose $g\colon \mathcal{D} \to \mathcal{S}$, where $\mathcal{D}$ is the support of a random variable $\mathbf{Y}$, and $\mathcal{S} \subset \mathbb{R}$.

- If $f: \mathcal{S} \to \mathbb{R}$ is convex, then

$$\mathbb{E}\left[(f \circ g)(\mathbf{y})\right] \geq f\Bigg(\mathbb{E}\left[g(\mathbf{y})\right]\Bigg)$$

- If $f: \mathcal{S} \to \mathbb{R}$ is concave, then

$$\mathbb{E}\left[(f \circ g)(\mathbf{y})\right] \leq f\Bigg(\mathbb{E}\left[g(\mathbf{y})\right]\Bigg)$$

Proof

Let $\ell(z) = a + b \cdot z$ be a line in $\mathbb{R}^2$ tangent to $f(z)$ at the point $z = \mathbb{E}\left[g(\mathbf{y})\right]$.

$$\mathbb{E}\left[(f \circ g)(\mathbf{y})\right] \geq \mathbb{E}\left[(\ell \circ g)(\mathbf{y})\right] = a + b\mathbb{E}\left[g(\mathbf{y})\right] = \ell\left(\mathbb{E}\left[g(\mathbf{y})\right]\right) = f\Bigg(\mathbb{E}\left[g(\mathbf{y})\right]\Bigg)$$

where the first inequality is due to convexity.

Back