Convex Functions

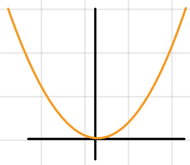
Raphael Joost
Chris Rüttimann

Simple functions

1. $f(x) = x^2, \quad x \in \mathbb{R}$

$\mathrm{dom}(f(x)) = \mathbb{R} \rightarrow$ convex

$f'(x) = 2x, \quad f''(x) = 2 \geq 0 \rightarrow$ convex



2. $f(x) = e^{x^2}, \quad x \in \mathbb{R}$

$\mathrm{dom}(f(x)) = \mathbb{R} \rightarrow$ convex

$f'(x) = 2x \cdot e^{x^2}, \quad f''(x) = \underbrace{4x^2}_{\geq 0} \cdot \underbrace{e^{x^2}}_{\geq 0} + \underbrace{2 \cdot e^{x^2}}_{\geq 0}$ for all x

3. $f(x,y) = x^2 + 3xy + 2y^2, \quad x \in \mathbb{R}, \quad y \in \mathbb{R}$

$\mathrm{dom}(f(x,y)) = \mathbb{R}^2 \rightarrow$ convex

$$\nabla f(x,y) = \left[\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y} \right]^T = \left[2x + 3y, 3x + 4y \right]^T$$

$$\nabla^2 f(x,y) = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x^2} & \frac{\partial f(x,y)}{\partial x \partial y} \\ \frac{\partial f(x,y)}{\partial x \partial y} & \frac{\partial f(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$$

calculate eigenvalues!

$$\det(A - \lambda E) = 0$$

$$(2-\lambda)(4-\lambda) - 9 = 0$$

$$\Rightarrow \lambda^2 - 6\lambda - 1 = 0$$

$$\Rightarrow \lambda_{1,2} = \frac{6 \pm \sqrt{36 + 4}}{2} = 3 \pm \sqrt{10}$$

$\Rightarrow$ One eigenvalue is negative

$\Rightarrow$ It is not positive semidefinite and therefore not convex

Log-sum-exp

$$f(\vec{x}) = \log(e^{x_1} + \dots + e^{x_n}) \qquad x \in \mathbb{R}^n$$

$$e^x > 0 \quad \forall \quad x \in \mathbb{R}$$

$$\nabla f(\vec{x}) = \begin{bmatrix} \frac{\partial f(\vec{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\vec{x})}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{e^{x_1}}{\sum_{i=1}^{n} e^{x_i}} \\ \vdots \\ \frac{e^{x_n}}{\sum_{i=1}^{n} e^{x_i}} \end{bmatrix}$$

Side calculation: $\frac{\partial \left(\log\left(\sum_{i=1}^{n} e^{x_i}\right)\right)}{\partial x_k} = \frac{\partial \log(u)}{\partial u} \cdot \frac{\partial u}{\partial x_k}$ with $u = \sum_{i=1}^{n} e^{x_i}$

$$= \frac{1}{u} \cdot e^{x_k} = \frac{e^{x_k}}{\sum_{i=1}^{n} e^{x_i}}$$

$$\nabla^2 f = \begin{bmatrix} \frac{\partial f(\vec{x})}{\partial x_1^2} & \dots & \frac{\partial f(\vec{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(\vec{x})}{\partial x_1 \partial x_n} & \dots & \frac{\partial f(\vec{x})}{\partial x_n^2} \end{bmatrix} = \begin{bmatrix} \frac{e^{x_1} \cdot \left(\sum_{i=2}^{n} e^{x_i}\right)}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2} \\ \\ \\ \end{bmatrix}$$

Quotient rule: $f(x) = \frac{g(x)}{h(x)}$, $f'(x) = \frac{g'(x)h(x) - g(x)h'(x)}{h(x)^2}$

Side calculation: $\frac{\partial}{\partial x_k} \frac{e^{x_k}}{\sum_{i=1}^{n} e^{x_i}} = \frac{e^{x_k} \cdot \left(\sum_{i=1}^{n} e^{x_i}\right) - e^{x_k} \cdot e^{x_k}}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2} = \frac{e^{x_k} \cdot \left(\left(\sum_{i=1}^{n} e^{x_i}\right) - e^{x_k}\right)}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2}$

$$= \frac{e^{x_k} \cdot \left(\sum_{i=1 \wedge i \neq k}^{n} e^{x_i}\right)}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2}$$

$$\frac{\partial}{\partial x_p} \frac{e^{x_k}}{\sum_{i=1}^{n} e^{x_i}} = \frac{0 - e^{x_k} \cdot e^{x_p}}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2} = -\frac{e^{x_k + x_p}}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2}$$

Geometric mean

It is equivalent to show that
$-f(x)$ is convex

$$-f(x) = -\left(\prod_{i=1}^{n} x_i\right)^{1/n}$$

$$-\left(\partial f / \partial x_k\right) = -\left(\prod_{i=1, i \neq k}^{n} x_i\right)^{1/n} \frac{1}{n} x_k^{1/n - 1}$$

$$-\left(\nabla^2 f(x)\right)_{kj} = -\left(\partial f / \partial x_k \partial x_j\right)$$

$$= -\left(\prod_{i=1, i \neq k, i \neq j}^{n}\right)^{1/n} \frac{1}{n^2} x_k^{1/n-1} x_j^{1/n-1}, \text{ for } k \neq j$$

$$-\left(\nabla^2 f(x)\right)_{kk} = -\left(\prod_{i=1, i \neq k}^{n}\right)^{1/n} \frac{1}{n}\left(\frac{1}{n} - 1\right) x_k^{1/n - 2}$$

$$\Rightarrow -\nabla^2 f(x) = -\frac{f(x)}{n^2} A, \quad A_{ij} = \begin{cases} (1-n) x_i^{-2} & \text{for } i = j \\ x_i^{-1} x_j^{-1} & \text{for } i \neq j \end{cases}$$

substitute $x_i^{-1} \rightarrow y_i$

$$v^T A v = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} v_i v_j$$

$$= \sum_{i=1}^{n} A_{ii} v_i v_i + \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} A_{ij} v_i v_j$$

$$= \underbrace{\sum_{i=1}^{n} \underbrace{(1-n)}_{<0} \underbrace{y_i^2 v_i v_i}_{>0}}_{<0} + \underbrace{\sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \underbrace{y_i y_j}_{>0 \text{ on } \mathbb{R}_{++}} v_i v_j}_{<0 \text{ or } >0}$$

The left term will dominate the right term
therefore is A negative definite
$\Rightarrow -\frac{f(x)}{n^2} A$ is positive definite ($f(x)$ is positive on $\mathbb{R}_{++}$)

$\Rightarrow -f(x)$ convex ✓