Aopr

Schaller, Fontana, Carrel

Exercise 2

Simple functions:

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

- $\text{dom } f(x) = \mathbb{R}$ ✓
- $\nabla f(x) = 2x$

$\nabla^2 f(x) = 2 \geq 0 \quad \forall x \in \text{dom}(f)$ ✓

$\Rightarrow f(x)$ is $\underline{\underline{\text{convex}}}$

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

- $\text{dom } f(x) = \mathbb{R}$ ✓
- $\nabla f(x) = 2x e^{x^2}$

$\nabla^2 f(x) = 4x^2 e^{x^2} + 2e^{x^2}$

$= \underbrace{(4x^2+2)}_{\geq 0}\underbrace{e^{x^2}}_{\geq 0} \geq 0 \quad \forall x \in \text{dom } f(x)$ ✓

$\Rightarrow \underline{\underline{f(x) \text{ convex}}}$

OR by combinaison rule
$\exp(g(x))$ convex if
$g(x)$ convex
$\Rightarrow x^2$ is $\Rightarrow$ ✓

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

- $\text{dom } f(x) = \mathbb{R}^2$ ✓
- $\nabla f(x) = [2x + 3y \quad 4y + 3x]$

$\nabla^2 f(x) = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} = A$. Check if $A \geq 0$ is semi-definite positive or not

Method of pivots:

$\begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} \Rightarrow \begin{pmatrix} 2 & 3 \\ 0 & -0{,}5 \end{pmatrix} \Rightarrow$ One positive and one negative pivot

$\Rightarrow$ one negative eigenvalue $\Rightarrow$ not positive semi-definite $\Rightarrow$ $\underline{\underline{\text{not convex}}}$

Log-sum-exp:

Show that $f(x) = \log(e^{x_1} + \dots + e^{x_n})$ is convex on $\mathbb{R}^n$

Proof:

First for $n=2$ to get an idea

$f(x) = \log(e^{x_1} + e^{x_2})$

$\nabla f(x) = \frac{1}{e^{x_1} + e^{x_2}} \begin{pmatrix} e^{x_1} \\ e^{x_2} \end{pmatrix}$

$\nabla^2 f(x) = \underbrace{\frac{e^{x_1} e^{x_2}}{(e^{x_1} + e^{x_2})^2}}_{\geq 0} \underbrace{\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}}_{= \begin{pmatrix} 1 \\ -1 \end{pmatrix}(1 \;\; -1) = AA^T} $ ✓

Now for $n \in \mathbb{R}_+$:

$$f(x) = \log\left(\sum_{i=1}^{n} e^{x_i}\right)$$

$$\nabla f(x) = \frac{1}{\sum_{i=1}^{n} e^{x_i}} \begin{pmatrix} e^{x_1} \\ \vdots \\ e^{x_n} \end{pmatrix} = \frac{1}{A} a \quad \text{where} \quad a = \begin{pmatrix} e^{x_1} \\ \vdots \\ e^{x_n} \end{pmatrix}, \quad A = \sum_{i=1}^{n} e^{x_i}$$

Calculate the Hessian,

be $g_i(x)$ the i-th component:

$$g_i(x) = \frac{e^{x_i}}{\sum_{i=1}^{n} e^{x_i}} = \frac{a_i}{A}$$

then

- $$\frac{\partial g_i(x)}{\partial x_i} = \frac{a_i}{A} - \frac{a_i^2}{A^2}$$
- $$\frac{\partial g_i(x)}{\partial x_j} = \frac{-a_i a_j}{A^2} \quad j \neq i$$

$$\Rightarrow \nabla^2 f(x) = \frac{1}{A^2}\left(A \operatorname{diag}(a) - aa^T\right)$$

$$= \operatorname{diag}(S(x)) - S(x)S(x)^T \quad \text{where} \quad S(x) = \frac{a}{A} = \begin{pmatrix} \frac{e^{x_1}}{\sum_{i=1}^{n} e^{x_i}} \\ \vdots \\ \frac{e^{x_n}}{\sum_{i=1}^{n} e^{x_n}} \end{pmatrix}$$

We want $\nabla^2 f(x) \succeq 0$, be $v \in \mathbb{R}^n$

$$v^T \nabla^2 f(x) v \geq 0 \iff v^T \operatorname{diag}(S(x)) v \geq v^T S(x) S(x)^T v$$

We have:

$$v^T S(x) S(x)^T v = \left(S(x)^T v\right)^2$$

$$= \left(\sum_{i=1}^{n} S(x)_i v_i\right)^2$$

inequality of Cauchy-Schwarz

$$\leq \sum_{i=1}^{n} S(x)_i v_i^2$$

$$= v^T \operatorname{diag}(S(x)) v$$

$$\Rightarrow f(x) = \log\left(\sum_{i=1}^{n} e^{x_i}\right) \text{ is } \underline{\underline{\text{convex}}} \quad \#$$

Geometric mean:

Show that $f(x) = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$ is concave on $\mathbb{R}^n_{++}$

ie. $\forall x, y \in \mathbb{R}^n_{++}$, $\forall \lambda \in [0,1]$, we have $f(\lambda x + (1-\lambda) y) \geq \lambda f(x) + (1-\lambda) f(y)$

Proof:

Mahler inequality (proven by arithmetic and geometric means)

$$f(\lambda x + (1-\lambda)y) = \prod_{i=1}^{n} (\lambda x_i + (1-\lambda) y_i)^{1/n} \geq \prod_{i=1}^{n} \lambda^{1/n} x_i^{1/n} + \prod_{i=1}^{n} (1-\lambda)^{1/n} y_i^{1/n}$$

$$= \left(\lambda^{1/n}\right)^n \prod_{i=1}^{n} x_i^{1/n} + \left((1-\lambda)^{1/n}\right)^n \prod_{i=1}^{n} y_i^{1/n} = \lambda f(x) + (1-\lambda) f(y) \quad \# \Rightarrow \underline{\underline{\text{concave}}}$$