Convexity of the Log–Sum–Exp Function: A Sequential Proof

Statement. Let $f : \mathbb{R}^N \to \mathbb{R}$ be defined by

$$f(x) \ = \ \log\Big(\sum_{i=1}^{N} e^{x_i}\Big).$$

Then f is *convex* everywhere, provided $x_0 = 0$.

Proof .

Let $S(x) = \sum_{j=1}^{N} e^{x_j}$ and define *softmax* weights

$$p_i(x) \ = \ \frac{e^{x_i}}{S(x)}, \qquad p_i > 0, \qquad \sum_{i=1}^{N} p_i = 1.$$

We compute the gradient:

$$\frac{\partial f}{\partial x_i}(x) \ = \ \frac{e^{x_i}}{\sum_{j=1}^{N} e^{x_j}} \ = \ p_i(x).$$

Hence $\nabla f(x) = (p_1, \dots, p_N)^\top$.

Now we compute the Hessian matrix. First, differentiate $p_i = e^{x_i}/S$ with respect to x_k:

$$\frac{\partial p_i}{\partial x_k} \ = \ \frac{\delta_{ik} e^{x_i}\, S - e^{x_i} e^{x_k}}{S^2} \ = \ p_i(\delta_{ik} - p_k),$$

where δ_{ik} is the Kronecker delta. Therefore

$$\nabla^2 f(x) \ = \ \operatorname{diag}(p) - pp^\top,$$

with $p = (p_1, \dots, p_N)^\top$.

We now show positive semidefiniteness. For any $v \in \mathbb{R}^N$,

$$\begin{aligned} v^\top \nabla^2 f(x)\, v &= \sum_{i=1}^{N} p_i v_i^2 - \Big(\sum_{i=1}^{N} p_i v_i\Big)^2 \\ &= \sum_{i=1}^{N} p_i \Big(v_i - \sum_{j=1}^{N} p_j v_j\Big)^2 \ \ \geq \ 0. \end{aligned}$$

Thus $\nabla^2 f(x)$ is positive semidefinite for every x, so f is convex. Since $x_0 = 0$ and we know that $p_i > 0$, it follwos that $(v_i - \sum_{j=1}^{N} p_j v_j) > 0$ for all $i.j \in N$. Thus, the matrix is positive definite and the statement follows.

□

Results of part 2 exercise 7 are reported below:

Part 2 — Exercise 7 (1-D)

Method	**Points**	**Estimate**	**Abs. error**
True (quadgk tol=1e-14)	adaptive	0.551493272637	0.000000000000
Monte Carlo (n=20)	20	0.569603523713	0.018110251076
Monte Carlo (n=400)	400	0.566812558205	0.015319285568
Gauss-Hermite (n=4)	4	0.551313876771	0.000179395865
Gauss-Hermite (n=5)	5	0.551549103736	0.000055831099
Gauss-Hermite (n=11)	11	0.551493432441	0.000000159804
Gauss-Hermite (n=12)	12	0.551493204065	0.000000068571

Part 2 — Exercise 7 (2-D)

True ≈ GH (35x35)	1225	0.714483805394	0.000000000000
Monte Carlo (n=20)	20	0.668660043026	0.045823762368
Monte Carlo (n=400)	400	0.721492812411	0.007009007017
Gauss-Hermite (4x4)	16	0.714371378972	0.000112426422
Gauss-Hermite (5x5)	25	0.714485614767	0.000001809372
Gauss-Hermite (11x11)	121	0.714483811263	0.000000005868
Gauss-Hermite (12x12)	144	0.714483809858	0.000000004464