Name: **Present group members:**

Worksheet 7-1: Q1

Justify each step (i)–(vi) in the following proof that $f(t) = t^2$ is convex. Assume $\lambda \in [0,1]$ and $t_1, t_2 \in \mathbb{R}$. Then

$$\begin{aligned}
f(\lambda t_1 + (1-\lambda)t_2) &= (\lambda t_1 + (1-\lambda)t_2)^2 && \text{(i)}\\
&= \lambda^2 t_1^2 + 2\lambda(1-\lambda)t_1 t_2 + (1-\lambda)^2 t_2^2 && \text{(ii)}\\
&= \lambda t_1^2 - \lambda(1-\lambda)t_1^2 + 2\lambda(1-\lambda)t_1 t_2 + (1-\lambda)t_2^2 - \lambda(1-\lambda)t_2^2 && \text{(iii)}\\
&= \lambda t_1^2 + (1-\lambda)t_2^2 - \lambda(1-\lambda)(t_1 - t_2)^2 && \text{(iv)}\\
&\le \lambda t_1^2 + (1-\lambda)t_2^2 && \text{(v)}\\
&= \lambda f(t_1) + (1-\lambda)f(t_2). && \text{(vi)}
\end{aligned}$$

(i) Definition of f.

(ii) Expand the square: $(a+b)^2 = a^2 + 2ab + b^2$ with $a = \lambda t_1$ and $b = (1-\lambda)t_2$.

(iii) Substitute $\lambda^2 = \lambda - \lambda(1-\lambda)$ and $(1-\lambda)^2 = (1-\lambda) - \lambda(1-\lambda)$ into step (ii). Both identities follow from factoring: $\lambda^2 + \lambda(1-\lambda) = \lambda(\lambda + (1-\lambda)) = \lambda$.

(iv) Factor $-\lambda(1-\lambda)$ out of the last three terms in step (iii), using $t_1^2 - 2t_1t_2 + t_2^2 = (t_1 - t_2)^2$.

(v) Since $\lambda \in [0,1]$ we have $\lambda(1-\lambda) \ge 0$, and $(t_1 - t_2)^2 \ge 0$, so $\lambda(1-\lambda)(t_1-t_2)^2 \ge 0$; subtracting a non-negative quantity can only decrease the value.

(vi) Definition of f (applied to t_1 and t_2).

For the rest of this worksheet, you can assume that the following "atom" functions are convex.

$$f(t) = mt + b \text{ for any } m, b \in \mathbb{R}. \quad f(t) = t^2 \quad f(t) = e^t$$
$$f(t) = t^4 \quad f(t) = -\ln(t) \text{ for } t > 0$$

Worksheet 7-1: Q2

Show that each of the following functions is convex. Use the theorems from the lecture to justify your answer.

(a) $f(x, y) = x^2 + 3y^2$

- *Let* $\phi(t) = t^2$, *which is convex (one of the atoms).*
- $x^2 = \phi([1\ 0]\,[x\ y]^T)$ *and* $y^2 = \phi([0\ 1]\,[x\ y]^T)$, *so each is convex on* $\mathbb{R}^2$ *by affine composition.*
- $3y^2$ *is convex by non-negative scalar multiplication.*
- *Therefore,* $f(x, y) = x^2 + 3y^2$ *is convex by the sum rule.*

(b) $h(x, y) = \exp(x + y)$.

- *Let* $f_1(x, y) = x + y = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$, *which is affine (hence convex).*
- *Let* $g(t) = \exp(t)$, *which is convex and non-decreasing on* $\mathbb{R}$ *(one of the atoms).*
- *Therefore,* $h(x, y) = g(f_1(x, y)) = \exp(x + y)$ *is convex by the composition rule.*

(c) $f(x, y) = x^2 + 2xy + 3y^2 + 2x - 3y$.

- *The gradient is* $\nabla f(x, y) = \begin{bmatrix} 2x + 2y + 2 \\ 2x + 6y - 3 \end{bmatrix}$.
- *The Hessian is constant:* $\nabla^2 f(x, y) = \begin{bmatrix} 2 & 2 \\ 2 & 6 \end{bmatrix}$.
- *Its eigenvalues are* $\lambda_{1,2} = 4 \pm 2\sqrt{2} > 0$.
- *Hence* $\nabla^2 f(x, y) \succeq 0$ *for all* (x, y), *so* f *is convex.*

(d) $f(x, y, z) = \exp(x - y + z) + \exp 2y + x$.

- *Let* $f_1(x, y, z) = x - y + z$ *and* $f_2(x, y, z) = 2y$; *both are affine (hence convex).*
- *Let* $g(t) = \exp(t)$, *convex and non-decreasing on* $\mathbb{R}$.
- *By composition,* $\exp(x - y + z) = g(f_1(x, y, z))$ *and* $\exp(2y) = g(f_2(x, y, z))$ *are convex.*
- *The function* x *is affine (hence convex), so*
- $f(x, y, z) = \exp(x - y + z) + \exp(2y) + x$ *is convex by the sum rule.*

(e) $f(x, y) = -\ln(xy)$, $x > 0, y > 0$.

- *Rewrite* $f(x, y) = -\ln(xy) = -\ln x - \ln y$ *on the domain* $x > 0,\ y > 0$.
- *Let* $\phi(t) = -\ln t$, *which is convex on* $(0, \infty)$ *(one of the atoms).*
- *Since* x *and* y *are affine in* (x, y), *the compositions* $\phi(x)$ *and* $\phi(y)$ *are convex on this domain.*
- *Therefore,* $f(x, y) = \phi(x) + \phi(y)$ *is convex by the sum rule.*

Worksheet 7-1: Q3

Are the following functions convex? You must justify your answer.

(a) $f(x, y) = xy$.

- *No. The Hessian is* $\nabla^2 f(x, y) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ *which has eigenvalues* $\lambda_1 = 1$ *and* $\lambda_2 = -1$.
- *So, the Hessian is not positive semidefinite, and* f *is not convex.*
- *Even stronger, note that the Hessian has no variables in it, so it is still not positive semidefinite even if you restrict the domain. This means that* $f(x, y) = xy$ *is not convex even if you restrict the domain.*

(b) $\|\mathbf{x}\|$ convex for any norm on $\mathbb{R}^n$. Use the definition of convex function for this part.

- *Pick any* $\lambda \in [0, 1]$ *and* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. *Then*

$$\begin{aligned} \|\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}\| &\leq \|\lambda \mathbf{x}\| + \|(1 - \lambda)\mathbf{y}\| \quad \textit{(by triangle inequality)} \\ &= \lambda \|\mathbf{x}\| + (1 - \lambda)\|\mathbf{y}\| \quad \textit{(by homogeneity of norms).} \end{aligned}$$

(c) $h(\mathbf{x}) = \exp(\|\mathbf{x}\|)$.

- *The function $h(\mathbf{x}) = \exp(\|\mathbf{x}\|)$ is convex because:*
  - $\|\mathbf{x}\|$ *is convex (as shown in the previous part).*
  - $g(t) = \exp t$ *is convex and non-decreasing in t (one of the atoms).*
  - *So, $h(\mathbf{x}) = g(\|\mathbf{x}\|)$ is convex by the composition with a non-decreasing convex function.*

(d) $\|\mathbf{x}\|^2$ for any norm on $\mathbb{R}^n$.

- *The function $h(\mathbf{x}) = \|\mathbf{x}\|^2$ is convex because:*
  - $\|\mathbf{x}\|$ *is convex (as shown in part (b)).*
  - $g(t) = t^2$ *is convex and non-decreasing in t for $t \geq 0$ (one of the atoms).*
  - *So, $h(\mathbf{x}) = g(\|\mathbf{x}\|)$ is convex by the composition with a non-decreasing convex function.*