Homework 1

1) f and g are convex functions: $\mathbb{R}^n \to \mathbb{R}$

$\forall \lambda \in [0,1]$, $x, y \in \mathbb{R}^n$ $\quad \lambda f(x) + (1-\lambda) f(y) \geq f(\lambda x + (1-\lambda) y)$

$\quad \lambda g(x) + (1-\lambda) g(y) \geq g(\lambda x + (1-\lambda) y)$

$+$ $\lambda f(x) + (1-\lambda) f(y) + \lambda g(x) + (1-\lambda) g(y) \geq f(\lambda x + (1-\lambda) y) + g(\lambda x + (1-\lambda) y)$

$\lambda (f+g)(x) + (1-\lambda)(f+g)(y) \geq (f+g)(\lambda x + (1-\lambda) y)$

$\hookrightarrow$ so the function $f+g : \mathbb{R}^n \to \mathbb{R}$ is convex

$\square$

2) $f(x) = \log \sum_{i=1}^{n} e^{x_i} = \log(e^{x_1} + \dots + e^{x_n})$

$$\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i} \log(e^{x_1} + \dots + e^{x_n}) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}$$

$$\nabla f(x) = \frac{1}{\sum_{i=1}^{n} e^{x_i}} \begin{pmatrix} e^{x_1} \\ e^{x_2} \\ \vdots \\ e^{x_n} \end{pmatrix}$$

$$\frac{\partial^2 f}{\partial x_i^2} = \frac{\partial}{\partial x_i} \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} = \frac{e^{x_i} \cdot \sum_{j=1}^{n} e^{x_j} - e^{2x_i}}{\left(\sum_{j=1}^{n} e^{x_j}\right)^2}$$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_j} \frac{e^{x_i}}{\sum_{k=1}^{n} e^{x_k}} = - \frac{e^{x_i} \cdot e^{x_j}}{\left(\sum_{j=1}^{n} e^{x_j}\right)^2}$$

$$H_f = -\frac{1}{\left(\sum_{k=1}^{n} e^{x_k}\right)^2} \begin{pmatrix} e^{2x_1} - e^{x_1} \sum_{k=1}^{n} e^{x_k} & e^{x_1+x_2} & \cdots & e^{x_1+x_n} \\ e^{x_1+x_2} & e^{2x_2} - e^{x_2} \sum_{k=1}^{n} e^{x_k} & & \vdots \\ \vdots & & \ddots & \\ e^{x_1+x_n} & \cdots & & e^{2x_n} - e^{x_n} \sum_{k=1}^{n} e^{x_k} \end{pmatrix}$$

3) $f(x) = \dfrac{1}{1+e^{-x}}$ $\qquad f'(x) = f(x)\cdot(1-f(x))$

$$f'(x) = +\frac{e^{-x}}{(1+e^{-x})^2} = \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}}{1+e^{-x}} =$$

$$= \frac{1}{1+e^{-x}}\left(1 - \frac{1}{1+e^{-x}}\right) = f(x)\cdot(1-f(x))$$

$\square$

5) $f: \mathbb{R}^n \to \mathbb{R}$

$$f(x) = \sum_{i=1}^{m} \log\left(1+e^{a_i^T x}\right)$$

$$\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\log\left(1+e^{a_1^T x}\right) + \dots + \log\left(1+e^{a_m^T x}\right)\right) =$$

$$= \frac{\partial}{\partial x_i}\left(\log\left(1+e^{a_{11}x_1+\dots+a_{1n}x_n}\right) + \dots + \log\left(1+e^{a_{m1}x_1+\dots+a_{mn}x_n}\right)\right) =$$

$$= \frac{a_{1i}\cdot e^{a_1^T x}}{1+e^{a_1^T x}} + \dots + \frac{a_{mi}\cdot e^{a_m^T x}}{1+e^{a_m^T x}} = \cancel{\frac{1}{1+e^{a_i^T x}}} \cdot \sum_{j=1}^{m}\left(\frac{a_{ji}\cdot e^{a_j^T x}}{1+e^{a_j^T x}}\right)$$

$$\cancel{\nabla f(x) = \left(\frac{\sum_{j=1}^{m} a_{j1}\cdot e^{a_j x}}{1+e^{a}}\right.}$$

$$\nabla f(x) = \left(\sum_{j=1}^{m}\frac{a_{j1}\cdot e^{a_j^T x}}{1+e^{a_j^T x}}, \dots, \sum_{j=1}^{m}\frac{a_{jn}\cdot e^{a_j^T x}}{1+e^{a_j^T x}}\right)^T$$

6) a) $f: \mathbb{R}^n \to \mathbb{R} \qquad f(x) = \log(1 + \|x\|^2)$

$$\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i} \log\left(1 + x_1^2 + \dots + x_n^2\right) = \frac{2x_i}{(1+\|x\|^2)}$$

$$\nabla f(x) = \frac{2}{(1+\|x\|^2)} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \frac{2x}{(1+\|x\|^2)}$$

~~b)~~ $\cancel{f: \mathbb{R}^n \to \mathbb{R} \qquad f(x) = \frac{1}{4}\|Ax - b\|^4}$

$$\cancel{\nabla f(x) = 4 \cdot \frac{1}{4} \cdot (Ax-b)^3 \, \nabla(Ax-b) = 2 \cdot A^T \cdot (Ax-b)^3}$$

$$\cancel{\nabla f(x) = \nabla_x(Ax-b) \; 4 \; \frac{1}{4} \; (Ax-b)^3 = A^T \cdot (Ax-b)^3}$$

c) $f: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \qquad f(x,y) = \frac{1}{2}\|xy^T - A\|^2$

$$\nabla_x f(x,y) = 2 \cdot \frac{1}{2} \cdot (xy^T - A) \cdot \nabla_x(xy^T - A) = (xy^T - A)y$$

$$\nabla_y f(x,y) = 2 \cdot \frac{1}{2} \cdot (xy^T - A) \cdot \nabla_y(xy^T - A) = (xy^T - A)x$$

$$\nabla f(x,y) = \begin{pmatrix} (xy^T - A)y \\ (xy^T - A)x \end{pmatrix}$$

* because $y_1, y_2 \geq 0$, $\lambda, 1-\lambda \geq 0$, but only one of them can be 0, not both

7) a) $f: \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R} \qquad f(x,y) = \frac{x^2}{y}$

We need that for all $x_1, x_2 \in \mathbb{R}$, $y_1, y_2 \in \mathbb{R}_{++}$, $\lambda \in [0,1]$:

$$\lambda \frac{x_1^2}{y_1} + (1-\lambda)\frac{x_2^2}{y_2} \geq \frac{(\lambda x_1 + (1-\lambda)x_2)^2}{\lambda y_1 + (1-\lambda) y_2}$$

· $y_1 > 0$
· $y_2 > 0$
· $(\lambda y_1 + (1-\lambda) y_2) > 0$, because *

$$\lambda^2 x_1^2 y_1 y_2 + \lambda(1-\lambda) x_1^2 y_2^2 + \lambda(1-\lambda) x_2^2 y_1^2 + (1-\lambda)^2 x_2^2 y_1 y_2$$
$$\geq$$
$$\lambda^2 x_1^2 y_1 y_2 + 2\lambda(1-\lambda) x_1 x_2 y_1 y_2 + (1-\lambda)^2 x_2^2 y_1 y_2$$

$$\lambda(1-\lambda) x_1^2 y_2^2 + \lambda(1-\lambda) x_2^2 y_1^2 \geq 2\lambda(1-\lambda) x_1 x_2 y_1 y_2$$

if $\lambda = 0$ or 1, this holds

otherwise:

$$x_1^2 y_2^2 + x_2^2 y_1^2 \geq 2 x_1 x_2 y_1 y_2$$

$$(x_1 y_2 - x_2 y_1)^2 \geq 0 \checkmark \to \text{trivial}$$

$\square$

b) $f: \mathbb{R}_{++} \times \mathbb{R}_{++} \to \mathbb{R}$

$$f(x, y) = x \cdot \log x - x \cdot \log y$$

~~We need that for all~~ $x_1, x_2, y_1, y_2 \in \mathbb{R}_{++}, \lambda \in [0, 1]$

$$\lambda \cdot x_1 \cdot \log x_1 - \lambda \cdot x_1 \cdot \log y_1 + (1-\lambda) \cdot x_2 \cdot \log x_2 - (1-\lambda) \cdot x_2 \cdot \log y_2$$

$$\geq$$

$$(\lambda \cdot x_1 + (1-\lambda) x_2) \left(\log(\lambda x_1 + (1-\lambda) x_2) - \log(\lambda y_1 + (1-\lambda) y_2) \right)$$

$$\log\left(\left(\frac{x_1}{y_1}\right)^{\lambda x_1} \cdot \left(\frac{x_2}{y_2}\right)^{(1-\lambda) x_2} \right) \geq \log\left(\left(\frac{\lambda x_1 + (1-\lambda) x_2}{\lambda y_1 + (1-\lambda) y_2} \right)^{\lambda x_1 + (1-\lambda) x_2} \right)$$

We have to check if the Hessian is PSD:

$$\frac{\partial f}{\partial x} = \log x + 1 - \log y$$

$$\frac{\partial f}{\partial y} = -\frac{x}{y}$$

$$\frac{\partial^2 f}{\partial^2 x} = \frac{1}{x} \qquad \Rightarrow H_f = \begin{pmatrix} \frac{1}{x} & -\frac{1}{y} \\ -\frac{1}{y} & \frac{x}{y^2} \end{pmatrix}$$

$$\frac{\partial^2 f}{\partial x \partial y} = -\frac{1}{y}$$

$$\frac{\partial^2 f}{\partial^2 y} = \frac{x}{y^2}$$

As it is symmetric, it is enough to check if $H_{11} \geq 0$ and $H_{11} \cdot H_{22} - H_{12}^2 \geq 0$ because of Sylvester's criterion:

$\frac{1}{x} \geq 0$ ✓, because $x > 0$

$\frac{1}{x} \cdot \frac{x}{y^2} - \left(-\frac{1}{y}\right)^2 = \frac{1}{y^2} - \frac{1}{y^2} = 0 \geq 0$ ✓ $\Rightarrow$ so f is convex on $\mathbb{R}_{++} \times \mathbb{R}_{++}$

$f: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$

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

$x_1 = 1 \quad x_2 = 0 \quad \lambda = \frac{1}{2} \Rightarrow \lambda x_1 + (1-\lambda) x_2 = \frac{1}{2}$

$y_1 = 0 \quad y_2 = 1 \quad \lambda y_1 + (1-\lambda) y_2 = \frac{1}{2}$

$$\frac{1}{2} \cdot f(1,0) + \frac{1}{2} \cdot f(0,1) \overset{?}{\geq} f\left(\frac{1}{2}, \frac{1}{2}\right)$$

$$0 + 0 \ngeq \frac{1}{4} \Rightarrow \text{not convex}$$

b) $f: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$

$f(x, y) = y \cdot e^x - x \cdot e^y$

$x_1 = 0 \quad x_2 = -1 \quad \lambda = \frac{1}{2}$

$y_1 = 0 \quad y_2 = -1$

$$\frac{1}{2} \cdot f(0,0) + \frac{1}{2} \cdot f(-1, +1) \overset{?}{\geq} f\left(-\frac{1}{2}, +\frac{1}{2}\right)$$

$$0 + \frac{1}{2} \cdot \left(+e^{-1} + e^{1}\right) \geq \frac{1}{2} \cdot e^{-\frac{1}{2}} + \frac{1}{2} \cdot e^{\frac{1}{2}}$$

$$e^{-1} + e \geq e^{-\frac{1}{2}} + e^{\frac{1}{2}}$$

$$e^{-1} - e^{-\frac{1}{2}} + e - e^{\frac{1}{2}} \geq 0$$

$$e^{-\frac{1}{2}}\left(e^{-\frac{1}{2}} - 1\right) - e\left(e^{-\frac{1}{2}} - 1\right) \geq 0$$

$$\left(e^{-\frac{1}{2}} - e^{1}\right)\left(e^{-\frac{1}{2}} - e^{0}\right) \geq 0$$

$x_1 = 2 \quad x_2 = 2 \quad \lambda = \frac{1}{2} \Rightarrow \lambda x_1 + (1-\lambda) x_2 = 2$

$y_1 = 0 \quad y_2 = 2 \quad \lambda y_1 + (1-\lambda) y_2 = 1$

$$\frac{1}{2} \cdot f(2,0) + \frac{1}{2} \cdot f(2,2) \overset{?}{\geq} f(2,1)$$

$$\frac{1}{2} \cdot (-2) + \frac{1}{2} \cdot 0 \overset{?}{\geq} e^2 - 2e$$

$$0 \ngeq (e-1)^2 \Rightarrow \text{not convex}$$

9) $$\min_{\substack{x_1>0 \\ x_2>0}} \frac{1}{x_1 \cdot x_2} + x_1 + x_2$$

find critical points:

$$\frac{\partial f}{\partial x_1} = -\frac{1}{x_1^2 \cdot x_2} + 1 = 0 \Rightarrow x_1^2 \cdot x_2 = 1$$

$$\frac{\partial f}{\partial x_2} = -\frac{1}{x_1 \cdot x_2^2} + 1 = 0 \Rightarrow x_1 \cdot x_2^2 = 1$$

$$\left.\begin{matrix} x_1^2 \cdot x_2 = 1 \\ x_1 \cdot x_2^2 = 1 \end{matrix}\right\} x_1 \cdot \frac{1}{x_1^4} = 1$$

$$\frac{1}{x_1^3} = 1$$

$$\underline{\underline{x_1 = 1}}$$

$$\underline{\underline{x_2 = 1}}$$

Check if local minimum:

$$\frac{\partial^2 f}{\partial^2 x_1} = \frac{2}{x_1^3 \cdot x_2}$$

$$\frac{\partial^2 f}{\partial^2 x_2} = \frac{2}{x_1 \cdot x_2^3} \rightarrow D(1,1) = \frac{2}{1^3 \cdot 1} \cdot \frac{2}{1 \cdot 1^3} - \left(\frac{1}{1^2 \cdot 1^2}\right)^2 = 3 > 0$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{1}{x_1^2 \cdot x_2^2} \qquad \frac{\partial^2 f(1,1)}{\partial^2 x_1} = \frac{2}{1^3 \cdot 1} = 2 > 0 \Rightarrow \text{local minimum}$$

Prove that f is convex

Because of Sylvester's criterion it is enough to check if $H_{11} \geq 0$ and $H_{11} \cdot H_{22} - H_{12}^2 \geq 0$, so H_f is PSD, so f is convex.

$$\frac{2}{x_1^3 \cdot x_2} \geq 0 \checkmark \text{, because } x_1, x_2 > 0$$

$$\frac{2}{x_1^3 \cdot x_2} \cdot \frac{2}{x_1 \cdot x_2^3} - \left(\frac{1}{x_1^2 \cdot x_2^2}\right)^2 = \frac{4}{x_1^4 \cdot x_2^4} - \frac{1}{x_1^4 \cdot x_2^4} = \frac{3}{x_1^4 \cdot x_2^4} \geq 0 \checkmark$$

thus H_f is PSD, thus f is convex, thus every local minimum is global minimum, so the solution of the problem: $(x_1, x_2) = (1,1)$

the minimum value: $\frac{1}{1\cdot 1} + 1 + 1 = \underline{\underline{3}}$

10) $f(x) = \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle + c$

We know that $\underset{\exists d \in \mathbb{R}}{}$ $f(x) \geq d \quad \forall x \in \mathbb{R}^n$

we need that A is PSD $\Leftrightarrow \langle Ax, x \rangle \geq 0 \ \forall x \in \mathbb{R}^n$

Let's suppose $\exists x' \in \mathbb{R}^n$ s.t. $\langle Ax', x' \rangle < 0$

in this case $\forall \lambda > 0$: $\langle A\lambda x', \lambda x' \rangle = \lambda^2 \cdot \langle Ax, x \rangle < 0$

and $\langle b, \lambda x' \rangle = \lambda \langle b, x' \rangle$, so even if $\langle b, x' \rangle > 0$,

we will be able to find a λ such that

$\lambda^2 \cdot \langle Ax', x' \rangle + \lambda \langle b, x' \rangle < d - c$, which is a contradiction,

so such x' cannot exist $\rightarrow \langle Ax, x \rangle \geq 0 \ \forall x \in \mathbb{R}^n$

$\hookrightarrow$ A is $\underline{\underline{\text{PSD}}}$

12) $f(x,y) = \sin(\pi x y^2)$

linear approximation of f at $(1,1)$:

$L_1(x,y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$

$f_x(x,y) = \pi y^2 \cdot \cos(\pi x y^2)$ $\qquad$ $f_y(x,y) = 2\pi x y \cos(\pi x y^2)$

$f_x(1,1) = \pi \cdot 1^2 \cdot \cos\pi = -\pi$ $\qquad$ $f_y(1,1) = 2\pi \cdot 1 \cdot 1 \cdot \cos\pi = -2\pi$

$f(1,1) = \sin\pi = 0$

$L(x,y) = 0 + (-\pi)\cdot(x-1) + (-2\pi)\cdot(y-1) = -\pi x - 2\pi y + 3\pi$

$L(1,1) = -\pi - 2\pi + 3\pi = 0$

so $f(1,1) = L(1,1)$

$\hookrightarrow$ it is always equal at (x_0, y_0)

13) $Q(x,y) = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) +$

$$+ \frac{1}{2}\left(f_{xx}(x_0,y_0)(x-x_0)^2 + 2f_{xy}(x_0,y_0)(x-x_0)(y-y_0) + f_{yy}(x_0,y_0)(y-y_0)^2\right)$$

$$f_{xx}(x,y) = -\pi^2 \cdot y^4 \cdot \sin(\pi x y^2)$$

$$f_{xx}(1,1) = -\pi^2 \cdot 1^4 \cdot \sin\pi = 0$$

$$f_{yy}(x,y) = 2\pi x \cos(\pi x y^2) - 4\pi^2 x^2 y^2 \sin(\pi x y^2)$$

$$f_{yy}(1,1) = 2\pi \cdot 1 \cdot \cos\pi - 4\pi^2 \cdot 1 \cdot \sin\pi = -2\pi$$

$$f_{xy}(x,y) = 2\pi y \cos(\pi x y^2) - 2\pi^2 x y^2 \overset{\sin}{\cancel{\cos}}(\pi x y^2)$$

$$f_{xy}(1,1) = 2\pi \cdot 1 \cdot \cos\pi - 2\pi^2 \cdot 1 \cdot \overset{\sin}{\cancel{\cos}}\pi = -2\pi$$

$$Q(x,y) = -\pi x - 2\pi y + 3\pi + \frac{1}{2}\cdot\left(0\cdot(x-1)^2 + 2\cdot(-2\pi)\cdot(x-1)(y-1) - 2\pi\cdot(y-1)^2\right) =$$

$$= -\pi x - 2\pi y + 3\pi + 2\pi x + 2\pi y - 2\pi x y - 2\pi - \pi y^2 + 2\pi y - \pi =$$

$$= -\pi y^2 - 2\pi x y + \pi x + 2\pi y$$

$$Q(1,1) = -\pi - 2\pi + \pi + 2\pi = 0$$

So $f(1,1) = Q(1,1)$ → it is always true at (x_0,y_0)

14) $f(x,y) = \sin(x^2+y) + y$

$$\left.\begin{aligned} \frac{\partial f}{\partial x} &= 2x\cdot\cos(x^2+y) = 0 \\ \frac{\partial f}{\partial y} &= \cos(x^2+y) + 1 = 0 \end{aligned}\right\}$$

if $x = 0$:

$\cos y + 1 = 0$

$\cos y = -1 \Rightarrow y = \pi + k\cdot 2\pi$

$z \in \mathbb{Z}$

~~$y = +\frac{\pi}{[illegible]} + k\cdot 2\pi \mid k \in \mathbb{Z}$~~

if $x \neq 0$:

$\cos(x^2+y) = 0$

$\cos(x^2+y) = -1$ ↯

So the critical points:

$$(x,y) = \left(0, +\frac{\pi}{[illegible]} + k\cdot 2\pi\right)$$

$k \in \mathbb{Z}$

16) $1-t \leq e^{-t}$

$$f(t) = e^{-t} + t - 1$$

$$\frac{\partial f}{\partial t} = -e^{-t} + 1 = 0 \iff e^{-t} = 1$$
$$t = 0$$

$$\frac{\partial^2 f}{\partial^2 t} = e^{-t} \Rightarrow \frac{\partial^2 f(0)}{\partial^2 t} = e^0 = 1 \Rightarrow \text{it is a } \overbrace{\ }^{\text{local}} \text{ minimum}$$

$$\Downarrow$$

$$e^{-t} > 0 \quad \forall t \in \mathbb{R} \Rightarrow f \text{ is convex on } \mathbb{R}$$

So all local minimums are global minimums, thus

$$f(t) \geq 0 \quad \forall t \iff 1-t \leq e^{-t} \quad \forall t$$

$\square$

6) b) $f: \mathbb{R}^n \to \mathbb{R}$

$$f(x) = \frac{1}{4} \|Ax - b\|^4$$

$$g(x) := \frac{1}{2} \cdot \|Ax - b\|^2 \rightarrow f(x) = g^2(x)$$

$$\nabla f(x) = 2 \cdot g(x) \cdot \nabla g(x)$$

From the lectures we know that $\nabla g(x) = A^T \cdot (Ax - b)$, hence

$$\nabla f(x) = 2 \cdot \frac{1}{2} \cdot \|Ax - b\|^2 \cdot A^T \cdot (Ax - b) = \|Ax - b\|^2 \cdot A^T \cdot (Ax - b)$$

11) a) $f(x) = \frac{1}{2}\|Ax - b\|^2$

We know that for $\forall A, x$: $\|Ax\| \le \|A\| \cdot \|x\|$ (for $\| \|_2$)

We also know that $\nabla f(x) = A^T(Ax - b)$

$$\|\nabla f(x) - \nabla f(y)\| = \|A^T(Ax - b) - A^T(Ay - b)\| =$$

$$= \|A^TAx - A^TAy + A^Tb - A^Tb\| = \|A^TA(x - y)\| \le$$

$$\le \|A^TA\| \cdot \|x - y\| \Rightarrow \text{so for } \forall L \ge \|A^TA\| \text{ f is L-smooth}$$

b) $f(x) = \sqrt{1 + \|x\|^2}$

$$\nabla f(x) = \frac{x}{\sqrt{1 + \|x\|^2}}$$

We need: $\forall x, y \in \mathbb{R}^n$:

$$\left\| \frac{x}{\sqrt{1+\|x\|^2}} - \frac{y}{\sqrt{1+\|y\|^2}} \right\| \le L \cdot \|x - y\|$$

$$\| \ \|^2 \le \| \ \|^2$$

$$\left(\frac{x}{\sqrt{1+\|x\|^2}} - \frac{y}{\sqrt{1+\|y\|^2}}\right)^T \left(\frac{x}{\sqrt{1+\|x\|^2}} - \frac{y}{\sqrt{1+\|y\|^2}}\right) \le L \cdot (x - y)^T(x - y)$$

$$\frac{x^Tx}{1+\|x\|^2} - \frac{2}{\sqrt{1+\|x\|^2}\sqrt{1+\|y\|^2}} x^Ty + \frac{y^Ty}{1+\|y\|^2} \le L\left(x^Tx - 2x^Ty + y^Ty\right)$$

L = 1 is a good choice: $\frac{x^Tx}{1+\|x\|^2} \le x^Tx$, because $1 + \|x\|^2 \ge 1$

$\frac{y^Ty}{1+\|y\|^2} \le y^Ty$, because $1 + \|y\|^2 \ge 1$

$-\frac{2\,x^Ty}{\sqrt{1+\|x\|^2}\sqrt{1+\|y\|^2}} \le -2x^Ty$, because $\sqrt{1+\|x\|^2} \cdot \sqrt{1+\|y\|^2} \ge 1$

So $\forall L \ge 1$ f is L-smooth

c) $f(x) = x^3 \qquad x \in [-3,3]$

$\nabla f(x) = f'(x) = 3x^2$

$\|\nabla f(x) - \nabla f(y)\| = \left|3x^2 - 3y^2\right| = 3\left|x^2 - y^2\right| =$

$= 3|x+y| \cdot |x-y| \leq 3 \cdot |3+3| \cdot |x-y| = 18|x-y|$, so

$\forall L \geq 18 \quad f$ is L-smooth

17) f convex $\Longleftrightarrow \forall x, y \quad f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle \quad \boxed{1)}$

$$\|y\|^2 = \|x + y - x\|^2$$
$$\|y\|^2 = \|x\|^2 + 2\langle x, y-x \rangle + \|y-x\|^2$$
$$\frac{\mu}{2}\|y\|^2 = \frac{\mu}{2}\|x\|^2 + \langle \mu x, y-x \rangle + \frac{\mu}{2}\|y-x\|^2 \quad \boxed{2)}$$

Adding 1) + 2):

$$f(y) + \frac{\mu}{2}\|y\|^2 \geq f(x) + \frac{\mu}{2}\|x\|^2 + \langle \nabla f(x), y-x \rangle + \langle \mu x, y-x \rangle + \frac{\mu}{2}\|y-x\|^2$$

because $\nabla F(x) = \nabla f(x) + \nabla \frac{\mu}{2}\|x\|^2 = \nabla f(x) + \mu x$

$$F(y) \geq F(x) + \langle \nabla F(x), y-x \rangle + \frac{\mu}{2}\|y-x\|^2$$

$\Longleftrightarrow$

F is μ-strongly convex

4) a) $f(x) = \|x\| = \sqrt{x_1^2 + \ldots + x_n^2}$

$f(x)$ is not differentiable at 0, hence it is not differentiable

f is diff. if $\forall x, \forall v \in \mathbb{R}^n$: $\lim\limits_{h\to 0} \frac{|f(x+hv) - f(x)|}{h}$ exists

$x = 0$:

$$\lim_{h\to 0} \frac{|f(x+hv)-f(x)|}{h}\Bigg|_{x=0} = \lim_{h\to 0} \frac{\|hv\|}{h} = \lim_{h\to 0} \frac{|h|\cdot\|v\|}{h} =$$

$= \|v\| \cdot \lim\limits_{h\to 0} \frac{|h|}{h}$, which does not exist, because

$$\lim_{h\to 0+} \frac{|h|}{h} = \lim_{h\to 0+} \frac{h}{h} = 1$$

$$\lim_{h\to 0-} \frac{|h|}{h} = \lim_{h\to 0-} -\frac{h}{h} = -1$$

b) $f(x) = \|x\|_1 = |x_1 + \ldots + x_n|$

$f(x)$ is not differentiable, because it is not differentiable at 0:

$$\lim_{h\to 0} \frac{|f(x+hv)-f(x)|}{h}\Bigg|_{x=0} = \lim_{h\to 0} \frac{|hv|}{h} = |v| \cdot \lim_{h\to 0} \frac{|h|}{h}$$, which

does not exist as we have previously seen

c) $f(x) = \frac{\|x\|^2}{x_1}$

we need $\forall x, v \in \mathbb{R}^n$: keep in mind that f is not defined if $x_1 = 0$

$$\lim_{h\to 0} \frac{|f(x+hv)-f(x)|}{h} = \lim_{h\to 0} \frac{\frac{\|x+hv\|^2}{x_1+hv_1} - \frac{\|x\|^2}{x_1}}{h} =$$

exists

15) $f(x) = x^4$ is a counter - example

$\nabla^2 f(x) = 12x^2 \Rightarrow$ if $x = 0$ $\nabla^2 f(x) \succ 0$ is not true,

but $f(x)$ is strictly convex:

First I show $g(x) = x^2$ is strictly convex:

$\forall x, y, \lambda \in (0,1)$:
$x \neq y$

$$\lambda x^2 + (1-\lambda) y^2 > (\lambda x + (1-\lambda) y)^2$$

$$\lambda x^2 + (1-\lambda) y^2 > \lambda^2 x^2 + 2\lambda(1-\lambda) xy + (1-\lambda)^2 y^2$$

$$\lambda(1-\lambda) x^2 - 2\lambda(1-\lambda) xy + (1-\lambda)\lambda y^2 > 0 \quad / : \lambda \neq 0, \ : (1-\lambda) \neq 0$$

$$x^2 - 2xy + y^2 > 0$$

$$(x-y)^2 > 0 \ \checkmark$$, because $x \neq y$

now x^4:

$\forall x, y, \lambda \in (0,1)$:
$x \neq y$

$$\lambda x^4 + (1-\lambda) y^4 \overset{?}{>} (\lambda x + (1-\lambda) y)^4$$

$$= ((\lambda x + (1-\lambda) y)^2)^2$$

$$< (\lambda x^2 + (1-\lambda) y^2)^2$$ because x^2 is strictly convex

$$< \lambda (x^2)^2 + (1-\lambda)(y^2)^2$$ because x^2 is strictly convex

$$= \lambda x^4 + (1-\lambda) y^4 \ \checkmark$$

$\Downarrow$

$f(x) = x^4$ is strictly convex