Homework 3

1

$\forall \boldsymbol{x}^*, \boldsymbol{y}^* \in M$, consider their convex combination $\theta \boldsymbol{x}^* + \overline{\theta} \boldsymbol{y}^*$. By definition of M, $\boldsymbol{x}^*, \boldsymbol{y}^* \in S$. S is convex, hence $\theta \boldsymbol{x}^* + \overline{\theta} \boldsymbol{y}^* \in S$. Now consider $f(\theta \boldsymbol{x}^* + \overline{\theta} \boldsymbol{y}^*)$. $\forall \boldsymbol{x} \in S, f(\theta \boldsymbol{x}^* + \overline{\theta} \boldsymbol{y}^*) \leq \theta f(\boldsymbol{x}^*) + \overline{\theta} f(\boldsymbol{y}^*)$ by convexity of f. Then $\theta f(\boldsymbol{x}^*) + \overline{\theta} f(\boldsymbol{y}^*) \leq \theta f(\boldsymbol{x}) + \overline{\theta} f(\boldsymbol{x}) = f(\boldsymbol{x}), \forall \boldsymbol{x} \in S$. So we know $\forall \boldsymbol{x} \in S, f(\theta \boldsymbol{x}^* + \overline{\theta} \boldsymbol{y}^*) \leq f(\boldsymbol{x})$, thus $\theta \boldsymbol{x}^* + \overline{\theta} \boldsymbol{y}^* \in M$. So M is convex.

2

Assume $f(\theta_1 \boldsymbol{x} + \overline{\theta}_1 \boldsymbol{y}) < \theta f(\boldsymbol{x}) + \overline{\theta}_1 f(\boldsymbol{y})$ for some θ_1. Without loss of generality, assume $\theta_1 \in (\theta_0, 1]$.
$\theta_0 \boldsymbol{x} + \overline{\theta}_0 \boldsymbol{y} = \frac{\theta_0}{\theta_1}(\theta_1 \boldsymbol{x} + \overline{\theta}_1 \boldsymbol{y}) + (1 - \frac{\theta_0}{\theta_1})\boldsymbol{y}$, a convex combination of $\theta_0 \boldsymbol{x} + \overline{\theta}_0 \boldsymbol{y}$ and $\boldsymbol{y}$. Then by convexity of f we know $f(\theta_0 \boldsymbol{x} + \overline{\theta}_0 \boldsymbol{y}) \leq \frac{\theta_0}{\theta_1} f(\theta_1 \boldsymbol{x} + \overline{\theta}_1 \boldsymbol{y}) + (1 - \frac{\theta_0}{\theta_1}) f(\boldsymbol{y}) < \frac{\theta_0}{\theta_1}(\theta_1 f(\boldsymbol{x}) + \overline{\theta}_1 f(\boldsymbol{y})) + (1 - \frac{\theta_0}{\theta_1}) f(\boldsymbol{y}) = \theta_0 f(\boldsymbol{x}) + \overline{\theta}_0 f(\boldsymbol{y})$. This contradicts with the fact that $f(\theta_0 \boldsymbol{x} + \overline{\theta}_0 \boldsymbol{y}) = \theta_0 f(\boldsymbol{x}) + \overline{\theta}_0 f(\boldsymbol{y})$. Similarly, when $\theta_1 \in [0, \theta_0)$, we can also deduce a contradiction. Thus our assumption does not hold. So $f(\theta \boldsymbol{x} + \overline{\theta} \boldsymbol{y}) = \theta f(\boldsymbol{x}) + \overline{\theta} f(\boldsymbol{y})$ holds for the same $\boldsymbol{x}, \boldsymbol{y}$ and any $\theta \in [0, 1]$.

3

We solve this problem by calculate the Hessian Matrix.

(a).

$H = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, the eigenvalues are 3,2,0. Thus H is positive semidefinite, f is convex.

(b).

$H = \begin{pmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{pmatrix}$, the leading principle minors are both positive, thus H is positive definite, f is convex.

(c).

$H = \begin{pmatrix} 0 & 2x_2 \\ 2x_2 & 2x_1 \end{pmatrix}$, the principle minors are $0, 2x_1, -4x_2^2$. So H is indefinite, f is neither convex nor concave.

(d).

$H = \begin{pmatrix} 0 & -\frac{1}{2}x_2^{-\frac{3}{2}} \\ -\frac{1}{2}x_2^{-\frac{3}{2}} & \frac{3}{4}x_1x_2^{-\frac{5}{2}} \end{pmatrix}$, the principle minors are $0, \frac{3}{4}x_1x_2^{-\frac{5}{2}}, -\frac{1}{4}x_2^{-3}$. So H is indefinite, f is neither convex nor concave.

(e).

$H = \begin{pmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & \alpha(\alpha-1)x_1^{\alpha}x_2^{-\alpha-1} \end{pmatrix}$, the principle minors are $0, \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha}, \alpha(\alpha-1)x_1^{\alpha}x_2^{-\alpha-1}$. So H is negative semidefinite, f is concave.

4

Let $\boldsymbol{x} = (x_1, x_2), \boldsymbol{y} = (y_1, y_2)$. Then $f(\theta\boldsymbol{x} + \overline{\theta}\boldsymbol{y}) = f(\theta x_1 + \overline{\theta}y_1, \theta x_2 + \overline{\theta}y_2) = f_1(\theta x_1 + \overline{\theta}y_1) + f_2(\theta x_2 + \overline{\theta}y_2) < \theta f_1(x_1) + \overline{\theta}f_1(y_1) + \theta f_2(x_2) + \overline{\theta}f_2(y_2) = \theta f(\boldsymbol{x}) + \overline{\theta}f(\boldsymbol{y})$, so f is strictly convex. When $f(x_1, x_2) = x_1^2 + x_2^4$, in this case f_1 and f_2 are both univariate functions, $f_1'' = 2 > 0, f_2'' = 12x_2^2 \geq 0$,but f_2 is always positive when $x_2 \neq 0$, so they are both strictly convex, thus f is strictly convex.

5

Necessity:
Since f is a differentiable function with an open convex domain C, $f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}), \forall \boldsymbol{x}, \boldsymbol{y} \in C$, also $f(\boldsymbol{x}) \geq f(\boldsymbol{y}) + \nabla f(\boldsymbol{y})^T(\boldsymbol{x} - \boldsymbol{y}), \forall \boldsymbol{x}, \boldsymbol{y} \in C$. By adding these two inequailities together we get

$$0 \geq \nabla f(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) + \nabla f(\boldsymbol{y})^T(\boldsymbol{x} - \boldsymbol{y}), \forall \boldsymbol{x}, \boldsymbol{y} \in C$$

or equivalently

$$(\nabla f(\boldsymbol{x})^T - \nabla f(\boldsymbol{y})^T)(\boldsymbol{y} - \boldsymbol{x}) \leq 0, \forall \boldsymbol{x}, \boldsymbol{y} \in C$$

$$(\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}))^T(\boldsymbol{x} - \boldsymbol{y}) \geq 0, \forall \boldsymbol{x}, \boldsymbol{y} \in C,$$

which is

$$\langle \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle \geq 0, \forall \boldsymbol{x}, \boldsymbol{y} \in C$$

Sufficiency:
We prove by showing that for any $\boldsymbol{x} \in C$ and any direction $\boldsymbol{d}$, $g(t) = f(\boldsymbol{x} + t\boldsymbol{d})$ is convex on dom g={ $t : \boldsymbol{x} + t\boldsymbol{d} \in$ C }
The hint tells us that the intersection of C with a straight line is an open interval when it is not empty, so by our definition dom g is an open interval. Thus to show $g(t)$ is convex, we only need to show $g'(t)$ is increasing.
$g'(t) = \boldsymbol{d}^T \nabla f(\boldsymbol{x} + t\boldsymbol{d})$, so $\forall t, s \in$ dom g, $(g'(t) - g'(s))(t - s)$

$$= (\boldsymbol{d}^T \nabla f(\boldsymbol{x} + t\boldsymbol{d}) - \boldsymbol{d}^T \nabla f(\boldsymbol{x} + s\boldsymbol{d}))(t - s)$$

$$= \boldsymbol{d}^T (\nabla f(\boldsymbol{x} + t\boldsymbol{d}) - \nabla f(\boldsymbol{x} + s\boldsymbol{d}))(t - s))$$

$$= (t\boldsymbol{d}^T - s\boldsymbol{d}^T)(\nabla f(\boldsymbol{x} + t\boldsymbol{d}) - \nabla f(\boldsymbol{x} + s\boldsymbol{d}))$$

$$= \langle \nabla f(\boldsymbol{x} + t\boldsymbol{d}) - \nabla f(\boldsymbol{x} + s\boldsymbol{d}), (t - s)\boldsymbol{d} \rangle$$

But we already have

$$\langle \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle \geq 0, \forall \boldsymbol{x}, \boldsymbol{y} \in C$$

As both $\boldsymbol{x} + t\boldsymbol{d}, \boldsymbol{x} + s\boldsymbol{d} \in C$, $\langle \nabla f(\boldsymbol{x} + t\boldsymbol{d}) - \nabla f(\boldsymbol{x} + s\boldsymbol{d}), (t - s)\boldsymbol{d} \rangle \geq 0$. Which means that $(g'(t) - g'(s))(t - s) \geq 0$. Thus $g'(t)$ is increasing, thus $g(t)$ is convex, by Proposition we have already proven in class, f is convex.