Unconstrained optimization problem. Let $f : \mathbb{R}^n \to \mathbb{R}$. An unconstrained optimization problem is to minimize/maximize $f(\mathbf{x})$, where $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$.
Necessary condition: If $\mathbf{x}^* \in \mathbb{R}^n$ is a local minimum or local maximum point of f over $\mathbb{R}^n$, then

$$(\nabla f)(\mathbf{x}^*) = 0, \quad \text{that is,} \quad \frac{\partial f}{\partial x_1}(\mathbf{x}^*) = \frac{\partial f}{\partial x_2}(\mathbf{x}^*) = \cdots = \frac{\partial f}{\partial x_n}(\mathbf{x}^*) = 0.$$

Sufficient condition: If $\mathbf{x}^* \in \mathbb{R}^n$ satisfies $(\nabla f)(\mathbf{x}^*) = 0$ and f is strictly convex (strictly concave) function in a neighbourhood of $\mathbf{x}^*$, then $\mathbf{x}^*$ is a local minimum (local maximum) point of f over $\mathbb{R}^n$.

Consider the Hessian matrix

$$(\nabla^2 f)(\mathbf{x}^*) = \left[\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)(\mathbf{x}^*)\right]_{n\times n}.$$

If the Hessian matrix is positive definite (negative definite), then f is strictly convex (strictly concave).

Note: The points $\mathbf{x}^*$ satisfying $(\nabla f)(\mathbf{x}^*) = 0$ are called critical points.

$n = 1$. Let $f : \mathbb{R} \to \mathbb{R}$. An unconstrained optimization problem is to minimize/maximize $f(x)$, where $x \in \mathbb{R}$.

Necessary condition: If $x^* \in \mathbb{R}$ is a local minimum or local maximum point of f over $\mathbb{R}$, then

$$f'(x^*) = 0.$$

Sufficient condition: Let $f'(x^*) = f''(x^*) = \cdots = f^{(k-1)}(x^*) = 0$, but $f^{(k)}(x^*) \neq 0$. The point x^* is a

(a) local minimum point if $f^{(k)}(x^*) > 0$ and k is even,
(b) local maximum point if $f^{(k)}(x^*) < 0$ and k is even,
(c) neither a local maximum or local minimum if k is odd.

$n = 2$. Let $f : \mathbb{R}^2 \to \mathbb{R}$. An unconstrained optimization problem is to minimize/maximize $f(x, y)$, where $(x, y) \in \mathbb{R}^2$.

Necessary condition: If $(x^*, y^*) \in \mathbb{R}^2$ is a local minimum or local maximum point of f over $\mathbb{R}^2$, then

$$\left(\frac{\partial f}{\partial x}\right)(x^*, y^*) = \left(\frac{\partial f}{\partial y}\right)(x^*, y^*) = 0.$$

Sufficient condition: The point $(x^*, y^*) \in \mathbb{R}^2$ satisfying the necessary condition is a local minimum (local maximum) point if

$$\begin{bmatrix} \left(\frac{\partial^2 f}{\partial x^2}\right)(x^*, y^*) & \left(\frac{\partial^2 f}{\partial x \partial y}\right)(x^*, y^*) \\ \left(\frac{\partial^2 f}{\partial y \partial x}\right)(x^*, y^*) & \left(\frac{\partial^2 f}{\partial y^2}\right)(x^*, y^*) \end{bmatrix} \qquad \text{(Hessian matrix)}$$

is positive definite (negative definite).

If (x^*, y^*) is a critical point, and if for every $r > 0$, there exist $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$ satisfying $(\alpha_1 - x_1^*)^2 + (\alpha_2 - x_2^*)^2 < r^2$ and $(\beta_1 - x_1^*)^2 + (\beta_2 - x_2^*)^2 < r^2$ such that $f(\alpha_1, \alpha_2) > f(x_1^*, x_2^*)$ and $f(\beta_1, \beta_2) < f(x_1^*, x_2^*)$, then (x_1^*, x_2^*) is called a saddle point. If the determinant of Hessian matrix is negative, then the point (x^*, y^*) is a saddle point.

Note: A real symmetric matrix A of order n is said to be positive definite (P.D.) if $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n$ with $\mathbf{x} \neq 0$. A real symmetric matrix A of order n is said to be negative definite (N.D.) if $\mathbf{x}^T A\mathbf{x} < 0$ for all $\mathbf{x} \in \mathbb{R}^n$ with $\mathbf{x} \neq 0$.

A real symmetric matrix A is positive definite if and only if all eigenvalues of A are positive. A real symmetric matrix A is negative definite if and only if all eigenvalues of A are negative.

A sufficient condition for a real symmetric matrix A to be positive definite is that all leading principal minors are positive, that is,

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \quad \cdots, \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} > 0.$$

A sufficient condition for a real symmetric matrix A to be negative definite is that all leading principal minors alternate in sign starting from negative, that is,

$$a_{11} < 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} > 0, \quad \cdots.$$

Example. Find the natures of the extreme points of the function

$$f(x) = 12x^5 - 45x^4 + 40x^3 + 5.$$

Solution. The necessary condition for extreme points is

$$f'(x) = 60x^4 - 180x^3 + 120x^2 = 60x^2(x-1)(x-2) = 0.$$

The extreme points are $x = 0, 1, 2$.

$$f''(x) = 240x^3 - 540x^2 + 240x.$$

Now $f''(1) = -60 < 0$. So $x = 1$ is a local maximum point. Again $f''(2) = 240 > 0$. So $x = 2$ is a local minimum point. Since $f''(0) = 0$, we require

$$f'''(x) = 720x^2 - 1080x + 240.$$

Since $f'''(0) = 240 \neq 0$ and 3 is odd, $x = 0$ is neither a local minimum point nor a local maximum point. $\square$

Example. Find the natures of extreme points of the function

$$f(x, y) = x^3 + y^3 + 2x^2 + 4y^2 + 6.$$

Solution. The necessary conditions for extreme points are

$$\frac{\partial f}{\partial x} = 3x^2 + 4x = 0, \quad \text{and} \quad \frac{\partial f}{\partial y} = 3y^2 + 8y = 0.$$

The extreme points are $(0, 0), (0, -8/3), (-4/3, 0), (-4/3, -8/3)$. The Hessian matrix is

$$H = \begin{bmatrix} 6x + 4 & 0 \\ 0 & 6y + 8 \end{bmatrix}.$$

The leading principal minors of H are $H_1 = 6x + 4$ and $H_2 = (6x + 4)(6y + 8)$.

1. For (0,0), $H_1 = 4, H_2 = 32$. So H is positive definite and hence (0,0) is a local minimum point.
2. For $(0, -8/3)$, $\det H = -32 < 0$. So $(0, -8/3)$ is a saddle point.
3. For $(-4/3, 0)$, $\det H = -32 < 0$. So $(-4/3, 0)$ is a saddle point.
4. For $(-4/3, 8/3)$, $H_1 = -4, H_2 = 32$. So H is negative definite and hence $(-4/3, -8/3)$ is a local maximum point.

□

Note: We have $f\left(-\frac{4}{3}, r\right) - f\left(-\frac{4}{3}, 0\right) = r^3 + 4r^2 > 0$ for all $r > 0$. Again for every $0 < r < 2$,

$$\begin{aligned} f\left(-\frac{4}{3} + r, 0\right) - f\left(-\frac{4}{3}, 0\right) &= \left(-\frac{4}{3} + r\right)^3 + 2\left(-\frac{4}{3} + r\right)^2 - \left(-\frac{4}{3}\right)^3 - 2\left(-\frac{4}{3}\right)^2 \\ &= r\left[\left(-\frac{4}{3} + r\right)^2 - \frac{4}{3}\left(-\frac{4}{3} + r\right) + \left(-\frac{4}{3}\right)^2\right] + 2r\left(-\frac{8}{3} + r\right) \\ &= r\left(\frac{16}{9} - \frac{8}{3}r + r^2 + \frac{16}{9} - \frac{4}{3}r + \frac{16}{9} - \frac{16}{3} + 2r\right) \\ &= r(r^2 - 2r) < 0. \end{aligned}$$

Therefore every neighbourhood of $\left(-\frac{4}{3}, 0\right)$ contain two points such that the value of f at one point is larger than $f\left(-\frac{4}{3}, 0\right)$, and the value of f at other point is smaller than $f\left(-\frac{4}{3}, 0\right)$. Therefore $\left(-\frac{4}{3}, 0\right)$ is a saddle point. Similar calculations can be done to show that $\left(0, -\frac{8}{3}\right)$ is a saddle point.

Exercises.

1. Find the critical points and their natures for the following functions.
   (a) $f(x) = x^2 - 6x^2 + 9x + 5$.
   (b) $f(x) = 2 + (x-1)^4$.
2. Find two numbers whose difference is 100 and whose product is a minimum.
3. Find two positive numbers whose product is 100 and whose sum is a minimum.
4. Find the point on the parabola $y^2 = 2x$ that is closest to the point (1,4).
5. Find the area of the largest rectangle that can be inscribed in a semicircle of radius 5. (The word inscribed means that the rectangle has two vertices on the semicircle and two vertices on the diameter.)
6. Find the critical points and their natures for the following functions.
   (a) $f(x, y) = x^3 - y^3 - 2xy + 6$.
   (b) $f(x, y) = x^4 - 2x^2 + y^3 - 3y$.
   (c) $f(x, y) = x^2 + y^4 + 2xy$.
   (d) $f(x, y) = y \cos x$.