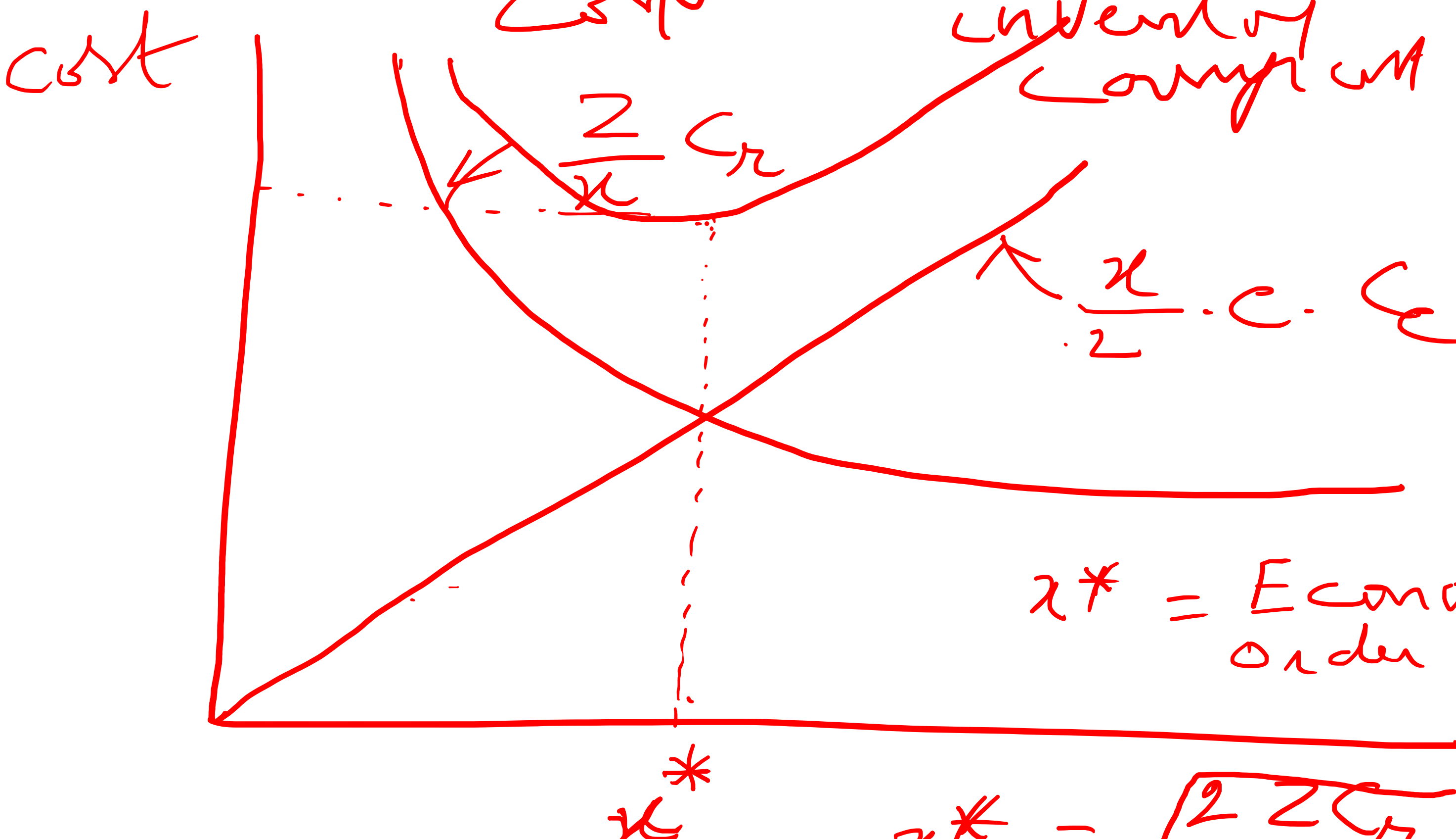
REVIEW OF LINEAR ALGEBRA
CONVEX AND CONCAVE FUNCTIONS

$$TC = \underbrace{\frac{Z}{x} C_r}_{\text{Annual ordering cost}} + \underbrace{\frac{x}{2} \cdot c \cdot C_e}_{\text{Annual inventory carrying cost}}$$

Z = Annual demand

C_r = ordering

$$TC = \underbrace{\frac{Z}{x} C_R}_{\text{Annual ordering cost}} + \underbrace{\frac{x}{2} \cdot C \cdot C_c}_{\text{Annual inventory carrying cost}} + \underbrace{Z \cdot C}_{\text{Annual item cost}}$$



x^* = Economic Order Quantity (EOQ)

$$x^* = \sqrt{\frac{2 Z C_R}{C \cdot C_c}}$$

Average ~~Inventory~~ value $= \dfrac{xc}{2}$

$$\frac{xc}{2} \le \cdots$$ (Constraint on inventory investment)

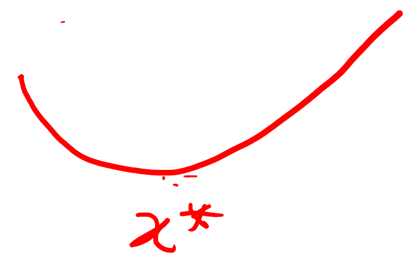
$$(10\,cm^3)\,x \le 1000\,cm^3$$ (Storage space constraint)

Conditions of optimality:

(1) Necessary condition

(2) Sufficiency condition

$f(x)$



convex function

$f(x)$ will have minimum value at x^* iff $\underbrace{\frac{df}{dx} = 0}_{\text{necessary condition}}$ and $\underbrace{\frac{d^2f}{dx^2} = +ve}_{\text{sufficiency condition}}$

$3x^2 - 4x + 3 = 0$

$$\frac{df}{dx} = \underbrace{4x^3 - 3x^2 + x} = 0$$

Numerical solution

Multi-variable unconstr

x^* ?

Multi-variable constrained

single variable unconstraint

single variable constrained

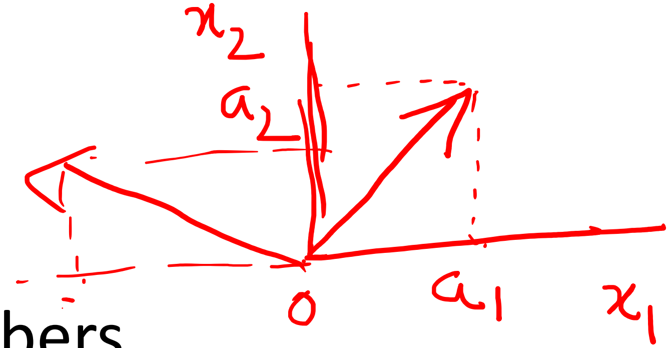
REVIEW OF LINEAR ALGEBRA

- Sets
- Vectors
- Matrices
- Convex Sets

Set Theory

- Set is a well-defined collection of things.
- Example: set S ={ x|x>= 0}, is a set of all non-negative numbers.
  - x = 2 is an element of S. Denoted as 2 ∈ S.
- Union of two sets: is another Set.
  - R = {x|x ∈ P or x ∈ Q or both}
- Intersection of two sets: is another set.
  - R = {x|x ∈ P *and* x ∈ Q}
- Subset: Denoted as $P \subset Q$, every element P is in Q.
- Disjoint Sets: No elements in common.
- Empty Set: Φ

Vectors



- Vector is an ordered set of real numbers.
- $a = (a_1, a_2, \ldots\ldots., a_n)$ is a vector of n elements or components.
- $a = (a_1, a_2, \ldots\ldots., a_n)$, $b = (b_1, b_2, \ldots\ldots., b_n)$
  - $a + b = c = (a_1 + b_1, a_2 + b_2, \ldots\ldots., a_n + b_n)$
  - $a - b = d = (a_1 - b_1, a_2 - b_2, \ldots\ldots., a_n - b_n)$
- For any scalar α positive or negative,
  - $\alpha a = (\alpha a_1, \alpha a_2, \ldots\ldots., \alpha a_n)$
- Vector (0, 0,….,0) is called null vector.
- Inner product or scalar product of two vectors, written as a.b is a number.
  - $a.b = (a_1 b_1, a_2 b_2, \ldots\ldots., a_n b_n)$

$a = (a_1, a_2) = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$

$a = (2, 3)$

$b = (3, 2)$

Vectors

- *Linearly Dependent:* Vectors a_1, a_2, a_3,……, a_n are linearly dependent, if there exist scalars α_1, α_2,…, α_n (not all zero), such that

$$\sum_{i=1}^{n} \alpha_i a_i = 0$$

- One vector can be written as linear combination of others,

$$a_1 = \lambda_2 a_2 + \lambda_3 a_3 + ... + \lambda_n a_n$$

$\underline{E}^n$ or R^n

- *Vector Space:* A set of all n-component vectors (*Euclidean n-space).*
- Spanning: A set of vectors span a vector space *V*, if every vector in *V* can be expressed as linear combination of vectors in that set.
- *Basis:* A set of linearly independent vectors that span the vector space *V*.

Linear Dependence

$$v_1' = \begin{bmatrix} 5 & 12 \end{bmatrix}$$

$$v_2' = \begin{bmatrix} 10 & 24 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 10 \\ 12 & 24 \end{bmatrix} = \begin{bmatrix} v_1' \\ v_2' \end{bmatrix}$$

$$2v_1' - v_2' = 0'$$

$$v_1 = \begin{bmatrix} 2 \\ 7 \end{bmatrix} v_2 = \begin{bmatrix} 1 \\ 8 \end{bmatrix} v_3 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$3v_1 - 2v_2$$

$$= \begin{bmatrix} 6 & 21 \end{bmatrix} - \begin{bmatrix} 2 & 16 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 5 \end{bmatrix} = v_3$$

$$3v_1 - 2v_2 - v_3 = 0$$

Matrices

- A matrix A of size m x n is a rectangular array of numbers with m rows and n columns

$$\underset{(2\,x\,3)}{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$ is a matrix of two rows and three columns

column vector

$\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$ $\begin{bmatrix} 3 \\ 6 \end{bmatrix}$

  - (i, j)th element of A is denoted by a_{ij}.
  - $a_{12} = 2$ and $a_{23} = 6$.
- The elements (a_{ij}) for i = j are called diagonal elements.
- Elements of each column of a matrix is called column vector.
- Elements of each row of a matrix is called row vector.
- A matrix with equal number of rows and columns is called square matrix.

Row vectors $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$

$\begin{pmatrix} 4 & 5 & 6 \end{pmatrix}$

Matrices

- Transpose of A: Matrix obtained by interchanging rows and columns of A.
  - Denoted by A' or A^T. $A^T = [a_{ji}]$
  - $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$
- Symmetric Matrix: $A^T = A$
- Identity Matrix: A square matrix whose diagonal elements are all 1 and off-diagonal elements are all zero. Denoted by I.
- Null Matrix: A matrix whose all elements are zero.

Matrix Operations

- Sum or difference of two matrices A and B is a matrix C.
  - $C = A \pm B$, $c_{ij} = a_{ij} \pm b_{ij}$
- Product AB is defined if and only if, the no. of columns of A is equal to no. of rows of B
- If A is an m x n matrix and B is an n x r matrix, then AB = C is a matrix of size m x r.
- The $(i, j)^{th}$ element of C is given by

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Matrix Addition

If $\quad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

and $\quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$

then $\quad C = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$

Matrix Multiplication

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad [2 \text{ x } 2]$$

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \quad [2 \text{ x } 3]$$

$$C = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \end{bmatrix}$$

[2 x 3]

Matrix Operations

- For any scalar α, $\alpha A = [\alpha a_{ij}]$
- $(A + B) + C = A + (B + C)$
- $A + B = B + A$
- $(A+B)C = AC + BC$
- $AB \neq BA$
- $(AB)C = A(BC)$
- $IA = AI = A$
- $(A + B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$

Determinant of a Square Matrix

- Denoted by $|A|$.
- If A is 2 x 2 matrix, then

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

- If A is n x n matrix, then $|A| = \sum_{i=1}^{n} a_{i1}(-1)^{i+1}|M_{i1}|$, where M_{i1} is a sub-matrix obtained by deleting row i and column 1 of A.
- Singular Matrix: Determinant is zero.
- Non-singular Matrix: Determinant is not zero.

Example

$$\underset{(3x3)}{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$|A| = 1\begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 4\begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} + 7\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}$$

$$= (45 - 48) - 4(18 - 24) + 7(12 - 15) = 0$$

A is singular

Inverse of a Matrix

- The inverse of a matrix A is denoted by A^{-1}.
- It is defined only for non-singular square matrices.
- $AA^{-1} = I$ (identity matrix)

$$Ax = b$$

- Unknown set of values x can be found by inverse of matrix A

$$x = A^{-1}Ax = A^{-1}b$$

Example

Inverse Matrix: 2 x 2

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

In words:

- Take the original matrix.
- Switch *a* and *d*.
- Change the signs of *b* and *c*.
- Multiply the new matrix by 1 over the determinant of the original matrix.

Condition of a Matrix

- Condition of a Matrix: Measure of the numerical difficulties likely to arise when the matrix is used in calculations.
- For example, consider $Q(x) = x^T A^{-1} x$
  - If small changes in x, produce large changes in Q(x), then matrix A is ill-conditioned.
- Condition Number of a Matrix: $K(A) = \left|\frac{\lambda_h}{\lambda_l}\right|$, where λ_h and λ_l are Eigen values of greatest and smallest modulus.
- If condition number is large, then ill-conditioned.
- If condition number is close to 1, then well-conditioned.

Quadratic Forms

- A function of n variables $f(x_1,x_2,\ldots..,x_n)$ is called a quadratic form if
  - $f(x_1, x_2, \ldots., x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij} x_i x_j = x^T Q x$, where $Q_{(nxn)} = [q_{ij}]$ and $x^T = (x_1,x_2,\ldots..,x_n)$.
  - Q is assumed to be symmetric.
  - Otherwise, Q can be replaced with the symmetric matrix $(Q+Q^T)/2$ without changing value of the quadratic form.

Definitions

- A matrix Q is positive definite when $x^TQx > 0$ for all $x \neq 0$.
  - $Q = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ is positive definite.
- A matrix Q is positive semi-definite when $x^TQx \geq 0$ for all x and there exists an $x \neq 0$ such that $x^TQx = 0$.
  - $Q = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ is positive semi-definite.
- A matrix Q is negative definite when $x^TQx < 0$ for all $x \neq 0$.
  - $Q = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ is negative definite.

Definitions

- A matrix Q is negative semi-definite when –Q is positive semi-definite.
  - $Q = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ is negative semi-definite.

- A matrix Q is indefinite when $x^TQx > 0$ for some x and < 0 for other x.
  - $Q = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}$ is indefinite.

Principal Minor

- <u>Principal minor of order k</u>: A sub-matrix obtained by deleting any n-k rows and their corresponding columns from an n x n matrix Q.

$n = 3 \quad k = 1$

- Consider $Q = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

(9) (5) (1) (1,2) (1,3) (2,3)

- Principal minors of order 1 are diagonal elements 1, 5, and 9.
- Principal minors of order 2 are $\begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix}$ and $\begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}$ ✓
- Principal minor of order 3 is Q.
- Determinant of a principal minor is called principal determinant.
    - There are 2^n-1 principal determinants for an n x n matrix.

Leading Principal Minor

- The leading principal minor of order k of an n x n matrix is obtained by deleting the last n-k rows and their corresponding columns.
- $$Q = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$
- Leading principal minor of order 1 is 1.
- Leading principal minor of order 2 is $\begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$
- Leading principal minor of order 3 is Q itself.
- No. of leading principal determinants of an n x n matrix is n.

Tests

- Tests for positive definite matrices
  - All diagonal elements must be positive.
  - All the leading principal determinants must be positive.
- Tests for positive semi-definite matrices
  - All diagonal elements are non-negative.
  - All the principal determinants are non-negative.
- Tests for negative definite and negative semi-definite matrices
  - Test the negative of the matrix for positive definiteness or positive semi-definiteness.
- Test for indefinite matrices
  - At least two of its diagonal elements are of opposite sign.

Positive definiteness

- **Test 2:** Another test that can be used to find the positive definiteness of a matrix **A** of order n involves evaluation of the determinants

$$A_1 = |a_{11}|$$

$$A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

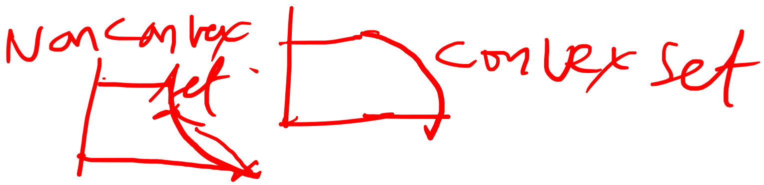
$$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \cdots a_{1n} \\ a_{21} & a_{22} & a_{23} \cdots a_{2n} \\ a_{31} & a_{32} & a_{33} \cdots a_{3n} \\ \vdots & & \\ a_{n1} & a_{n2} & a_{n3} \cdots a_{nn} \end{vmatrix}$$

- The matrix **A** will be **positive definite** if and only if all the values $A_1, A_2, A_3, \ldots A_n$ are positive
- The matrix **A** will be **negative definite** if and only if the sign of A_j is $(-1)^j$ for $j=1,2,\ldots,n$
- If some of the A_j are positive and the remaining A_j are zero, the matrix **A** will be **positive semidefinite**

Convex Sets

- A set S is convex set if for any two points in the set the line joining those two points is also in the set.
- S is a convex set if for any two vectors $x^{(1)}$ and $x^{(2)}$ in S, the vector $x = \lambda\, x^{(1)} + (1-\lambda)\, x^{(2)}$ is also in S for λ between 0 and 1.



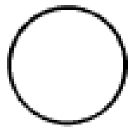


Figure A.1. Convex set.

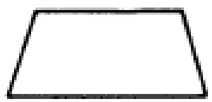


Figure A.2. Convex set.

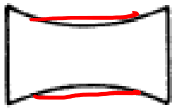


Figure A.3. Nonconvex set.

Convex Sets

- The intersection of convex sets is a convex set (Figure A.4).
- The union of convex sets is not necessarily a convex set (Figure A.4).

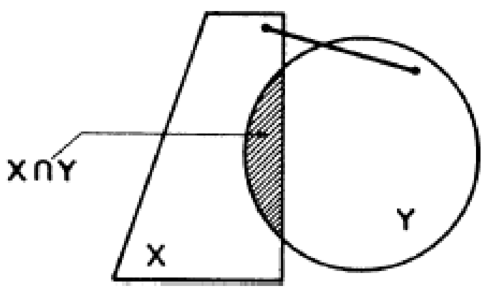


Figure A.4. Intersection and union of convex sets.

Convex Sets

- A convex combination of vectors $x^{(1)}$, $x^{(2)}$,…., $x^{(n)}$ is a vector x such that
  - $x = \lambda_1 x^{(1)} + \lambda_2 x^{(2)} +\ldots+ \lambda_k x^{(k)}$
  - $\lambda_1 + \lambda_2 +\ldots+ \lambda_k = 1$
  - $\lambda_i \geq 0$ for $i = 1,2,\ldots,k$
- Extreme point: A point in convex set that cannot be expressed as the midpoint of any two points in the set
  - Convex set $S = \{(x_1, x_2) | 0 \leq x_1 \leq 2, 0 \leq x_2 \leq 2\}$
  - This set has four extreme points (0,0), (0,2),(2,0) and (2,2)

Convex Function

- Convex function: A function is convex on set D if and only if for any two points $x^{(1)}$ and $x^{(2)} \in D$ and $0 \leq \lambda \leq 1$,
  - $f[\lambda x^{(1)} + (1-\lambda)x^{(2)}] \leq \lambda f(x^{(1)}) + (1-\lambda)f(x^{(2)})$

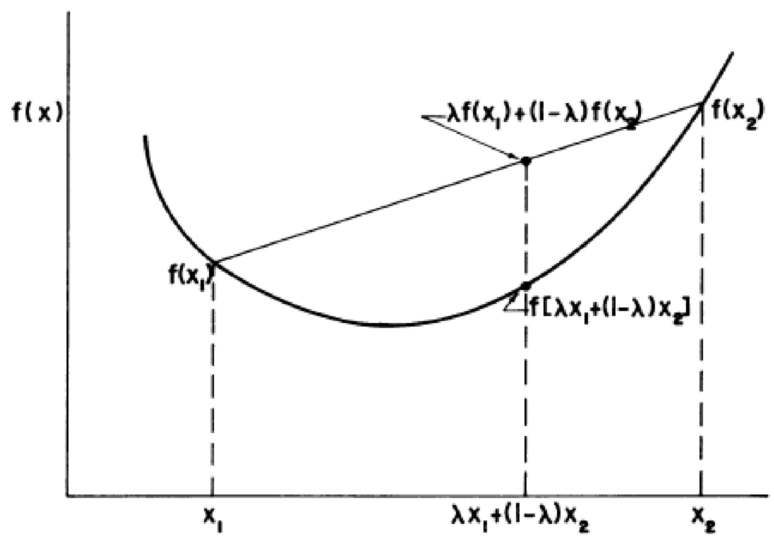


Figure B.1. Convex function.

Properties of Convex Functions

Properties of Convex Functions

1. The chord joining any two points on the curve always falls entirely on or above the curve between those two points.
2. The slope or first derivative of $f(x)$ is *increasing* or at least *nondecreasing* as x increases.
3. The second derivative of $f(x)$ is always *nonnegative* for all x in the interval.
4. The linear approximation of $f(x)$ at any point in the interval always *underestimates* the true function value.
5. For a convex function, a local minimum is always a global minimum.

- If f and g are concave then
  - $-f$ is convex
  - $1/f$ is convex if $f > 0$
  - $1/f$ is concave if $f < 0$
  - $f + g$ is concave
  - αf is concave for every $\alpha \geq 0$
  - $\log f$ concave
  - f is continuous on the interior of its domain

- A function can be concave over one region and convex over another region.

- A linear function is both concave and convex.

How to determine a given function is convex?

- A twice differentiable function $f(x)$ of a single variable defined on the interval I is
    - a convex function iff $f''(x) \geq 0$ for all x in the interval I.
    - a concave function iff $f''(x) \leq 0$ for all x in the interval I.

Examples

Example 1:

Is $f(x) = x^2 – 2x + 2$ concave or convex on any interval?

$f‘(x) = 2x - 2$

$f“(x) = 2$;

$f“(x) > 0$, so, $f\,x)$ is convex for all values of x.

Example 2:

Is $f(x) = x^3 - x^2$ concave or convex on any interval?

$f‘(x) = 3x^2 - 2x$; $f“(x) = 6x – 2$

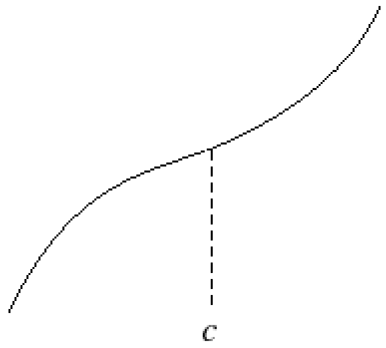
When $6x – 2 = 0$, $6x = 2$; $x = 1/3$; So, $f(x)$ is concave in the interval $(-\infty, 1/3)$

When $6x – 2 > 0$, $6x > 2$; $x > 1/3$; So, $f(x)$ is convex in the interval $(1/3, \infty)$

Exercise Problems

1. Determine the concavity/convexity of $f(x) = -(1/3)x^2 + 8x - 3$.

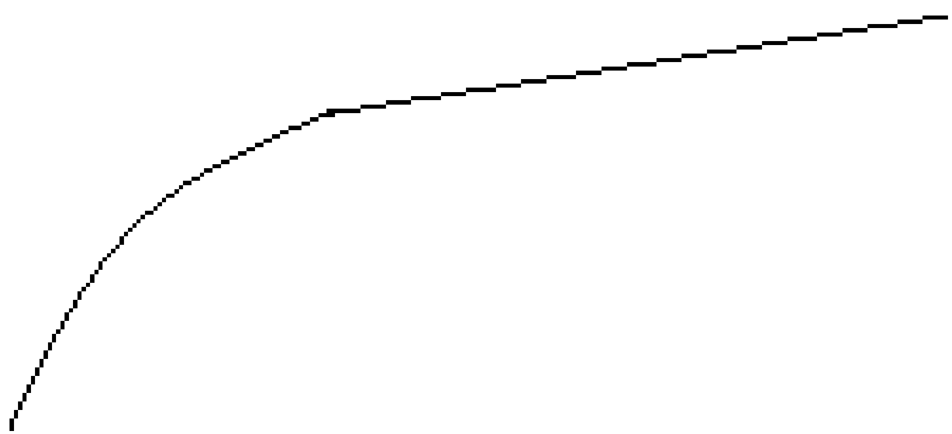
2. A competitive firm receives the price $p > 0$ for each unit of its output, and pays the price $w > 0$ for each unit of its single input. Its output from using x units of the variable input is $f(x) = x^{1/4}$. Is this production function concave? Is the firm's profit concave in x?



A point at which a twice-differentiable function changes from being convex to concave, or vice versa, is an inflection point.
An example of an inflection point is shown in the above figure.

Strict convexity and concavity

- The inequalities in the definition of concave and convex functions are weak: such functions may have linear parts, as in the following figure.



a concave, but not strictly concave, function

- A concave function that has **no** linear parts is said to be *strictly concave*.
- If $f''(x) < 0$ for all $x \in (a,b)$, then f is strictly concave on $[a, b]$, *but the converse is not true*:
- If f is strictly concave, then its second derivative is **not** necessarily negative at all points.
- Consider the function $f(x) = -x^4$. It is concave, but its second derivative at 0 is zero, *not negative*.
- That is, f is strictly concave on $[a, b]$ if $f''(x) < 0$ for all $x \in (a, b)$, **but** if f is strictly concave on $[a, b]$ then $f''(x)$ is *not* necessarily negative for all $x \in (a, b)$.

Convexity of a Multivariable Function

- The Hessian matrix of a function $f(x_1, x_2, \ldots, x_n)$ is an n x n symmetric matrix given by

$y = f(x_1, x_2)$

$\frac{\partial^2 f}{\partial x_1^2}$ $\frac{\partial^2 f}{\partial x_1}$ $\frac{\partial^2 f}{\partial x_1 \partial x_1}$

$$H_f(x_1, x_2, \ldots, x_n) = \left[\frac{\delta^2 f}{\delta x_1 \delta x_2} \right] = \nabla^2 f$$

- <u>Test for convexity of a function:</u> A function f is convex if the Hessian matrix of f is positive definite or positive semi-definite for all values of $x_1, x_2, \ldots, x_n$.
- <u>Test for concavity of a function:</u> A function f is concave if the Hessian matrix of f is negative definite or negative semi-definite for all values of $x_1, x_2, \ldots, x_n$.

$f(x_1, x_2)$

Matrix of Partial derivatives

$$H = \begin{array}{c c} & \begin{array}{cc} x_1 & \qquad x_2 \end{array} \\ \begin{array}{c} x_1 \\[2em] x_2 \end{array} & \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\[1em] \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{pmatrix} \end{array} = \underline{\underline{\nabla^2 f(x)}}$$

$$\nabla f(x) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\[1em] \dfrac{\partial f}{\partial x_2} \end{pmatrix}$$ Gradient vector

Example

$$f(x_1, x_2, x_3) = 3x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 - 2x_1x_3 + 2x_2x_3 - 6x_1 - 4x_2 - 2x_3$$

$6x_1 \; -2x_2 \; -2x_3 \; -6$

$$\nabla f(x_1, x_2, x_3) = \begin{pmatrix} 6x_1 - 2x_2 - 2x_3 - 6 \\ 4x_2 - 2x_1 + 2x_3 - 4 \\ 2x_3 - 2x_1 + 2x_2 - 2 \end{pmatrix}$$

- H is a symmetric.
- All diagonal elements are positive.

$$H_f(x_1, x_2, x_3) = \begin{bmatrix} 6 & -2 & -2 \\ -2 & 4 & 2 \\ -2 & 2 & 2 \end{bmatrix}$$

- The leading principal determinants are

$$|6| > 0 \qquad \begin{vmatrix} 6 & -2 \\ -2 & 4 \end{vmatrix} = 20 > 0$$

$$|H_f| = 16 > 0$$

- H is a positive-definite matrix , which implies *f* is a convex function

$$\begin{array}{c|ccc} & x_1 & x_2 & x_3 \\ x_1 & \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ x_2 & \frac{\partial^2 f}{\partial x_2 \partial x_1} & & \\ x_3 & [illegible] & & \end{array}$$

Optimality Criteria

- In considering optimization problems, two questions generally must be addressed:

1. **Static Question.** How can one determine whether a given point x^* is the optimal solution?
2. **Dynamic Question.** If x^* is not the optimal point, then how does one go about finding a solution that is optimal?

Optimality Criteria

- Local and global optimum

A function $f(x)$ defined on a set S attains its *global minimum* at a point $x^{**} \in S$ if and only if

$$f(x^{**}) \leqslant f(x) \qquad \text{for all } x \in S$$

A function $f(x)$ defined on S has a *local minimum* (*relative minimum*) at a point $x^* \in S$ if and only if

$$f(x^*) \leqslant f(x) \qquad \text{for all } x \text{ within a distance } \varepsilon \text{ from } x^*$$

that is, there exists an $\varepsilon > 0$ such that, for all x satisfying $|x - x^*| < \varepsilon$, $f(x^*) \leqslant f(x)$.

Stationary Point and Inflection Point

- A stationary point is a point *x** at which

$$\left.\frac{df}{dx}\right|_{x=x^*} = 0$$

- An *inflection point* or *saddle-point* is a stationary point that does not correspond to a local optimum (minimum or maximum).
- To distinguish whether a stationary point is a local minimum, a local maximum, or an inflection point, we need the *sufficient* conditions of optimality.

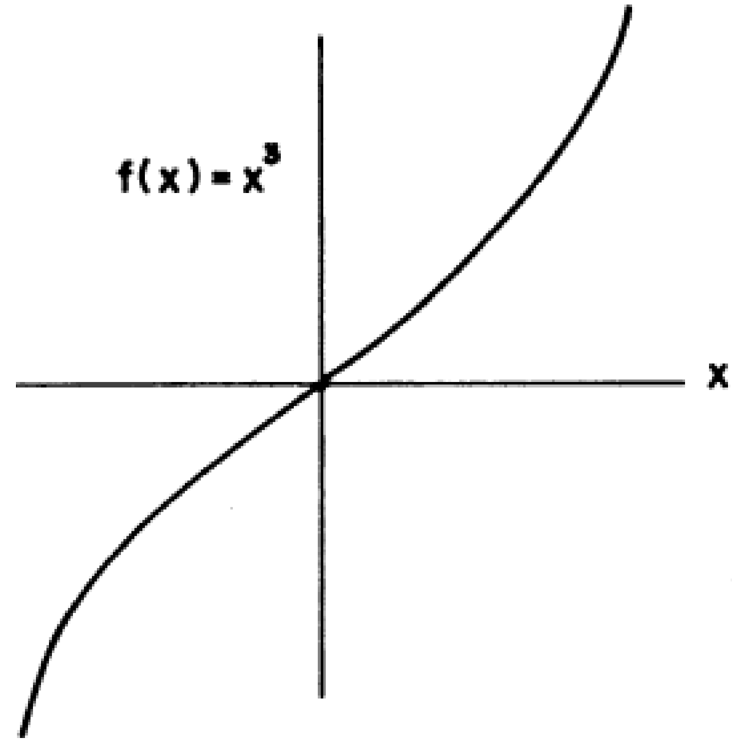
Theorem

- Suppose at a point x^* the first derivative is zero and the first nonzero higher order derivative is denoted by n.
  - If n is odd, then x^* is a point of inflection.
  - If n is even, then x^* is a local optimum.
- Moreover:
  - If that derivative is positive, then the point x^* is a local minimum.
  - If that derivative is negative, then the point x^* is a local maximum.

An Example

$$f(x) = x^3$$

$$\left.\frac{df}{dx}\right|_{x=0} = 0 \qquad \left.\frac{d^2f}{dx^2}\right|_{x=0} = 0 \qquad \left.\frac{d^3f}{dx^3}\right|_{x=0} = 6$$



- Thus the first non-vanishing derivative is 3 (odd), and $x = 0$ is an inflection point.

An Example

$$f(x) = 5x^6 - 36x^5 + \frac{165}{2}x^4 - 60x^3 + 36$$

$$\frac{df}{dx} = 30x^5 - 180x^4 + 330x^3 - 180x^2 = 30x^2(x-1)(x-2)(x-3)$$

Stationary points are $x = 0, 1, 2, 3$

$$\frac{d^2f}{dx^2} = 150x^4 - 720x^3 + 990x^2 - 360x$$

x	$f(x)$	d^2f/dx^2	
0	36	0	- Inflection point
1	27.5	60	-Local minimum
2	44	-120	-Local maximum
3	5.5	540	-Local minimum

At $x = 0$, $\dfrac{d^3f}{dx^3} = 600x^3 - 2160x^2 + 1980x - 360 = -360$

Example

Determine the maximum and minimum values of the function:

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

Solution:

$f'(x) = 60\, x^4 - 180\, x^3 + 120\, x^2$

$\quad = 60\,(x^4 - 3\, x^3 + 2\, x^2) = 60\, x^2\,(x - 1)\,(x - 2)$

$f'(x) = 0$ at $x = 0$, $x = 1$, and $x = 2$.

The second derivative is: $f''(x) = 60(4x^3 - 9x^2 + 4x)$

At $x = 1$, $f''(x) = -60$ and hence $x = 1$ is a relative maximum. Therefore,

$$f_{max} = f\,(x = 1) = 12$$

At $x = 2$, $f''(x) = 240$ and hence $x = 2$ is a relative minimum. Therefore,

$$f_{min} = f\,(x = 2) = -11$$

At x = 0, $f''(x)$ = 0 and hence we must investigate the next derivative.

$$f'''(x) = 60(12x^2 - 18x + 4) = 240 \text{ at } x = 0$$

Since $f'''(x) = 0$ at $x = 0$,

x = 0 is neither a maximum nor a minimum, and it is an inflection point.

Multivariable optimization with no constraints

- **Necessary condition**

If $f(\mathbf{X})$ has an extreme point (maximum or minimum) at $\mathbf{X}=\mathbf{X}^*$ and if the first partial derivatives of $f(\mathbf{X})$ exist at $\mathbf{X}^*$, then

$$\frac{\partial f}{\partial x_1}(\mathbf{X}^*) = \frac{\partial f}{\partial x_2}(\mathbf{X}^*) = \cdots = \frac{\partial f}{\partial x_n}(\mathbf{X}^*) = 0$$

- **Sufficient condition**

A sufficient condition for a stationary point **X*** to be an extreme point is that the matrix of second partial derivatives (Hessian matrix) of $f(\mathbf{X}^*)$ evaluated at **X*** is

Positive definite when **X*** is a **relative minimum point**

Negative definite when **X*** is a **relative maximum point**

Saddle point

- In the case of a function of two variables $f\,(x,\ y)$, the Hessian matrix may be neither positive nor negative definite at a point (x^*,y^*) at which $$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$$

  In such a case, the point (x^*,y^*) is called a saddle point.
- The characteristic of a saddle point is that it corresponds to a relative minimum or a relative maximum of $f\,(x,y)$ with respect to one variable, say, x (the other variable being fixed at $y=y^*$) and a relative maximum or a relative minimum of $f\,(x,y)$ with respect to the second variable y (the other variable being fixed at x*).

Saddle point

Example: Consider the function $f\,(x,y)=x^2 - y^2$

For this function:

$$\frac{\partial f}{\partial x} = 2x \text{ and } \frac{\partial f}{\partial y} = -2y$$

These first derivatives are zero at $x^* = 0$ and $y^* = 0$.

The Hessian matrix of f at (x^*,y^*) is given by:

$$\mathbf{H} = \begin{vmatrix} 2 & 0 \\ 0 & -2 \end{vmatrix}$$

Since this matrix is neither positive definite nor negative definite,

the point ($x^*=0$, $y^*=0$) is a saddle point.

Saddle point

It can be seen from the figure that $f(x, y^*) = f(x, 0)$

has a relative minimum and $f(x^*, y) = f(0, y)$

has a relative maximum at the saddle point (x^*, y^*).

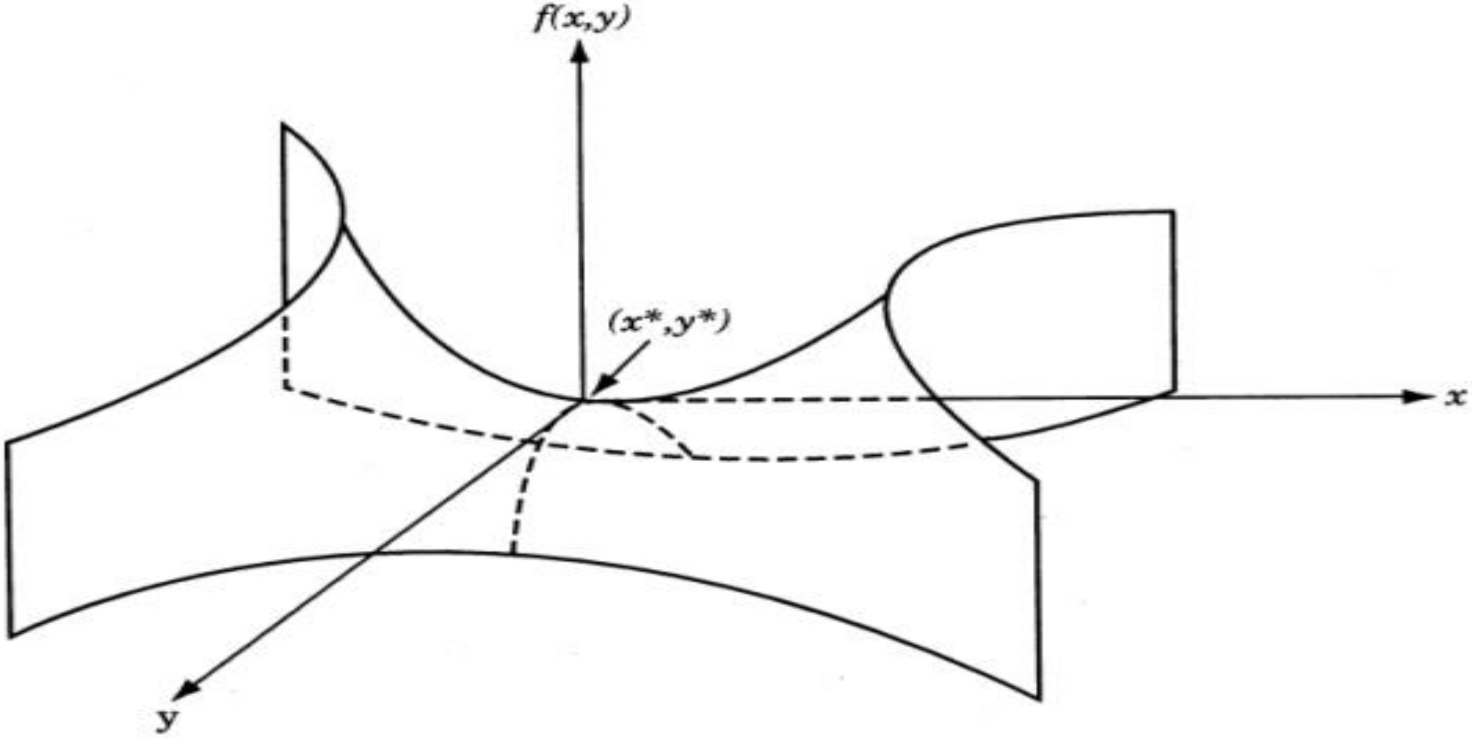


Figure 2.5 Saddle point of the function $f(x,y) = x^2 - y^2$.

Example

Find the extreme points of the function

$$f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$$

Solution: The necessary conditions for the existence of an extreme point are:

$$\frac{\partial f}{\partial x_1} = 3x_1^2 + 4x_1 = x_1(3x_1 + 4) = 0$$

$$\frac{\partial f}{\partial x_2} = 3x_2^2 + 8x_2 = x_2(3x_2 + 8) = 0$$

These equations are satisfied at the points: (0,0), (0,-8/3), (-4/3,0), and (-4/3,-8/3)

To find the nature of these extreme points, we have to use the sufficiency conditions. The second order partial derivatives of f are given by:

$$\frac{\partial^2 f}{\partial x_1^{\;2}} = 6x_1 + 4 \qquad \frac{\partial^2 f}{\partial x_2^{\;2}} = 6x_2 + 8 \qquad \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0$$

The Hessian matrix of f is given by:

$$\mathbf{H} = \begin{bmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{bmatrix}$$

$$\mathbf{H} = \begin{bmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{bmatrix}$$

$H_1 = |6x_1 + 4|$ and $\mathbf{H_2} = \begin{vmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{vmatrix}$

the values of H_1 and H_2 and

the nature of the extreme point are as given in the next slide:

Example

Point $\mathbf{X}$	Value of $\mathbf{J}_1$	Value of $\mathbf{J}_2$	Nature of $\mathbf{J}$	Nature of $\mathbf{X}$	$f(\mathbf{X})$
(0,0)	+4	+32	Positive definite	Relative minimum	6
(0,-8/3)	+4	-32	Indefinite	Saddle point	418/27
(-4/3,0)	-4	-32	Indefinite	Saddle point	194/27
(-4/3,-8/3)	-4	+32	Negative definite	Relative maximum	50/3

Find all local maxima, local minima and saddle points for

$f(x_1, x_2) = x_1^2 x_2 + x_2^3 x_1 - x_1 x_2$

The necessary conditions for the maximum of *f* give:

$$\frac{\partial f}{\partial x_1} = 2x_1 x_2 + x_2{}^3 - x_2 = 0$$

$$\frac{\partial f}{\partial x_2} = x_1{}^2 + 3\, x_1 x_2{}^2 - x_1 = 0$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = 2x_1 + 3\, x_2{}^2 - 1 = 0$$

$$\frac{\partial^2 f}{\partial x_1^2} = 2\, x_2 = 0$$

$$\frac{\partial^2 f}{\partial x_2^2} = 6\, x_1 x_2 = 0$$

$$\frac{\partial f}{\partial x_1} = 2x_1x_2 + x_2^{\ 3} - x_2 = 0$$

$2x_1\, x_2 + x_2^{\ 3} - x_2 = 0$

i.e., $x_2\,(2\, x_1 + x_2^{\ 2} - 1) = 0$

i.e., $x_2 = 0$ ... (Eq. 1)

or $2\, x_1 + x_2^{\ 2} - 1 = 0$ ... (Eq. 2)

$$\frac{\partial f}{\partial x_2} = x_1^{\ 2} + 3\, x_1x_2^{\ 2} - x_1 = 0$$

$x_1^2 + 3\, x_1\, x_2^2 - x_1 = 0$

i.e., $x_1\,(x_1 + 3\, x_2^{\ 2} - 1) = 0$

i.e., $x_1 = 0$ ... (Eq. 3)

or $x_1 + 3\, x_2^{\ 2} - 1 = 0$ ... (Eq. 4)

- For (x_1, x_2) to be a stationary point,

Eq. (1) and Eq. (3),
Eq. (1) and Eq. (4),
Eq. (2) and Eq. (3),
Eq. (2) and Eq. (4) must hold.

Eq. (1) and Eq. (3) hold at $x_2 = 0$ and $x_1 = 0$.
Eq. (1) and Eq. (4) hold at $x_2 = 0$ and $x_1 = 1$.
Eq. (2) and Eq. (3) hold at $x_1 = 0, x_2 = 1$ and
$x_1 = 0, x_2 = -1$

- Eq. (2) and Eq. (4) hold when

$2\,x_1 + x_2^{\,2} - 1 = 0$ .... (A)

$x_1 + 3\,x_2^{\,2} - 1 = 0$ .... (B)

From (A), $x_2^{\,2} = 1 - 2\,x_1$ Substitute this in (B), we get,

$x_1 + 3\,(1 - 2\,x_1) - 1 = 0$

$-\,5\,x_1 = -\,2$ Hence, $x_1 = 2/5$.

Substituting this in (A), we get, $x_2^{\,2} = 1 - 2\,(2/5)$

$x_2^{\,2} = 1/5$. Thus, $x_2 = +\sqrt{\frac{1}{5}}$, $x_2 = -\sqrt{\frac{1}{5}}$

Thus, $f(x_1, x_2)$ has the following stationary points:

Summary of the Stationary Points

X	Value of Leading Principal Minor of Order 1	Value Leading Principal Minor of Order 2	Nature of Hessian matrix	Nature of X
(0, 0)	0	-1	Indefinite	Saddle point
(1, 0)	0	-1	Indefinite	Saddle point
(0, 1)	2	-4	Indefinite	Saddle point
(0, -1)	-2	-4	Indefinite	Saddle point
(2/5, 1/√5)	2/√5	20/25	Positive Definite	Minimum Point
(2/5, -1/√5)	- 2/√5	20/25	Negative Definite	Maximum Point