# Operations Research III: Theory

# Convex Analysis

Ling-Chieh Kung

Department of Information Management
National Taiwan University

# Road map

- **Motivating examples**.
- Convex analysis.
- Single-variate NLPs.
- Multi-variate convex analysis.

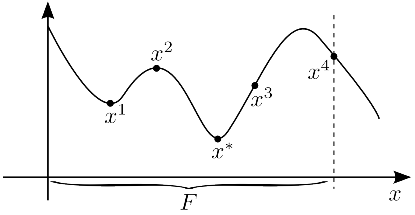
# Difficulties of NLP

- Compared with LP, NLP is much more **difficult**.

**Observation 1**

*In an NLP, a local minimum* **is not always** *a global minimum.*

- Over the feasible region $F$, $x^1$ is a local minimum but not a global minimum. How about other points?




- A greedy search may be trapped at a local minimum.

# Difficulties of NLP

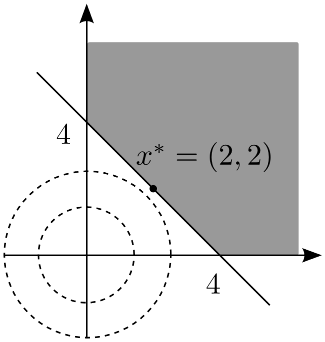
**Observation 2**

*In an NLP which has an optimal solution, there may* **exist no** *extreme point optimal solution.*

- For example:

$$\begin{array}{rl} \min\limits_{x_1 \geq 0, x_2 \geq 0} & x_1^2 + x_2^2 \\ \text{s.t.} & x_1 + x_2 \geq 4. \end{array}$$

- The optimal solution $x^* = (2, 2)$ is not an extreme point.
- The two extreme points are not optimal.



# Difficulties of NLP

- **No one** has invented an efficient algorithm for solving general NLPs (i.e., finding a global optimum).
- For an NLP:
  - We want to have a condition that makes a local minimum always a global minimum.
  - We want to have a condition that guarantees an extreme point optimal solution (when there is an optimal solution).
- To answer these questions, we need **convex analysis**.
  - Let's define convex sets and convex and concave functions.
  - Then we define convex programs and show that they have the first desired property.

# Road map

- Motivating examples.
- **Convex analysis**.
- Single-variate NLPs.
- Multi-variate convex analysis.

# Convex sets

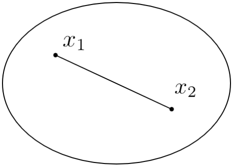
- Let's start by defining **convex sets** and **convex functions**:

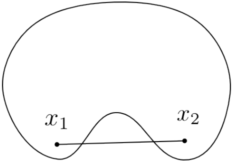
**Definition 1 (Convex sets)**

*A set $F \subseteq \mathbb{R}^n$ is convex if*

$$\lambda x_1 + (1-\lambda)x_2 \in F$$

*for all $\lambda \in [0,1]$ and $x_1, x_2 \in F$.*



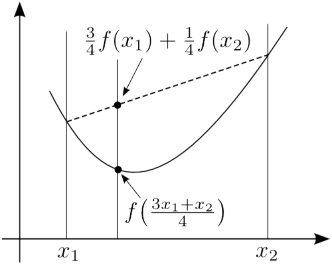
# Convex functions

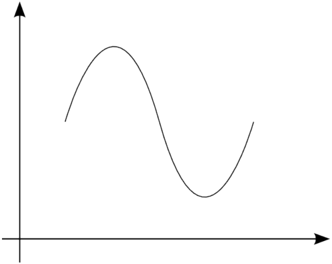
### Definition 2 (Convex functions)

*For a convex domain $F \subseteq \mathbb{R}^n$, a function $f : \mathbb{R}^n \to \mathbb{R}$ is convex over $F$ if*

$$f\Big(\lambda x_1 + (1-\lambda)x_2\Big) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

*for all $\lambda \in [0,1]$ and $x_1, x_2 \in F$.*



# Concave functions and some examples

**Definition 3 (Concave functions)**

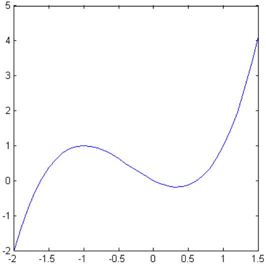
*For a convex domain $F \in \mathbb{R}^n$, a function $f : \mathbb{R}^n \to \mathbb{R}$ is* **concave** *over $F$ if $-f$ is convex.*

- Convex sets?
  - $X_1 = [10, 20]$.
  - $X_2 = (10, 20)$.
  - $X_3 = \mathbb{N}$.
  - $X_4 = \mathbb{R}$.
  - $X_5 = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq 4\}$.
  - $X_6 = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \geq 4\}$.
- Convex functions?
  - $f_1(x) = x + 2, x \in \mathbb{R}$.
  - $f_2(x) = x^2 + 2, x \in \mathbb{R}$.
  - $f_3(x) = \sin x, x \in [0, 2\pi]$.
  - $f_4(x) = \sin x, x \in [\pi, 2\pi]$.
  - $f_5(x) = \log x, x \in (0, \infty)$.
  - $f_6(x, y) = x^2 + y^2, (x, y) \in \mathbb{R}^2$.

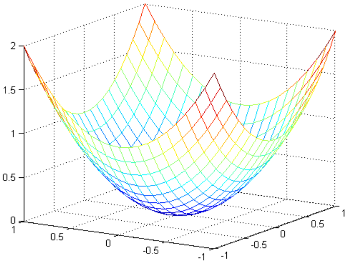
# Local vs. global optima

**Proposition 1 (Global optimality of convex functions)**

*For a convex (concave) function $f$ over a convex domain $F$, a local minimum (maximum) is a global minimum (maximum).*




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

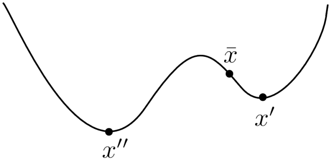



$f(x, y) = x^2 + y^2.$

# Local vs. global optima

*Proof.* Suppose a local minimum $x'$ is not a global minimum and there exists $x''$ such that $f(x'') < f(x')$. Consider a small enough $\lambda > 0$ such that $\bar{x} = \lambda x'' + (1 - \lambda)x'$ satisfies $f(\bar{x}) > f(x')$. Such $\bar{x}$ exists because $x'$ is a local minimum. Now, note that

$$\begin{aligned} f(\bar{x}) &= f(\lambda x'' + (1 - \lambda)x') \\ &> f(x') \\ &= \lambda f(x') + (1 - \lambda)f(x') \\ &> \lambda f(x'') + (1 - \lambda)f(x'), \end{aligned}$$



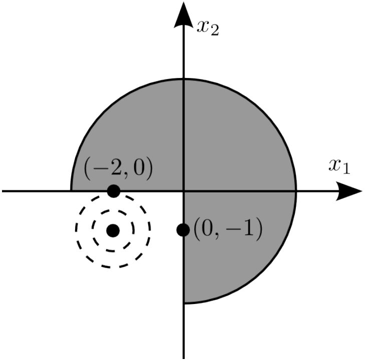

which violates the fact that $f(\cdot)$ is convex. Therefore, by contradiction, the local minimum $x'$ must be a global minimum.

# Convexity of the feasible region is required

- Consider the following example

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & (x_1 + 2)^2 + (x_2 + 1)^2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 \leq 9 \\ & x_1 \geq 0 \text{ or } x_2 \geq 0. \end{aligned}$$

  Note that the feasible region is not convex.
- The local minimum $(0, -1)$ is not a global minimum. The unique global minimum is $(-2, 0)$.



# Extreme points and optimal solutions

- Now we know if we minimize a convex function over a convex feasible region, a local minimum is a global minimum.
- What may happen if we **minimize a concave function**?
- One "goes down" on a concave function if she moves "towards its boundary".
- We thus have the following proposition:

**Proposition 2**

*For any concave function that has a global minimum over a convex feasible region, there exists a global minimum that is an extreme point.*

*Proof.* Beyond the scope of this course. □

# Special case: LP

- Now we know when we minimize $f(\cdot)$ over a convex feasible region $F$:
  - If $f(\cdot)$ is **convex**, search for a **local minimum**.
  - If $f(\cdot)$ is **concave**, search among the **extreme points** of $F$.
- For any LP, we have both!

**Proposition 3**

*The feasible region of an LP is convex.*

*Proof.* First, note that the feasible region of an LP is the intersection of several half spaces (each one is determined by an inequality constraint) and hyperplanes (each one is determined by an equality constraint). It is trivial to show that half spaces and hyperplanes are always convex. It then remains to show that the intersection of convex sets are convex, which is left as an exercise. □

# Special case: LP

## Proposition 4

*A linear function $f : \mathbb{R}^n \to \mathbb{R}$ is both convex and concave.*

*Proof.* To show that a function $f$ is convex and concave, we need to show that $f(\lambda x^1 + (1-\lambda)x^2) = \lambda f(x^1) + (1-\lambda) f(x^2)$, which is exactly the separability of linear functions: Let $f(x) = c^T x + b$ be a linear function, $c \in \mathbb{R}^n, b \in \mathbb{R}$, then

$$f\Big(\lambda x^1 + (1-\lambda)x^2\Big) = c^T \Big(\lambda x^1 + (1-\lambda)x^2\Big) + b$$
$$= \lambda(c^T x^1 + b) + (1-\lambda)(c^T x^2 + b) = \lambda f(x^1) + (1-\lambda) f(x^2).$$

Therefore, a linear function is both convex and concave. □

- To solve an LP, use a **greedy search** focusing on **extreme points**.
- This is exactly the simplex method.

# Convex Programming

- Consider a general NLP

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq b_i \quad \forall i = 1, ..., m. \end{aligned}$$

- If the feasible region $F = \{x \in \mathbb{R}^n | g_i(x) \leq b_i \forall i = 1, ..., m\}$ is convex and $f$ is convex over $F$, a local minimum is a global minimum.
- In this case, the NLP is called a **convex program** (CP).

**Definition 4 (Convex programs)**

*An NLP is a CP if its feasible region is convex and its objective function is convex over the feasible region.*

- Efficient algorithms exist for solving CPs.
- The subject of formulating and solving CPs is **Convex Programming**.

# A sufficient condition for CP

- When is an NLP a CP?

**Proposition 5**

*For an NLP*

$$\min_{x \in \mathbb{R}^n} \left\{ f(x) \middle| g_i(x) \leq b_i \forall i = 1, ..., m \right\},$$

*if $f$ and $g_i$s are all convex functions, the NLP is a CP.*

*Proof.* We only need to prove that the feasible region is convex, which is implied if $F_i = \{x \in \mathbb{R}^n | g_i(x) \leq b_i\}$ is convex for all $i$. For two points $x_1$, $x_2 \in F_i$ and an arbitrary $\lambda \in [0, 1]$, we have

$$\begin{aligned} g_i(\lambda x_1 + (1-\lambda)x_2) &\leq \lambda g_i(x_1) + (1-\lambda) g_i(x_2) \\ &\leq \lambda b_i + (1-\lambda) b_i = b_i, \end{aligned}$$

which implies that $F_i$ is convex. Repeating this argument for all $i$ completes the proof. □

# Convex programming

- We have learned some algorithms for solving nonlinear programs.
- We now should be convinced that:
  - People **can** efficiently solve CPs.
  - People **cannot** efficiently solve general NLPs.
- We will now focus on how to analytically solve NLPs.
  - Analytical solutions are the foundations for managerial insights.

# Road map

- Motivating examples.
- Convex analysis.
- **Solving single-variate NLPs**.
- Multi-variate convex analysis.

# Solving single-variate NLPs

- Here we discuss how to analytically solve single-variate NLPs.
  - "Analytically solving a problem" means to express the solution as a **function** of problem parameters **symbolically**.
- Even though solving problems with only one variable is restrictive, we will see some useful examples in the remaining semester.
- We will focus on **twice differentiable** functions and try to utilize **convexity** (if possible).

# Convexity of twice differentiable functions

- For a general function, we may need to use the definition of convex functions to show its convexity.
- For single-variate twice differentiable functions (i.e., the second-order derivative exists), there are useful properties:

**Proposition 6**

*For a twice differentiable function $f : \mathbb{R} \to \mathbb{R}$ over an interval $(a, b)$:*

- *$f$ is convex over $(a, b)$ if and only if $f''(x) \geq 0$ for all $x \in (a, b)$.*
- *$\bar{x}$ is a local minimum over $(a, b)$ only if $f'(\bar{x}) = 0$.*
- *If $f$ is convex over $(a, b)$, $x^*$ is a global minimum over $(a, b)$ if and only if $f'(x^*) = 0$.*

*Proof.* For the first two, see your Calculus textbook. The last one is a combination of the second one and the convexity of $f$. □

- Note that the two boundary points may need special considerations.

# Convexity of twice differentiable functions

- The condition $f'(x) = 0$ is called the **first order condition** (FOC).
  - For all functions, FOC is **necessary** for a local minimum.
  - For convex functions, FOC is also **sufficient** for a global minimum.
- To solve an NLP, convexity is the key.

# EOQ revisited

- Recall our EOQ (economic order quantity) problem.
- A decision maker determines the order quantity in each order in the following environment:
  - Demand is deterministic and occurs at a constant rate.
  - Regardless the order quantity, a fixed ordering cost is incurred.
  - No shortage is allowed.
  - The ordering lead time is zero.
  - The inventory holding cost is constant.

# Parameters and the decision variable

- Parameters:

$$
\begin{aligned}
D &= \text{annual demand (units)}, \\
K &= \text{unit ordering cost (\$)}, \\
h &= \text{unit holding cost per year (\$), and} \\
p &= \text{unit purchasing cost (\$)}.
\end{aligned}
$$

- Decision variable:

$$
q = \text{order quantity per order (units)}.
$$

- Objective: Minimizing annual total cost.
- For all our calculations, we will use **one year** as our time unit. Therefore, $D$ can be treated as the demand **rate**.

# Annual costs

- Annual holding cost $= h \times \frac{q}{2} = \frac{hq}{2}$.
  - For one year, the length of the time period is 1 and the inventory level is $\frac{q}{2}$ **in average**.
- Annual purchasing cost $= pD$.
  - We need to buy $D$ units regardless the order quantity $q$.
- Annual ordering cost $= K \times \frac{D}{q} = \frac{KD}{q}$.
  - The number of orders in a year is $\frac{D}{q}$.
- The NLP for optimizing the ordering decision (with the constant purchasing cost ignored) is

$$\min_{q \geq 0} \frac{KD}{q} + \frac{hq}{2}.$$

# Convexity of the EOQ model

- For

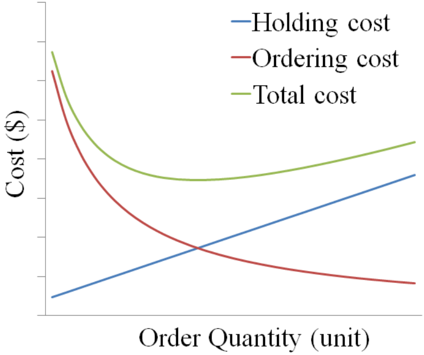
$$TC(q) = \frac{KD}{q} + \frac{hq}{2},$$

we have

$$TC'(q) = -\frac{KD}{q^2} + \frac{h}{2} \text{ and}$$

$$TC''(q) = \frac{2KD}{q^3} > 0.$$

Therefore, $TC(q)$ is convex in $q$.



# Optimizing the order quantity

- Let $q^*$ be the quantity satisfying the FOC:

$$TC'(q^*) = -\frac{KD}{(q^*)^2} + \frac{h}{2} = 0 \quad \Rightarrow \quad q^* = \sqrt{\frac{2KD}{h}}.$$

- As this quantity is feasible, it is optimal.
- The resulting annual holding and ordering cost is $TC(q^*) = \sqrt{2KDh}$.
- The optimal order quantity $q^*$ is called the **EOQ**. It is:
  - Increasing in the ordering cost $K$.
  - Increasing in the annual demand $D$.
  - Decreasing in the holding cost $h$.

  Why?

# Road map

- Motivating examples.
- Convex analysis.
- Single-variate NLPs.
- **Multi-variate convex analysis**.

# Convex analysis

- We have learned how to solve single-variate NLPs.
  - An optimal solution either satisfies the **FOC** or is a boundary point.
  - If the NLP is a **CP**, a feasible point satisfying the FOC is optimal.
- The above facts actually apply to **multi-variate NLPs**.
- We need to be able to determine whether a multi-variate function is convex, concave, or neither.
- We will still focus on **twice differentiable** functions.
  - Let's extend the notion of derivatives first.

# Partial derivatives

- For a function $f : \mathbb{R}^n \to \mathbb{R}$, its $i$th **partial derivative** is $\dfrac{\partial f(x)}{\partial x_i}$.
  - E.g., the partial derivatives for

$$f(x_1, x_2, x_3) = x_1^2 + x_2 x_3 + x_3^3$$

are

$$\frac{\partial f(x)}{\partial x_1} = 2x_1, \frac{\partial f(x)}{\partial x_2} = x_3 \text{ and } \frac{\partial f(x)}{\partial x_3} = x_2 + 3x_3^2.$$

- It also has **second-order partial derivatives**:
  - For the same $f$, we have

$$\frac{\partial^2 f(x)}{\partial x_1^2} = 2, \frac{\partial^2 f(x)}{\partial x_2^2} = 0, \frac{\partial^2 f(x)}{\partial x_3^2} = 6x_3,$$

$$\frac{\partial^2 f(x)}{\partial x_1 x_2} = \frac{\partial^2 f(x)}{\partial x_2 x_1} = 0, \frac{\partial^2 f(x)}{\partial x_1 x_3} = \frac{\partial^2 f(x)}{\partial x_3 x_1} = 0, \frac{\partial^2 f(x)}{\partial x_2 x_3} = \frac{\partial^2 f(x)}{\partial x_3 x_2} = 1.$$

# Symmetry of second-order derivatives

- For a second-order derivatives, we have the following fact:

**Proposition 7**

*For a twice differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, if its second-order derivatives are all continuous, then*

$$\frac{\partial^2 f(x)}{\partial x_i x_j} = \frac{\partial^2 f(x)}{\partial x_j x_i}$$

*for all $i = 1, ..., n$, $j = 1, ..., n$.*

- For all functions we will see in this course, the above property holds.

# Multi-variate convex functions

- For $f : \mathbb{R} \to \mathbb{R}$, $f$ is convex if and only if $f''(x) \geq 0$ for all $x$.
- For $f : \mathbb{R}^n \to \mathbb{R}$, is it true that $f$ is convex if and only if $\frac{\partial^2 f(x)}{\partial x_i^2} \geq 0$ for all $x_i$, $i = 1, ..., n$?
- Consider $f(x_1, x_2) = x_1^2 + 4x_1x_2 + x_2^2 + x_1 + x_2$. Is it convex at $(0, 0)$?
  - We have

$$\frac{\partial f(0,0)}{\partial x_1} = (2x_1 + 4x_2 + 1)\Big|_{(x_1,x_2)=(0,0)} = 1 \quad \text{and} \quad \frac{\partial^2 f(0,0)}{\partial x_1^2} = 2 > 0.$$

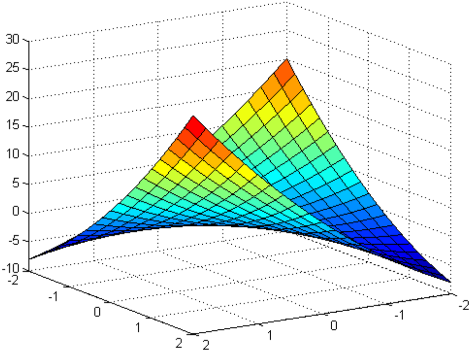
  - We also have

$$\frac{\partial f(0,0)}{\partial x_2} = (2x_2 + 4x_1 + 1)\Big|_{(x_1,x_2)=(0,0)} = 1 \quad \text{and} \quad \frac{\partial^2 f(0,0)}{\partial x_1^2} = 2 > 0$$

  - Is $f$ convex at $(0, 0)$?

# Multi-variate convex functions

- This is necessary but **insufficient**!
- $\frac{\partial^2}{\partial x_1^2} f(0,0) \geq 0$ and $\frac{\partial^2}{\partial x_2^2} f(0,0) \geq 0$ only imply that $f$ is convex **along the two axes**!
  - Along $(1,-1)$, e.g., $f$ is not convex.
- We need to test whether $f$ is convex **in all directions**.




$$f(x_1, x_2) = x_1^2 + 4x_1x_2 + x_2^2 + x_1 + x_2.$$

# Gradients and Hessians

- For a function $f : \mathbb{R}^n \to \mathbb{R}$, collecting its first- and second-order partial derivatives generates its **gradient** and **Hessian**:

**Definition 5 (Gradients and Hessians)**

*For a multi-variate twice differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, its gradient and Hessian are*

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} \quad \text{and} \quad \nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \ddots & \\ \vdots & & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}.$$

- In this course, all Hessians are **symmetric**.

# Example

- For $f(x_1, x_2, x_3) = x_1^2 + x_2 x_3 + x_3^3$, the gradient is

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \frac{\partial f(x)}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 2x_1 \\ x_3 \\ x_2 + 3x_3^2 \end{bmatrix}.$$

- The Hessian is

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f(x)}{\partial x_3 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_3 \partial x_2} & \frac{\partial^2 f(x)}{\partial x_3^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 6x_3 \end{bmatrix}.$$

- What are $\nabla f(3, 2, 1)$ and $\nabla^2 f(3, 2, 1)$?

# Convexity of twice differentiable functions

- Recall the following theorem for single-variate functions:

**Proposition 8**

*For a single-variate twice differentiable function $f(x)$:*

- *$f$ is convex in $[a, b]$ if $f''(x) \geq 0$ for all $x \in [a, b]$.*
- *$\bar{x}$ is an interior local min only if $f'(\bar{x}) = 0$.*
- *If $f$ is convex, $x^*$ is a global min if and only if $f'(x^*) = 0$.*

- We have an analogous theorem for multi-variate functions:

**Proposition 9**

*For a multi-variate twice differentiable function $f(x)$:*

- *$f$ is convex in $F$ if $\nabla^2 f(x)$ is positive semi-definite for all $x \in F$.*
- *$\bar{x}$ is an interior local min only if $\nabla f(x) = 0$.*
- *If $f$ is convex, $x^*$ is a global min if and only if $\nabla f(x^*) = 0$.*

  - What is **positive semi-definiteness** (PSD)?

# Positive semi-definite matrices

- Positive semi-definite Hessians in $\mathbb{R}^n$ are **generalizations** of nonnegative second-order derivatives in $\mathbb{R}$.

**Definition 6 (Positive semi-definite matrices)**

*A symmetric matrix $A$ is positive semi-definite if $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$.*

- Example 1: For $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, we have

$$x^T A x = 2x_1^2 + 2x_1 x_2 + 2x_2^2 = (x_1 + x_2)^2 + x_1^2 + x_2^2 \geq 0 \quad \forall x \in \mathbb{R}^2.$$

- Example 2: For $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$, we have $x^T A x = x_1^2 + 4x_1 x_2 + x_2^2$, which is negative when $x_1 = 1$ and $x_2 = -1$.

# Positive semi-definite matrices

- Given a function $f$, when is its Hessian $\nabla^2 f$ PSD?

**Proposition 10**

*For a symmetric matrix $A$, the following statements are equivalent:*

- *$A$ is positive semi-definite.*
- *$A$'s eigenvalues are all nonnegative.*
- *$A$'s principal minors are all nonnegative.*

  - $A$'s eigenvalues $\lambda$ and eigenvectors $x$ satisfy $Ax = \lambda x$.
  - $A$'s level-$k$ principal minors is the determinant of a $k \times k$ submatrix whose diagonal is part of $A$' diagonal. A sufficient condition is for $A$'s leading principal minors to be all positive, where the level-$k$ leading principal minor is the top-left $k \times k$ principal minor.
- Given a function $f$, we will (1) find its Hessian, (2) find its eigenvalues or principal minors, (3) determine over what region the Hessian is PSD. We then conclude that over that region, the function is convex.

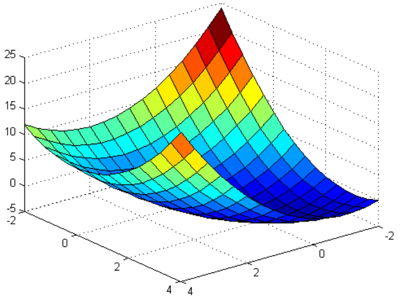
# An example

- Consider the NLP

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2),$$

where

$$f(x_1, x_2) = x_1^2 + x_2^2 + x_1 x_2 - 2x_1 - 4x_2.$$




- Its gradient and Hessian are

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 + x_2 - 2 \\ x_1 + 2x_2 - 4 \end{bmatrix} \quad \text{and} \quad \nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

# An example

$$f(x_1, x_2) = x_1^2 + x_2^2 + x_1 x_2 - 2x_1 - 4x_2.$$

- To find the eigenvalues of $\nabla^2 f(x_1, x_2)$, recall that

$$Ax = \lambda x \quad \Leftrightarrow \quad (A - \lambda I)x = 0 \quad \Leftrightarrow \quad \det(A - \lambda I) = 0.$$

- For our $\nabla^2 f(x_1, x_2)$, we have

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0 \quad \Leftrightarrow \quad 3 - 4\lambda + \lambda^2 = 0 \quad \Leftrightarrow \quad \lambda = 1 \text{ or } 3.$$

- Or by leading principal minors:

$$\begin{vmatrix} 2 \end{vmatrix} = 2 \quad \text{and} \quad \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3.$$

- So $\nabla^2 f(x_1, x_2)$ is PSD and thus $\min_{x \in \mathbb{R}^2} f(x_1, x_2)$ is a CP. The FOC requires $2x_1^* + x_2^* - 2 = 0$ and $x_1^* + 2x_2^* - 4 = 0$, i.e., $(x_1^*, x_2^*) = (0, 2)$.

# Another example

- Consider $f(x_1, x_2) = x_1^3 + 4x_1x_2 + \frac{1}{2}x_2^2 + x_1 + x_2$. When is it convex?
- Its Hessian is
$$\begin{bmatrix} 6x_1 & 4 \\ 4 & 1 \end{bmatrix}.$$
- When is the Hessian positive semi-definite?
  - We need the first leading principal minor $6x_1 \geq 0$.
  - We need the second leading principal minor $6x_1 - 16 \geq 0$.
  - We need $1 \geq 0$.
- Therefore, the function is convex if and only if $x_1 \geq \frac{8}{3}$.

# An example for solving a multi-variate NLP

- A retailer sells products 1 and 2 at prices $p_1$ and $p_2$. For product $i$, the demand is

$$q_i = a - p_i + bp_{3-i}, \quad i = 1, 2,$$

  where $a > 0$ and $b \in [0, 1)$. The retailer sets $p_1$ and $p_2$ to maximize its total profit. Assume that there is no production cost.
- Questions:
  - Explain why $b \in [0, 1)$ is reasonable.
  - Formulate the retailer's problem.
  - Is this a convex program?
  - Solve the retailer's problem.
  - How do the optimal prices change with $a$ and $b$? Does that make sense?

# An example for solving multi-variate NLP

- Suppose $b \geq 1$, it means other product's price will have the equal or even more impact on your own product's price, which is not reasonable. If $b < 0$, the higher the other product's price, the lower our demand quantity. This is possible if the two products are complementary instead of substituting. In this problem, we will focus on the case with $b \in [0, 1)$.
- $\max_{p_1, p_2} \quad p_1(a - p_1 + bp_2) + p_2(a + bp_1 - p_2)$.
- Let $f(p) = -\left[p_1(a - p_1 + bp_2) + p_2(a + bp_1 - p_2)\right]$, we have $\nabla^2 f(p) = \begin{bmatrix} 2 & -2b \\ -2b & 2 \end{bmatrix}$, which is positive semi-definite if $b \in [0, 1)$. Therefore, $f(p)$ is convex and $-f(p)$, our objective function, is concave. The problem is thus a convex program.

# An example for solving multi-variate NLP

- $\nabla f(p) = 0$ requires $-a + 2p_1 - 2bp_2 = 0$ and $-a + 2p_2 - 2bp_1 = 0$, which lead to $p_1 = p_2 = \frac{a}{2(1-b)}$.
- When $a$ increases, the two prices increase. This makes sense because one will sell a product at a higher price when the base demand becomes higher. When $b$ increases, the two prices also increase. This also makes sense because the effective demands tend to be larger.