1 Overview
This is the note for lecture 2.

2 Background on Matrix
If $A \in \mathbb{R}^{n \times n}$, then $A$ can be rewritten as $A = Q \Lambda Q^T$, where $\Lambda$ is eigenvector and $Q = [q_1, q_2, ..., q_n]$;

**Definition 1 (Positive Semidefinite)** $A \in S^n$ is positive semidefinite if for every $x \in \mathbb{R}^n$, $X^T AX \geq 0$. We can write it as $A \succeq 0$. Similarly, $A$ is positive definite if for every $x \in \mathbb{R}^n$ and $x \neq 0$, $X^T AX > 0$. We can write it as $A \succ 0$.

**Lemma 1** $A$ is positive (semi)definite iff all of its eigenvalues are nonnegative (positive).

**Proof:** if $a \succeq 0$ and $(\lambda_i, q_i)$ is an eigenvalue-eigenvector pair of $A$, then $0 \leq q_i^T A q_i = q_i^T (\sum_{i,j} \lambda_i q_i q_j^T) q_i = q_i^T (\lambda_i q_i q_i^T) q_i = \lambda_i$ (nonnegative).

Conversely, since any vector $x \in \mathbb{R}^n$ is a linear combination of $q_i$, if all $\lambda_i \geq 0$, then for $x \in \mathbb{R}^n$, we have:

$$x^T AX = [\sum_i \alpha_i q_i^T][\sum_j \lambda_j q_j q_j^T][\sum_k \alpha_k q_k] = \sum_i \alpha_i \lambda_i \alpha_i q_i^T q_i q_i^T q_i = \sum i \alpha_i^2 \lambda_i \geq 0;$$

For positive definite situation, just change $A \succeq 0$ to $A \succ 0$, the proof procedures are same.

**Lemma 2** $A \succeq 0$ iff there exists a $B \in \mathbb{R}^{m \times n}$, such that $A = B^T \bullet B$. 


3.1 Some definitions:

**Convex sets, Convex and Concave function:**

- **Convex set:** $C \in R^n$ is convex if for each $X_1, X_2 \in C$, $\lambda X_1 + (1-\lambda) X_2 \in C$ for all $0 \leq \lambda \leq 1$.

- **Convex function:** A function $f : C \rightarrow R$, where $C \in R^n$ is a convex set, is called convex function if for every $0 \leq \lambda \leq 1$ and every $X_1, X_2 \in C$, we have $f[\lambda X_1 + (1-\lambda) X_2] \leq \lambda f(X_1) + (1-\lambda) f(X_2)$.

- **Concave function:** $f$ is called concave if $-f$ is convex.

**Lemma 3** If $A$ is symmetric matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$, then:

1. $\lambda_1 = \max X^T A X$
2. $\lambda_n = \min X^T A X$

**Proof:** $X^T A X = (\sum \alpha_i q_i^T)(\sum \lambda_i q_i q_i^T)(\sum \alpha_k q_k) = \sum \alpha_i^2 \lambda_i$.

Write $y_i = \alpha_i^2$. Then it becomes a linear programming problem:

$$\begin{align*}
\text{Max:} & \quad \sum \lambda_i y_i \\
\text{s.t.} & \quad i = 1 \\
& \quad y_i \geq 0.
\end{align*} \tag{1}$$

Obviously, it has an optimal solution: $y_1 = 1, y_i = 0, \quad i = 2, \ldots, m$.

3 Convex sets, Convex and Concave function:

3.1 Some definitions:

- **Convex set:** $C \in R^n$ is convex if for each $X_1, X_2 \in C$, $\lambda X_1 + (1-\lambda) X_2 \in C$ for all $0 \leq \lambda \leq 1$.

- **Convex function:** A function $f : C \rightarrow R$, where $C \in R^n$ is a convex set, is called convex function if for every $0 \leq \lambda \leq 1$ and every $X_1, X_2 \in C$, we have $f[\lambda X_1 + (1-\lambda) X_2] \leq \lambda f(X_1) + (1-\lambda) f(X_2)$.

- **Concave function:** $f$ is called concave if $-f$ is convex.

**Lemma 4** If $f_1 : C \rightarrow R$ and $f_2 : C \rightarrow R$ for some convex set $C \in R^n$ are convex function, then $g : C \rightarrow R^n, g(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x)$ is also convex for any $\alpha_1, \alpha_2 \geq 0$.

**Proof:** $g[\lambda X_1 + (1-\lambda) X_2] = \alpha_1 f_1[\lambda X_1 + (1-\lambda) X_2] + \alpha_2 f_2[\lambda X_1 + (1-\lambda) X_2] \\
\leq \alpha_1 [\lambda f_1(X_1) + (1-\lambda) f_1(X_2)] + \alpha_2 [\lambda f_2(X_1) + (1-\lambda) f_2(X_2)] \\
= \lambda [\alpha_1 f_1(X_1) + \alpha_2 f_2(X_1)] + (1-\lambda) [\alpha_1 f_1(X_2) + \alpha_2 f_2(X_2)] \\
= \lambda g(X_1) + (1-\lambda) g(X_2).$ Therefore, $g$ is a convex function.
3.2 Epigraph of function:

Definition 2 (Epigraph of function) If any \( f : C \to R \), \( C \in R^n \), then epigraph of \( f \) is: \( \text{epi} f = \{(X,Y) : X \in R^n, Y \in R, Y \geq f(X)\} \).

Lemma 5 If a convex function \( f : C \to R \), with \( C \in R^n \), then \( f \) is a convex function iff its epigraph is a convex set.

Proof: If \( f \) is a convex, and \((x_1, y_1) \) and \((x_2, y_2) \in \text{epi} f \), then \( f(x_1) \leq y_1 \), \( f(x_2) \leq y_2 \). It means that: \( f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \leq \lambda y_1 + (1-\lambda)y_2 \).

Conversely, let \( \text{epi} f \) be convex and the points \( \{[x_1, f(x_1)], [x_2, f(x_2)]\} \in \text{epi} f \), then: \( f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \leq \lambda y_1 + (1-\lambda)y_2 \). Therefore, \( f \) is a convex function.

3.3 Level Sets:

Definition 3 (Level Sets) If \( f : C \in R^n \to R \), then \( L_c = \{X : f(x) \leq C\} \).

Lemma 6 If \( f : C \in R^n \to R \) is convex over convex set \( C \), then for every \( C, L_c \) is also convex.

Proof: If \( f \) is a convex, and \((x_1) \) and \((x_2) \in L_c \), then \( f(x_1) \leq C \), \( f(x_2) \leq C \). It means that:
\[
f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \leq C \Rightarrow [\lambda x_1 + (1-\lambda)x_2 \in L_c].
\]
Therefore, \( L_c \) is convex.

3.4 Jensen Inequality:

Assume \( \lambda_1, \lambda_2, \ldots, \lambda_k \geq 0 \), and \( \sum \lambda_i = 1 \). If \( f : C \in R^n \to R \) is convex over convex set \( C \), then for any \( x_1, x_2, \ldots, x_k \in C \), we have:
\[
f(\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \ldots + \lambda_n f(x_n).
\]

Lemma 7 Let \( C \in R^n \) be a convex set and \( f_1 : C \to R, f_2 : C \to R \), be convex functions, then the function: \( g(x) = \max\{f_1(x), f_2(x)\} \) is also convex.

Proof: Given \( x_1, x_2 \in C \) for any \( 0 \leq \lambda \leq 1 \), we have:
\[
g(\lambda x_1 + (1-\lambda)x_2) = f_i(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f_i(x_1) + (1-\lambda)f_i(x_2) \leq \lambda g(x_1) + (1-\lambda)g(x_2),
\]
where \( i = 1 \) or \( 2 \). Therefore, \( g(x) \) is convex. This lemma can be extended to \( i \) functions by induction.

Lemma 8 Assume a function \( \{f(X;Y) : X \subseteq R^n, Y \in R^n\} \), and \( C \) is convex, if for every \( y \), \( h(X,Y) \) is convex, then \( g(x) = \sup_y f(X;Y) \) is also convex.
Theorem 1 \ Let \( C \subseteq \mathbb{R}^n \) be an open convex set, if \( f \) is a convex function on \( C \), then \( f \) is continuous.

Theorem 2 \ Let \( C \subseteq \mathbb{R}^n \) be a convex set and \( f : C \to \mathbb{R}^n \) be a convex function. Consider the minimization problem to minimize \( f(X) \) on \( C \). Find \( X^* \) such that \( f(X^*) = \min_X f(X) \), and the set of \( X^* \)s, where the minimum is attained, is convex.

Proof: Let \( Z^* = f(x_1) = f(x_2) = \min f(x) \). Now consider \( \lambda x_1 + (1 - \lambda)x_2 \) for all \( 0 \leq \lambda \leq 1 \), we have: \( f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) = Z^* \). Therefore, the set of \( X^* \)s is convex. \( \blacksquare \)
Definition 2: Epigraph of function

Definition 3: Level Sets
Lemma 7: \( g(x) \) is convex

Theorem 2: Minimum set is convex