I. NOTATION

- \( \mathbb{R} \): The set of real numbers.
- \( \mathbb{R}_+ \): The set of nonnegative real numbers.
- \( \mathbb{R}_{++} \): The set of positive real numbers.
- \( \mathbb{S}_k \): The set of symmetric \( k \times k \) matrices.
- \( \mathbb{S}_k^+ \): The set of symmetric positive semi-definite \( k \times k \) matrices.
- \( \mathbb{S}_k^{++} \): The set of symmetric positive definite \( k \times k \) matrices.
- \( \text{dom} f \): The domain of the function \( f \). Let \( f : \mathbb{R}^n \to \mathbb{R}^m \), then \( \text{dom} f \triangleq \{ x \in \mathbb{R}^n : f(x) \text{ exists} \} \). For example, \( \text{dom} \log = \mathbb{R}_{++} \)

II. CONVEX OPTIMIZATION

In the previous lecture, we discussed about convex set. This lecture we will continue the discussion about convex set and start discussing about convex functions.

**Definition 1 (Supporting hyperplanes)** Suppose \( C \subseteq \mathbb{R}^n \), and \( x_0 \) is a point on the boundary of \( C \). If \( a \neq 0 \) satisfies \( a^T x \leq a^T x_0 \) for all \( x \in C \), then the hyperplane \( \{ x \in \mathbb{R}^n : a^T x = a^T x_0 \} \) is called a supporting hyperplane to \( C \) at the point \( x_0 \).

The geometric interpretation is that the hyperplane \( \{ x \in \mathbb{R}^n : a^T x = a^T x_0 \} \) is tangent to \( C \) at \( x_0 \), and the halfspace \( \{ x \in \mathbb{R}^n : a^T x \leq a^T x_0 \} \) contains \( C \). This is illustrated in Fig. 1.

**Definition 2 (Dual cones)** Let \( K \) be a cone. Then, the set

\[
K^* = \{ y : x^T y \geq 0 \text{ for all } x \in K \},
\]

is called the dual cone of \( K \).
Geometrically, \( y \in K^* \) iff \(-y\) is the normal of a hyperplane that supports \( K \) at the origin. This is illustrated in Fig. 2.

**Example 1 (Subspace)** The dual cone of a subspace \( V \subseteq \mathbb{R}^n \) is its orthogonal complement \( V^\perp = \{ y : y^T x = 0 \text{ for all } x \in V \} \).

**Example 2 (Nonnegative quadrants)** The cone \( \mathbb{R}_+ \) is its own dual

\[
y^T x \geq 0 \text{ for all } x \geq 0 \iff y \geq 0. \tag{2}
\]

We shall call such a cone *self-dual*. 
In the following, we present equivalent definitions of convex functions, and present some examples of convex functions.

**Definition 3 (Convex function)** A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if $\text{dom} f$ is a convex set and if for all $x, y \in \text{dom} f$, and $\theta$ with $0 \leq \theta \leq 1$, we have

$$f(\theta x + \bar{\theta} y) \leq \theta f(x) + \bar{\theta} f(y),$$

(3)

where $\bar{\theta} \triangleq 1 - \theta$. A function is strictly convex if strict inequality hold in (3) for $x \neq y$ and $0 < \theta < 1$. Also, we say that $f$ is concave if $-f$ is convex, and strictly concave if $-f$ is strictly convex.

Note that this definition is true also for vectors. Next, we present special criteria to verify the convexity of a function.

**Lemma 1 (Restriction of a convex function to a line)** The function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex iff the function $g : \mathbb{R} \rightarrow \mathbb{R}$,

$$g(t) = f(x + tv) \quad \text{dom} g = \{ t \in \mathbb{R} : x + tv \in \text{dom} f \},$$

(4)

is convex in $t$ for any $x \in \text{dom} f$ and $v \in \mathbb{R}^n$. Therefore, checking convexity of multivariate functions can be carried out by checking convexity of univariate functions.

**Example 3** Let $f : \mathbb{S}^n \rightarrow \mathbb{R}$ with

$$f(X) = -\log \det X, \quad \text{dom} f = \mathbb{S}^n_{++}. \quad (5)$$

Then

$$g(t) = -\log \det (X + tV)$$

$$= -\log \det X - \log \det (I + tX^{-1/2}VX^{-1/2})$$

$$= -\log \det X - \sum_{i=1}^{n} \log (1 + t\lambda_i),$$

(6)

where $I$ is the identity matrix and $\lambda_i$, $i = 1, \ldots, n$ are the eigenvalues of the matrix $X^{-1/2}VX^{-1/2}$. Since $-\sum_{i=1}^{n} \log (1 + t\lambda_i)$ is convex in $t$, as sum of convex functions (prove that the inner term of the sum is convex), and because $-\log \det X$ is constant with respect
to $t$, we have proven that for any choice of $V$ and any $X \in \text{dom} f$, $g$ is convex. Hence $f$ is also convex.

Next, we present first-order conditions which assures that a function is convex.

**Lemma 2 (First-order condition)** Let $f : \mathbb{R}^n \to \mathbb{R}$ denote a differentiable function, i.e. $\text{dom} f$ is open and for all $x \in \text{dom} f$ the gradient vector

\[
\nabla f(x) \triangleq \left[ \frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n} \right]^T,
\]

exists. Then $f$ is convex iff $\text{dom} f$ is convex and for all $x, y \in \text{dom} f$

\[
f(y) \geq f(x) + \nabla f(x)^T (y - x).
\]

**Remark 1** For $n = 1$ Lemma 2 implies that $f$ is convex iff $f(y) \geq f(x) + f'(x)(y - x)$.

Before we prove this lemma let us interpret it. The r.h.s. of (8) is the first order taylor approximation of $f(y)$ in the vicinity of $x$. According to (8), the first order taylor approximation in case where $f$ is convex, is a global underestimate of $f$. This is a very important property used in algorithm designs and performance analysis. The inequality in (8) is illustrated in Fig. 3. In the following, we prove Lemma 2 for the case of $n = 1$. Generalization is straightforward.

![Fig. 3](image-url)  

Fig. 3. If $f$ is convex and differentiable, then $f(y) \geq f(x) + \nabla f(x)^T (y - x)$ for all $x, y \in \text{dom} f$. 
**Proof:** Assume first that \( f \) is convex and \( x, y \in \text{dom} f \). Since \( \text{dom} f \) is convex, using the definition, for all \( 0 < t \leq 1 \), \( x + t (y - x) \in \text{dom} f \), and by convexity of \( f \),

\[
f (x + t (y - x)) \leq (1 - t) f (x) + tf (y).
\]

If we divide both sides by \( t \), we obtain

\[
f (y) \geq f (x) + \frac{f (x + t (y - x)) - f (x)}{t} (y - x),
\]

and taking the limit as \( t \to 0 \) yields (8).

To show sufficiency, assume the function satisfies (8) for all \( x, y \in \text{dom} f \). Choose any \( x \neq y \), and \( 0 \leq \theta \leq 1 \), and let \( z = \theta x + \theta y \). Applying (8) twice yields

\[
f (x) \geq f (z) + f' (z) (x - z), \quad f (y) \geq f (z) + f' (z) (y - z).
\]

Hence

\[
\theta f (x) + \overline{\theta} f (y) \geq \theta f (z) + \theta f' (z) (x - z) + \overline{\theta} f (z) + \overline{\theta} f' (z) (y - z)
\]

\[
= f (z) - z f' (z) + \theta x f' (z) + \overline{\theta} y f' (z)
\]

\[
= f (\theta x + \overline{\theta} y),
\]

which proves that \( f \) is convex.

Now, we present second-order conditions which assures that a function is convex.

**Theorem 1 (Second-order conditions)** Let \( f : \mathbb{R}^n \to \mathbb{R} \) denote a twice differentiable function, i.e. \( \text{dom} f \) is open and for all \( x \in \text{dom} f \) the Hessian matrix, \( \nabla^2 f (x) \in \mathbb{S}^n \),

\[
\nabla^2 f (x)_{i,j} \triangleq \frac{\partial^2 f (x)}{\partial x_i \partial x_j},
\]

exists. Then \( f \) is convex iff \( \text{dom} f \) is convex and \( \nabla^2 f (x) \succeq 0 \), for all \( x \in \text{dom} f \).

**Examples of convex/concave functions**

- Convex functions:
  - Affine: \( f (x) = ax + b \) on \( \mathbb{R} \), for any \( a, b \in \mathbb{R} \).
- Exponential: \( f(x) = \exp(ax) \) on \( \mathbb{R} \), for any \( a \in \mathbb{R} \).
- Powers: \( f(x) = x^\alpha \) on \( \mathbb{R}^+ \), for \( \alpha \geq 1 \) or \( \alpha \leq 0 \).
- Powers of absolute values: \( |x|^p \) on \( \mathbb{R} \), for \( p \geq 1 \).
- Negative entropy: \( x \log x \) on \( \mathbb{R}^+ \).
- Norms: Every norm (follows from triangle inequality).
- Max function: \( f(x) = \max\{x_1, \ldots, x_n\} \) on \( \mathbb{R}^n \).
- Quadratic-over-linear function: Let \( f : \mathbb{R}^2 \to \mathbb{R} \), such that
  \[
  f(x, y) = \frac{x^2}{y}. \tag{12}
  \]
  Then
  \[
  \nabla^2 f(x, y) = \frac{2}{y^2} \begin{bmatrix}
    y & -x \\
    -x & \frac{x^2}{y}
  \end{bmatrix}. \tag{13}
  \]
  Therefore, \( f \) is convex for any \( y > 0 \).
- Quadratic function: Let \( f : \mathbb{R}^n \to \mathbb{R} \), such that
  \[
  f(x) = \frac{1}{2} x^T P x + q^T x + r, \tag{14}
  \]
  \( q, r \in \mathbb{R}^n \) and \( P \in \mathbb{S}^n \). Since
  \[
  \nabla f(x) = Px + q, \tag{15}
  \]
  then
  \[
  \nabla^2 f(x) = P. \tag{16}
  \]
  Therefore, if \( P \succeq 0 \) then \( f \) is convex.
- Log-sum-exp: \( f(x) = \log(\exp(x_1) + \cdots + \exp(x_n)) \) on \( \mathbb{R}^n \).

- Concave functions
  - Affine: \( f(x) = ax + b \) on \( \mathbb{R} \), for any \( a, b \in \mathbb{R} \).
  - Powers: \( f(x) = x^\alpha \) on \( \mathbb{R}^+ \), for \( 0 \leq \alpha \leq 1 \).
  - Logarithm: \( \log(x) \) is concave on \( \mathbb{R} \) and defined as 0 for \( x = 0 \).
  - Log-determinant: \( f(X) = \log \det(X) \) on \( \mathbb{S}^n_+ \).

**REFERENCES**