COM524500 Optimization for Communications
Summary: Convex Sets and Convex Functions

1 Convex Sets

Affine Sets

- A set \( C \subseteq \mathbb{R}^n \) is said to be affine if
  \[
  x_1, x_2 \in C \implies \theta x_1 + (1 - \theta) x_2 \in C, \forall \theta \in \mathbb{R}
  \]
  \( 1 \)

- A point
  \[
  y = \sum_{i=1}^{k} \theta_i x_i
  \]
  \( 2 \)

where \( \theta_1 + \theta_2 + \ldots + \theta_k = 1 \), is an affine combination of the points \( x_1, \ldots, x_k \).

- An affine set can always be expressed as
  \[
  C = V + x_o
  \]
  \( 3 \)

where \( x_o \in C \), and \( V \) is a subspace.

- The affine hull of a set \( C \) (not necessarily affine) is
  \[
  \text{aff} C = \{ \theta_1 x_1 + \ldots + \theta_k x_k \mid x_1, \ldots, x_k \in C, \theta_i \in \mathbb{R}, i = 1, \ldots, k, \theta_1 + \ldots + \theta_k = 1 \}
  \]
  \( 4 \)

The affine hull is the smallest affine set that contains \( C \).

Convex Sets

- A set \( C \subseteq \mathbb{R}^n \) is said to be convex if
  \[
  x_1, x_2 \in C \implies \theta x_1 + (1 - \theta) x_2 \in C, \forall \theta \in [0,1]
  \]
  \( 5 \)

- A point
  \[
  y = \sum_{i=1}^{k} \theta_i x_i
  \]
  \( 6 \)

where \( \theta_1, \ldots, \theta_k \geq 0, \theta_1 + \theta_2 + \ldots + \theta_k = 1 \), is a convex combination of the points \( x_1, \ldots, x_k \).

- The convex hull of a set \( C \) (not necessarily convex) is
  \[
  \text{conv} C = \{ \theta_1 x_1 + \ldots + \theta_k x_k \mid x_1, \ldots, x_k \in C, \theta_i \geq 0, i = 1, \ldots, k, \theta_1 + \ldots + \theta_k = 1 \}
  \]
  \( 7 \)

The convex hull is the smallest convex set that contains \( C \).

Convex Cones

- A set \( C \subseteq \mathbb{R}^n \) is said to be a convex cone if
  \[
  x_1, x_2 \in C \implies \theta_1 x_1 + \theta_2 x_2 \in C, \forall \theta_1, \theta_2 \geq 0
  \]
  \( 8 \)
• A point

\[ y = \sum_{i=1}^{k} \theta_i x_i, \]  

where \( \theta_1, \ldots, \theta_k \geq 0 \), is a conic combination of the points \( x_1, \ldots, x_k \).

• The conic hull of a set \( C \) (not necessarily convex) is

\[ \text{conic}C = \{ \theta_1 x_1 + \ldots + \theta_k x_k \mid x_1, \ldots, x_k \in C, \theta_i \geq 0, i = 1, \ldots, k \} \]  

Some Examples of Convex Sets

• Hyperplane: \( \{ x \mid a^T x = b \} \).

• Halfspace: \( \{ x \mid a^T x \leq b \} \).

• Norm ball associated with norm \( \| \cdot \| \):

\[ B(x_c, r) = \{ x \mid \| x - x_c \| \leq r \} \]  

where \( x_c \) is the center and \( r \) is the radius. When \( \| \cdot \| \) is the 2-norm it is known as the Euclidean norm.

• Ellipsoid:

\[ E = \{ x \mid (x - x_c)^T P^{-1} (x - x_c) \leq 1 \} \]  

where \( P \succ 0 \).

• Norm cone associated with \( \| \cdot \| \):

\[ K = \{ (x, t) \mid \| x \| \leq t \} \]  

When \( \| \cdot \| \) is the 2-norm \( K \) is called the 2nd-order cone or the ice cream cone. A norm cone is not only convex but also a convex cone.

• Polyhedron:

\[ P = \{ x \mid Ax \preceq b, Cx = d \} \]  

\[ = \{ x \mid a_j^T x \leq b_j, j = 1, \ldots, m, c_j^T x = d_j, j = 1, \ldots, p \} \]  

A bounded polyhedron is called a polytope.

• Simplex: Given a set of vectors \( v_0, \ldots, v_k \) that are affine independent, a simplex is

\[ C = \text{conv} \{ v_0, \ldots, v_k \} = \{ \theta_0 v_0 + \ldots + \theta_k v_k \mid \theta \succeq 0, 1^T \theta = 0 \} \]  

A simplex is a polyhedron.

• PSD cone: \( S^n_+ = \{ X \in S^n \mid X \succeq 0 \} \) is a convex cone. (recall that \( S^n \) is the set of all real \( n \times n \) symmetric matrices.)

• The empty set \( \emptyset \) is convex. A singleton \( \{ x_0 \} \) is convex.
Convexity Preserving Operations

- Intersection:
  \[ S_1, S_2 \text{ convex} \implies S_1 \cap S_2 \text{ convex} \quad (16) \]
  \[ S_\alpha \text{ convex for every } \alpha \in \mathcal{A} \implies \bigcap_{\alpha \in \mathcal{A}} S_\alpha \text{ convex} \quad (17) \]

- Affine mapping: If \( S \subseteq \mathbb{R}^n \) is convex and \( f : \mathbb{R}^n \to \mathbb{R}^m \) is affine, then the image of \( S \) under \( f \)
  \[ f(S) = \{ f(x) \mid x \in S \} \quad (18) \]
is convex. Similarly, if \( C \subseteq \mathbb{R}^m \) is convex and \( f : \mathbb{R}^n \to \mathbb{R}^m \) is affine, then the inverse image of \( C \)
under \( f \)
  \[ f^{-1}(C) = \{ x \mid f(x) \in C \} \quad (19) \]
is convex.

- Image under perspective function: The perspective function \( P : \mathbb{R}^{n+1} \to \mathbb{R}^n \), with domain \( \text{dom}P = \mathbb{R}^n \times \mathbb{R}^+ \)
is given by
  \[ P(z, t) = z/t \quad (20) \]
If \( C \subseteq \text{dom}P \), then \( P(C) \) is convex.

Proper Cones and Generalized Inequalities

- A cone \( K \subseteq \mathbb{R}^n \) is proper if
  - \( K \) is convex
  - \( K \) is closed
  - \( K \) is solid; i.e., \( \text{int}K \neq \emptyset \)
  - \( K \) is pointed; i.e., \( x \in K, -x \in K \implies x = 0 \)

- Generalized inequality associated with a proper cone \( K \):
  \[ x \preceq_K y \iff y - x \in K \quad (21) \]
  \[ x \prec_K y \iff y - x \in \text{int}K \quad (22) \]

- Properties of generalized inequalities
  - \( x \preceq_K y, u \preceq_K v \implies x + u \preceq_K y + v \)
  - \( x \preceq_K y, y \preceq_K z \implies x \preceq_K z \)
  - \( x \preceq_K y, \alpha \geq 0 \implies \alpha x \preceq_K \alpha y \)
  - \( x \preceq_K y, y \preceq_K x \implies y = x \)

- Some examples:
  - \( K = \mathbb{R}^n_+ \). Then, \( x \preceq_K y \iff x_i \leq y_i \) for all \( i \).
  - \( K = S^n_+ \). Then, \( X \preceq_K Y \) means that \( Y - X \) is PSD.

- Minimum and minimal elements: A point \( x \in S \) is the minimum element of \( S \) if
  \[ y \in S \implies x \preceq_K y \quad (23) \]
provided that such an \( x \) exists. The minimum element, if it exists, is unique. A point \( x \in S \) is a
minimal element of \( S \) if
  \[ y \in S, y \preceq_K x \implies y = x \quad (24) \]
Separating Hyperplane Theorem
Suppose that $C, D \subseteq \mathbb{R}^n$ are convex and that $C \cap D = \emptyset$. Then there exist $a \neq 0$ and $b$ such that
\begin{align*}
a^T x \leq b & \implies x \in C \quad (25) \\
a^T x \geq b & \implies x \in D \quad (26)
\end{align*}
The hyperplane \{ $x \mid a^T x = b$ \} is called a separating hyperplane for $C$ and $D$.

Supporting Hyperplanes
Suppose that $x_o \in \text{bd}C$ (a boundary point of $C$). If there exist $a \neq 0$ such that $a^T x \leq a^T x_o$ for all $x \in C$, then \{ $x \mid a^T x = b$ \} is called a supporting hyperplane to $C$ at point $x_o$.

For any nonempty convex set $C$, and for any $x_o \in \text{bd}C$, there exists a supporting hyperplane to $C$ at $x_o$.

Dual Cones
- The dual cone of a cone $K$ is
  \begin{equation}
  K^* = \{ y \mid y^T x \geq 0 \ \forall x \in K \} \quad (27)
  \end{equation}
  A cone $K$ is called self-dual if $K = K^*$.
- If $K$ is proper then $K^*$ is also proper.
- Some examples: $\mathbb{R}_+^n$ and $S_n^+$ are self-dual. The dual cone of a norm cone $K^* = \{(x,t) \mid \|x\| \leq t \}$ is
  \begin{equation}
  K^* = \{(x,t) \mid \|x\|_* \leq t \} \quad (28)
  \end{equation}
  where \|.$\|_*$ is the dual norm of $\|.$\|. 
2 Convex Functions

Definition

- A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if $\text{dom}f$ is convex and for all $x, y \in \text{dom}f$, $0 \leq \theta \leq 1$,
  \[
f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)
  \]  \hspace{1cm} (29)

- A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is strictly convex if $\text{dom}f$ is convex and for all $x, y \in \text{dom}f$, $x \neq y$, $0 < \theta < 1$,
  \[
f(\theta x + (1 - \theta)y) < \theta f(x) + (1 - \theta)f(y)
  \]  \hspace{1cm} (30)

- A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is concave if $-f$ is convex.

Fundamental Properties

- $f$ is convex if and only if it is convex when restricted to any line that intersects its domain; i.e., for all $x \in \text{dom}f$ and $\nu$,
  \[
g(t) = f(x + t\nu)
  \]  is convex over $\{t \mid x + t\nu \in \text{dom}f\}$.

- First order condition: Suppose that $f$ is differentiable. A function $f$ with a convex domain $\text{dom}f$ is convex if and only if
  \[
f(y) \geq f(x) + \nabla f(x)^T(y - x)
  \]  \hspace{1cm} (32)
  for all $x, y \in \text{dom}f$.

- Second order condition: Suppose that $f$ is twice differentiable. A function $f$ with a convex domain $\text{dom}f$ is convex if and only if its Hessian
  \[
  \nabla^2 f(x) \succeq 0
  \]  \hspace{1cm} (33)
  for all $x \in \text{dom}f$. A function $f$ with a convex domain $\text{dom}f$ is strictly convex if
  \[
  \nabla^2 f(x) \succ 0
  \]  \hspace{1cm} (34)
  for all $x \in \text{dom}f$ (the converse is not true).

- Sublevel sets: The sublevel set of $f$ is
  \[
  C_\alpha = \{x \in \text{dom}f \mid f(x) \leq \alpha\}
  \]  \hspace{1cm} (35)
  If $f$ is convex, then $C_\alpha$ is convex for every $\alpha$ (the converse is not true).

- Epigraph: The epigraph of $f$ is
  \[
  \text{epi}f = \{(x, t) \mid x \in \text{dom}f, f(x) \leq t\}
  \]  \hspace{1cm} (36)
  $f$ is convex if and only if $\text{epi}f$ is convex.
Examples

• Examples on $\mathbb{R}$:
  - $e^{ax}$ is convex on $\mathbb{R}$
  - $\log x$ is concave on $\mathbb{R}_{++}$
  - $x \log x$ is convex on $\mathbb{R}_{++}$
  - $\log \int_{-\infty}^{\infty} e^{-t^2/2} dt$ is concave on $\mathbb{R}$

• Examples on $\mathbb{R}^n$:
  - A linear function $a^T x + b$ is convex and concave.
  - A quadratic function $x^T Px + 2q^T x + r$ is convex if and only if $P \succeq 0$, and is strictly convex if $P \succ 0$.
  - Every norm $\|x\|$ is convex.
  - $\max\{x_1, \ldots, x_n\}$ is convex.
  - The geometric mean $(\prod_{i=1}^{n} x_i)^{1/n}$ is concave on $\mathbb{R}_{++}^n$.

• Examples on $\mathbb{R}^{n \times m}$
  - $\text{tr}(AX)$ is linear on $\mathbb{R}^{n \times m}$, and hence is convex and concave.
  - The negative logarithmic determinant function $-\log \det X$ is convex on $\mathbb{S}^n_{++}$.
  - $\text{tr}(X^{-1})$ is convex on $\mathbb{S}^n_{++}$.

Jensen Inequality

• For a convex $f$,
  \[ f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y) \] (37)
  holds for any $x, y \in \text{dom} f$ and $0 \leq \theta \leq 1$.

• Extension: For a convex $f$,
  \[ f(Ez) \leq Ef(z) \] (38)
  for any random variable $z$.

• Jensen inequality can be used to derive certain inequalities; e.g., the arithmetic-geometric mean inequality:
  \[ \sqrt{ab} \leq \frac{a + b}{2}, \quad a, b \geq 0 \] (39)
  and
  \[ \left( \prod_{i=1}^{n} x_i \right)^{1/n} \leq \frac{1}{n} \sum_{i=1}^{n} x_i, \quad x_i \geq 0, i = 1, \ldots, n \] (40)
Convexity Preserving Operations

- Nonnegative weighted sums:
  \[ f_1, \ldots, f_m \text{ convex, } w_1, \ldots, w_m \geq 0 \implies \sum_{i=1}^{m} w_i f_i \text{ convex} \quad (41) \]
  \[ f(x, y) \text{ convex in } x \text{ for each } y \in A, w(y) \geq 0 \text{ for each } y \in A \implies \int_{A} w(y)f(x, y)dy \text{ convex} \quad (42) \]
  Example: \( f(x) = \sum_{i=1}^{n} x_i \log x_i \) is convex on \( \mathbb{R}^n_{++} \).

- Composition with an affine mapping:
  \[ g(x) = f(Ax + b) \quad (43) \]
  is convex if \( f \) is convex.

- Pointwise maximum and supremum:
  \[ f_1, f_2 \text{ convex} \implies g(x) = \max\{f_1(x), f_2(x)\} \text{ convex} \quad (44) \]
  \[ f(x, y) \text{ convex in } x \text{ for each } y \in A \implies g(x) = \sup_{y \in A} f(x, y) \text{ convex} \quad (45) \]

Examples:
- A piecewise linear function \( f(x) = \max_{i=1, \ldots, L} a_i^T x + b_i \) is convex.
- \( f(x) = \sup_{y \in C} \|x - y\| \) is convex for any set \( C \).
- The largest eigenvalue of \( X \)
  \[ f(X) = \lambda_{\max}(X) = \sup_{\|y\|_2 = 1} \|xy\|_2 = \sup_{\|y\|_2 = 1} \text{ tr}(Xyy^T) \quad (46) \]
  is convex on \( \mathbb{S}^n \).
- The 2-norm of \( X \)
  \[ f(X) = \|X\|_2 = \sup_{\|y\|_2 = 1} \|Xy\|_2 \quad (47) \]
  is convex on \( \mathbb{R}^{n \times m} \).

- Composition: Let \( f(x) = h(g(x)) \), where \( h : \mathbb{R} \to \mathbb{R} \) and \( g : \mathbb{R}^n \to \mathbb{R} \). Let
  \[ \tilde{h}(x) = \begin{cases} h(x), & x \in \text{dom} h \\ \infty, & \text{otherwise} \end{cases} \quad (48) \]
  Then, \( f \) is convex if \( \tilde{h} \) is convex and nondecreasing, and \( g \) is convex.
  \( f \) is convex if \( \tilde{h} \) is convex and nonincreasing, and \( g \) is concave.

- Minimization:
  \[ f(x, y) \text{ convex in } (x, y), C \text{ convex nonempty} \implies g(x) = \inf_{y \in C} f(x, y) \text{ convex} \quad (49) \]
  provided that \( g(x) > -\infty \) for some \( x \).

Examples:
\(- \operatorname{dist}(x, S) = \inf_{y \in S} \|x - y\| \) is convex for convex \( S \).

- The Schur complement

\[
\begin{bmatrix}
A & B \\
B^T & C
\end{bmatrix} \succeq 0 \iff C \succeq 0, A - BC^T B \succeq 0
\] (50)

may be proven by the convex minimization property.

- Perspective: The perspective of a function \( f : \mathbb{R}^{n+1} \rightarrow \mathbb{R} \)

\[
g(x, t) = tf(x/t), \quad \text{dom} g = \{(x, t) \mid x/t \in \text{dom} f, t > 0\} \] (51)

If \( f \) is convex then \( g \) is convex.

### Conjugate Function

\[
f^*(y) = \sup_{x \in \text{dom} f} (y^T x - f(x))
\] (52)

- \( f^* \) is convex irrespective of the convexity of \( f \).
- \( \text{dom} f^* \) consists of \( y \) for which the supremum is finite; i.e., \( \text{dom} f^* = \{y \mid f^*(y) < \infty\} \).
- Examples:
  - The conjugate function of \( f(x) = \frac{1}{2}x^T Q x, \quad Q \succ 0 \)
    is \( f^*(y) = \frac{1}{2} y^T Q^{-1} y \).
  - The conjugate function of \( f(X) = -\log \det X, \quad \text{dom} f = S^n_+ \) is \( f^*(Y) = \log \det (-Y^{-1}) - n \),
    \( \text{dom} f^* = -S^n_+ \).

### Quasiconvex Functions

- Definition:
  - A function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is quasiconvex (or unimodal) if \( \text{dom} f \) is convex and the sublevel set

\[
S_\alpha = \{x \in \text{dom} f \mid f(x) \leq \alpha\}
\] (53)

is convex for every \( \alpha \).
  - A function \( f \) is quasiconcave if \(-f\) is quasiconvex.
  - A function \( f \) is quasilinear if \( f \) is quasiconvex and quasiconcave.

- Examples:
  - \( \log x \) is quasilinear on \( \mathbb{R}_{++} \).
  - A linear fractional function

\[
f(x) = \frac{a^T x + b}{c^T x + d} \quad \text{dom} f = \{x \mid c^T x + d > 0\}
\] (54)

is quasilinear.
  - \( \text{rank} X \) is quasiconcave on \( S^n_+ \) (proven using the modified Jensen inequality).
- Modified Jensen inequality: \( f \) is quasiconvex if and only if for any \( x, y \in \text{dom} f \), and \( 0 \leq \theta \leq 1 \),

\[
f(\theta x + (1 - \theta) y) \leq \max\{f(x), f(y)\}
\] (55)
First-order condition: Suppose that $f$ is differentiable. $f$ is quasiconvex if and only if $\text{dom} f$ is convex and for all $x, y \in \text{dom} f$
\[ f(y) \leq f(x) \implies \nabla f(x)^T(y - x) \leq 0 \] (56)

Second-order condition: Suppose $f$ is differentiable. If $f$ is quasiconvex then for all $x \in \text{dom} f, y \in \mathbb{R}^n$,
\[ y^T \nabla f(x) = 0 \implies y^T \nabla^2 f(x)y \geq 0 \] (57)

Convexity with respect to Generalized Inequality

Let $K$ be a proper cone. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is $K$-convex if for all $x, y \in \text{dom} f$ and $0 \leq \theta \leq 1$,
\[ f(\theta x + (1 - \theta)y) \preceq_K \theta f(x) + (1 - \theta)f(y) \] (58)

For $K = \mathbb{R}^n_+$, a $K$-convex function is a function for which each component function $f_i$ is convex.

Consider $K = S^n_+$.
- $f(X) = X^T X$ is $K$-convex on $\mathbb{R}^{n \times m}$.
- $f(X) = X^{-1}$ is $K$-convex on $S^n_+$. 

9