ELEG5481: Signal Processing Optimization Techniques
Summary: Convex Sets and Convex Functions

1 Convex Sets

Affine Sets

- A set $C \subseteq \mathbf{R}^n$ is said to be affine if

$$x_1, x_2 \in C \quad \Longrightarrow \quad \theta x_1 + (1-\theta)x_2 \in C, \forall \theta \in \mathbf{R} \tag{1}$$

- A point

$$y = \sum_{i=1}^{k} \theta_i x_i, \tag{2}$$

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

$$C = V + x_o \tag{3}$$

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

- The affine hull of a set C (not necessarily affine) is

$$\mathbf{aff}C = \{\theta_1 x_1 + \ldots + \theta_k x_k \mid x_1, \ldots, x_k \in C,\ \theta_i \in \mathbf{R}, i = 1, \ldots, k,\ \theta_1 + \ldots + \theta_k = 1\} \tag{4}$$

The affine hull is the smallest affine set that contains C.

Convex Sets

- A set $C \subseteq \mathbf{R}^n$ is said to be convex if

$$x_1, x_2 \in C \quad \Longrightarrow \quad \theta x_1 + (1-\theta)x_2 \in C, \forall \theta \in [0,1] \tag{5}$$

- A point

$$y = \sum_{i=1}^{k} \theta_i x_i, \tag{6}$$

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

$$\mathbf{conv}C = \{\theta_1 x_1 + \ldots + \theta_k x_k \mid x_1, \ldots, x_k \in C, \theta_i \geq 0, i = 1, \ldots, k, \theta_1 + \ldots + \theta_k = 1\} \tag{7}$$

The convex hull is the smallest convex set that contains C.

Convex Cones

- A set $C \subseteq \mathbf{R}^n$ is said to be a convex cone if

$$x_1, x_2 \in C \quad \Longrightarrow \quad \theta_1 x_1 + \theta_2 x_2 \in C, \forall \theta_1, \theta_2 \geq 0 \tag{8}$$

- A point
$$y = \sum_{i=1}^{k} \theta_i x_i, \tag{9}$$
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
$$\textbf{conic} C = \{\theta_1 x_1 + \ldots + \theta_k x_k \mid x_1, \ldots, x_k \in C, \theta_i \geq 0, i = 1, \ldots, k\} \tag{10}$$

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 $\|.\|$:
$$B(x_c, r) = \{x \mid \|x - x_c\| \leq r\} \tag{11}$$
where x_c is the center and r is the radius. When $\|.\|$ is the 2-norm it is known as the Euclidean norm.
- Ellipsoid:
$$\mathcal{E} = \{x \mid (x - x_c)^T P^{-1} (x - x_c) \leq 1\} \tag{12}$$
where $P \succ 0$.
- Norm cone associated with $\|.\|$:
$$K = \{(x, t) \mid \|x\| \leq t\} \tag{13}$$
When $\|.\|$ 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:
$$\begin{aligned} \mathcal{P} &= \{x \mid Ax \preceq b,\ Cx = d\} \\ &= \{x \mid a_j^T x \leq b_j, j = 1, \ldots, m,\ c_j^T x = d_j, j = 1, \ldots, p\} \end{aligned} \tag{14}$$
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
$$C = \textbf{conv}\{v_0, \ldots, v_k\} = \{\theta_0 v_0 + \ldots + \theta_k v_k \mid \theta \succeq 0, \mathbf{1}^T \theta = 0\} \tag{15}$$
A simplex is a polyhedron.
- PSD cone: $\mathbf{S}_+^n = \{X \in \mathbf{S}^n \mid X \succeq 0\}$ is a convex cone. (recall that $\mathbf{S}^n$ is the set of all real $n \times n$ symmetric matrices.)
- The empty set $\emptyset$ is convex. A singleton $\{x_o\}$ is convex.

Convexity Preserving Operations

- Intersection:

$$S_1, S_2 \text{ convex} \quad \Longrightarrow \quad S_1 \cap S_2 \text{ convex} \tag{16}$$

$$S_\alpha \text{ convex for every } \alpha \in \mathcal{A} \quad \Longrightarrow \quad \bigcap_{\alpha \in \mathcal{A}} S_\alpha \text{ convex} \tag{17}$$

- Affine mapping: If $S \subseteq \mathbf{R}^n$ is convex and $f : \mathbf{R}^n \to \mathbf{R}^m$ is affine, then the image of S under f

$$f(S) = \{f(x) \mid x \in S\} \tag{18}$$

is convex. Similarly, if $C \subseteq \mathbf{R}^m$ is convex and $f : \mathbf{R}^n \to \mathbf{R}^m$ is affine, then the inverse image of C under f

$$f^{-1}(C) = \{x \mid f(x) \in C\} \tag{19}$$

is convex.

- Image under perspective function: The perspective function $P : \mathbf{R}^{n+1} \to \mathbf{R}^n$, with domain $\mathbf{dom} P = \mathbf{R}^n \times \mathbf{R}_{++}$ is given by

$$P(z, t) = z/t \tag{20}$$

If $C \subseteq \mathbf{dom} P$, then $P(C)$ is convex.

Proper Cones and Generalized Inequalities

- A cone $K \subseteq \mathbf{R}^n$ is proper if
  - K is convex
  - K is closed
  - K is solid; i.e., $\mathbf{int} K \neq \emptyset$
  - K is pointed; i.e., $x \in K, -x \in K \Longrightarrow x = 0$
- Generalized inequality associated with a proper cone K:

$$x \preceq_K y \Longleftrightarrow y - x \in K \tag{21}$$

$$x \prec_K y \Longleftrightarrow y - x \in \mathbf{int} K \tag{22}$$

- Properties of generalized inequalities
  - $x \preceq_K y, u \preceq_K v \Longrightarrow x + u \preceq_K y + v$
  - $x \preceq_K y, y \preceq_K z \Longrightarrow x \preceq_K z$
  - $x \preceq_K y, \alpha \geq 0 \Longrightarrow \alpha x \preceq_K \alpha y$
  - $x \preceq_K y, y \preceq_K x \Longrightarrow y = x$
- Some examples:
  - $K = \mathbf{R}^n_+$. Then, $x \preceq_K y \Longleftrightarrow x_i \leq y_i$ for all i.
  - $K = \mathbf{S}^n_+$. 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

$$y \in S \quad \Longrightarrow \quad x \preceq_K y \tag{23}$$

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

$$y \in S, y \preceq_K x \quad \Longrightarrow \quad y = x \tag{24}$$

Dual Cones

- The dual cone of a cone K is
$$K^* = \{y \mid y^T x \geq 0 \ \forall x \in K\} \tag{25}$$
A cone K is called self-dual if $K = K^*$.
- If K is proper then K^* is also proper.
- Some examples: $\mathbf{R}_+^n$ and $\mathbf{S}_+^n$ are self-dual. The dual cone of a norm cone $K^* = \{(x,t) \mid \|x\| \leq t\}$ is
$$K^* = \{(x,t) \mid \|x\|_* \leq t\} \tag{26}$$
where $\|.\|_*$ is the dual norm of $\|.\|$.

2 Convex Functions

Definition

- A function $f : \mathbf{R}^n \to \mathbf{R}$ is convex if $\mathbf{dom} f$ is convex and for all $x, y \in \mathbf{dom} f$, $0 \le \theta \le 1$,

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y) \tag{27}$$

- A function $f : \mathbf{R}^n \to \mathbf{R}$ is strictly convex if $\mathbf{dom} f$ is convex and for all $x, y \in \mathbf{dom} f$, $x \neq y$, $0 < \theta < 1$,

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y) \tag{28}$$

- A function $f : \mathbf{R}^n \to \mathbf{R}$ 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 $x \in \mathbf{dom} f$ and ν,

$$g(t) = f(x + t\nu) \tag{29}$$

  is convex over $\{t \mid x + t\nu \in \mathbf{dom} f\}$.

- First order condition: Suppose that f is differentiable. A function f with a convex domain $\mathbf{dom} f$ is convex if and only if

$$f(y) \ge f(x) + \nabla f(x)^T (y - x) \tag{30}$$

  for all $x, y \in \mathbf{dom} f$.

- Second order condition: Suppose that f is twice differentiable. A function f with a convex domain $\mathbf{dom} f$ is convex if and only if its Hessian

$$\nabla^2 f(x) \succeq 0 \tag{31}$$

  for all $x \in \mathbf{dom} f$. A function f with a convex domain $\mathbf{dom} f$ is stricly convex if

$$\nabla^2 f(x) \succ 0 \tag{32}$$

  for all $x \in \mathbf{dom} f$ (the converse is not true).

- Sublevel sets: The sublevel set of f is

$$C_\alpha = \{x \in \mathbf{dom} f \mid f(x) \le \alpha\} \tag{33}$$

  If f is convex, then C_α is convex for every α (the converse is not true).

- Epigraph: The epigraph of f is

$$\mathbf{epi} f = \{(x, t) \mid x \in \mathbf{dom} f, f(x) \le t\} \tag{34}$$

  f is convex if and only if $\mathbf{epi} f$ is convex.

Examples

- Examples on $\mathbf{R}$:
  - e^{ax} is convex on $\mathbf{R}$
  - $\log x$ is concave on $\mathbf{R}_{++}$
  - $x \log x$ is convex on $\mathbf{R}_{++}$
  - $\log \int_{-\infty}^{x} e^{-t^2/2} dt$ is concave on $\mathbf{R}$
- Examples on $\mathbf{R}^n$:
  - A linear function $a^T x + b$ is convex and concave.
  - A quadratic function $x^T P x + 2q^T x + r$ 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 $(\prod_{i=1}^{n} x_i)^{\frac{1}{n}}$ is concave on $\mathbf{R}_{++}^n$.
- Examples on $\mathbf{R}^{n \times m}$
  - $\mathrm{tr}(AX)$ is linear on $\mathbf{R}^{n \times m}$, and hence is convex and concave.
  - The negative logarithmetic determinant function $-\log\det X$ is convex on $\mathbf{S}_{++}^n$.
  - $\mathrm{tr}(X^{-1})$ is convex on $\mathbf{S}_{++}^n$.

Jensen Inequality

- For a convex f,
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y) \tag{35}$$
holds for any $x, y \in \mathbf{dom} f$ and $0 \leq \theta \leq 1$.
- Extension: For a convex f,
$$f(\mathbf{E}z) \leq \mathbf{E} f(z) \tag{36}$$
for any random variable z.
- Jensen inequality can be used to derive certain inequalities; e.g., the arithmetic-geometric mean inequality:
$$\sqrt{ab} \leq \frac{a+b}{2}, \qquad a, b \geq 0 \tag{37}$$
and
$$\left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad x_i \geq 0, i = 1, \ldots, n \tag{38}$$

Convexity Preserving Operations

- Nonnegative weighted sums:

$$f_1, \ldots, f_m \text{ convex}, w_1, \ldots, w_m \geq 0 \implies \sum_{i=1}^{m} w_i f_i \text{ convex} \quad (39)$$

$$f(x,y) \text{ convex in } x \text{ for each } y \in \mathcal{A}, w(y) \geq 0 \text{ for each } y \in \mathcal{A} \implies \int_{\mathcal{A}} w(y) f(x,y) dy \text{ convex} \quad (40)$$

  Example: $f(x) = \sum_{i=1} x_i \log x_i$ is convex on $\mathbf{R}^n_{++}$.

- Composition with an affine mapping:

$$g(x) = f(Ax + b) \quad (41)$$

  is convex if f is convex.

- Pointwise maximum and supremum:

$$f_1, f_2 \text{ convex} \implies g(x) = \max\{f_1(x), f_2(x)\} \text{ convex} \quad (42)$$

$$f(x,y) \text{ convex in } x \; for\ each\ y \in \mathcal{A} \implies g(x) = \sup_{y \in \mathcal{A}} f(x,y) \text{ convex} \quad (43)$$

  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

$$\begin{aligned} f(X) &= \lambda_{\max}(X) \\ &= \sup_{\|y\|_2 = 1} y^T X y = \sup_{\|y\|_2 = 1} \operatorname{tr}(X y y^T) \end{aligned} \quad (44)$$

    is convex on $\mathbf{S}^n$.
  - The 2-norm of X

$$\begin{aligned} f(X) &= \|X\|_2 \\ &= \sup_{\|y\|_2 = 1} \|Xy\|_2 \end{aligned} \quad (45)$$

    is convex on $\mathbf{R}^{n \times m}$.

- Composition: Let $f(x) = h(g(x))$, where $h : \mathbf{R} \to \mathbf{R}$, and $g : \mathbf{R}^n \to \mathbf{R}$. Let

$$\tilde{h}(x) = \begin{cases} h(x), & x \in \mathbf{dom} h \\ \infty, & \text{otherwise} \end{cases} \quad (46)$$

  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:

$$f(x,y) \text{ convex in } (x,y), C \text{ convex nonempty} \implies g(x) = \inf_{y \in C} f(x,y) \text{ convex} \quad (47)$$

  provided that $g(x) > -\infty$ for some x.
  Examples:

- $\text{dist}(x, S) = \inf_{y \in S} \|x - y\|$ is convex for convex S.
- The Schur complement

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0 \iff C \succeq 0, A - BC^\dagger B^T \succeq 0 \tag{48}$$

  may be proven by the convex minimization property.

- Perspective: The perspective of a function f is a function $g : \mathbf{R}^{n+1} \to \mathbf{R}$

$$g(x, t) = tf(x/t), \qquad \mathbf{dom} g = \{(x, t) \mid x/t \in \mathbf{dom} f, t > 0\} \tag{49}$$

  If f is convex then g is convex.

Quasiconvex Functions

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

$$S_\alpha = \{x \in \mathbf{dom} f \mid f(x) \leq \alpha\} \tag{50}$$

    is convex for every α.
  - 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 $\mathbf{R}_{++}$.
  - A linear fractional function

$$f(x) = \frac{a^T x + b}{c^T x + d}, \qquad \mathbf{dom} f = \{x \mid c^T x + d > 0\} \tag{51}$$

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

$$f(\theta x + (1 - \theta) y) \leq \max\{f(x), f(y)\} \tag{52}$$

- First-order condition: Suppose that f is differentiable. f is quasiconvex if and only if $\mathbf{dom} f$ is convex and for all $x, y \in \mathbf{dom} f$

$$f(y) \leq f(x) \implies \nabla f(x)^T (y - x) \leq 0 \tag{53}$$

- Second-order condition: Suppose f is differentiable. If f is quasiconvex then for all $x \in \mathbf{dom} f$, $y \in \mathbf{R}^n$,

$$y^T \nabla f(x) = 0 \implies y^T \nabla^2 f(x) y \geq 0 \tag{54}$$

Convexity with respect to Generalized Inequality

- Let K be a proper cone. A function $f : \mathbf{R}^n \to \mathbf{R}^m$ is K-convex if for all $x, y \in \mathbf{dom} f$ and $0 \leq \theta \leq 1$,

$$f(\theta x + (1 - \theta) y) \preceq_K \theta f(x) + (1 - \theta) f(y) \tag{55}$$

- For $K = \mathbf{R}^n_+$, a K-convex function is a function for which each component function f_i is convex.
- Consider $K = \mathbf{S}^n_+$.
  - $f(X) = X^T X$ is K-convex on $\mathbf{R}^{n \times m}$.
  - $f(X) = X^{-1}$ is K-convex on $\mathbf{S}^n_{++}$.