Convex set : $x, y \in C$, $\theta \in [0,1] \rightarrow \theta x + (1-\theta) y \in C$ ($\bar{\theta}$) a convex combination

If $A_1, A_2 \cdots A_i$ are convex sets, $\cup A_i$ is a convex set

Affline space : $S = \{x \mid Ax = b\}$ is convex   polyhedra : $\{x \mid Ax \le b\}$

Convex hull : the smallest convex set containing S

$\text{conv} S = \{ \sum_{i=1}^{m} \theta_i x_i : m \in \mathbb{N}, x_i \in S$ iff $\sum_{0}^{m} c_i x_i = 0$ and $\sum_{0}^{m} c_i = 0 \Rightarrow c_i = 0$

Affinely independent for $x_0, x_1 \cdots x_m$, if $x_1 - x_0 \cdots x_m - x_0$ are linearly independent

Simplex $\text{conv}\{x_0, x_1 \cdots x_n\} = \{\theta_0 x_0 + \theta_1 x_1 + \cdots + \theta_m x_m\}$ probability simplex : $\Delta_n = \{\theta \mid \theta \ge 0, 1^T\theta = 1\}$

unit simplex : $\Delta_n' = \{\theta' \mid \theta' \ge 0, 1^T \theta' \le 1\}$

$\text{dist}(x, C) = \inf_{z \in C} \|x - z\|$ theorem : if C nonempty, closed and convex, $\exists$ unique $\hat{x}$ $\text{dist}(x, C) = \|x - \hat{x}\|$

projection onto a convex set $\hat{x}$: unique   here $\hat{x} = P_C(x)$

any z in C : $\langle x - \hat{x}, z - \hat{x} \rangle \le 0$ $\|P_C(x) - P_C(y)\| \le \|x - y\|$

for $x_0 \notin C$: exists a ω , $\sup_{x \in C} \langle \omega, x \rangle < \langle \omega, x_0 \rangle$

boundary of a set C is $\partial C = \bar{C} \setminus \text{int}\, C$

Supporting hyperplane theorem : $x_0 \in \partial C$, $\exists \omega$, $\langle \omega, x \rangle \le \langle \omega, x_0 \rangle$, $\forall x \in C$

if C is convex : $\text{int}\, C = \text{int}\, \bar{C}$ and $\partial C = \partial \bar{C}$

Separating hyperplane theorem : $\exists \omega$: $\omega^T x_1 \le b$, $\omega^T x_2 \ge b$

if $C_1 \cap C_2 = \emptyset$, $C = C_1 - C_2$ is convex and $\vec{0} \notin C$

Convex function : $f(\theta x + \bar{\theta} y) \le \theta f(x) + \bar{\theta} f(y)$ strict convex : $\le \rightarrow <$

$-f$ is (strictly) concave

affine function $f(x) = \omega^T x + b$ convex and concave (not strict)

f is convex iff $\forall x \in \text{dom} f$ and any direction d

extanded-value Extension : $\tilde{f} = \begin{cases} f(x) & , x \in S \\ \infty & , x \notin S \end{cases}$

effective domain $\text{dom} \tilde{f} = \text{dom} f = S = \{x : \tilde{f}(x) < \infty\}$

First-order condition for convexity : $f(y) \ge f(x) + \nabla f(x)^T (y - x)$

for univariate convex function : f' is increasing , $f''(x) \ge 0$ (semidefinite)

for strict condition, not necessary

$\nabla f(x^*) = 0$ for a convex function f, x^* is a global minimum

Negative entropy $f'(x) = \log x + 1$, $f''(x) = x^{-1} > 0$

Quadratic functions $f(x) = x^T Q x + b^T x + c$

Least squares loss $f(x) = \|Ax - y\|_2^2$

Log-sum-exp function $f(x) = \log(\sum_{i=1}^{n} e^{x_i})$

α - sublevel set : $C_\alpha = \{x \in \text{dom} f : f(x) \le \alpha\}$ if f is convex, C_α is convex

epigraph : $\text{epi} f = \{(x, y) : x \in S, y \ge f(x)\}$ and hypograph the opposite

f is convex iff $\text{epi} f$ is convex

$y = f(x_0) + \nabla f(x_0)^T (x - x_0)$ is a hyperplane of epi f

Jensen's inequality $f(\sum_{i=1}^{m} \theta_i x_i) \le \sum_{i=1}^{m} \theta_i f(x_i)$ e.g. $\sum_{i=1}^{n} \frac{1}{n} x_i \le \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2}$

Hölder's inequality: conjugate exponents p and q s.t. $p^{-1} + q^{-1} = 1$

$\sum_{i=1}^{n} |x_i y_i| \le \|x\|_p \|y\|_q$

Minkowski's inequality: $\|x+y\|_p \le \|x\|_p + \|y\|_p$

Nonnegative combination: $f(x) = \sum_{i=1}^{m} c_i f_i(x)$ is convex $\Leftarrow$ f_i are convex

Affine composition $f(x) = g(Ax+b)$ is convex $\Leftarrow$ g is convex e.g. $f(x) = \log(\sum_{i=1}^{n} e^{w_i^T x + b_i})$

Scalar composition $f(x) = h(g(x))$ $\begin{cases} \text{convex} \begin{cases} h \text{ convex } \nearrow \ g \text{ convex} \\ h \text{ convex } \searrow \ g \text{ concave} \end{cases} \\ \text{concave} \begin{cases} h \text{ concave } \nearrow \ g \text{ concave} \\ h \text{ concave } \searrow \ g \text{ convex} \end{cases} \end{cases}$

Vector composition $x \succeq y$ (component wise) $h(x) \ge (\le) h(y)$, h is increasing (decreasing)

Pointwise maximum f_i are convex and $f(x) = \max f_i(x)$ is convex Hinge function $(x)^+ = \max\{x, 0\}$

$f(x) = \sup f_i(x)$ is convex $\phi(\lambda) = \sup\{f(x) + \lambda^T g(x)\}$

Partial minimization: $g(x)$ convex and $\emptyset \ne C$ is convex: $f(x) = \inf_{y \in C} g(x, y)$ is convex

Convex optimization $\min_x f(x)$ s.t. cond1, cond2 ... feasible set X

Optimal value $f^* = \inf_{x \in X} f(x)$ can take $\begin{cases} +\infty & \text{infeasible} \\ -\infty & \text{unbounded low} \end{cases}$

Optimal point $f^* = f(x^*)$ not always attainable Local optimal: only on $\|x - x^*\| < \delta$

ε-suboptimal $f(x_0) \le f^* + \varepsilon$

the cond of a convex optimization: $\underbrace{g(x) \le 0}_{\text{convex}}$ $\underbrace{h(x) = 0}_{\text{aiffine}}$

the set of solution X_{opt} is convex

first-order optimality condition $\nabla f(x^*)^T (x - x^*) \ge 0$

linear program (LP) is: $\min_x c^T x$ s.t. $Bx \le d$, $Ax = b$

$Ax = b$, $x \ge 0$ (standard)

$Ax \le b$ (inequality)

quadratic program (QP): $\min \frac{1}{2} x^T Q x + c^T x$ s.t. $Bx \le d$ $Ax = b$

quadratically constrained quadratic program (QCQP): $\min \frac{1}{2} x Q^T x + c^T x$ s.t. $\frac{1}{2} x^T Q_i x + c_i^T x + d_i \le 0$

$Ax = b$

General unconstrained QP: $\min_x f(x) = \frac{1}{2} x^T Q x + b^T x + c$

Geometric program: $\min f(x)$ s.t. $g_i(x) \le 1$, $h_j(x) = 1$

h_j are minomials: f $R^n_{++} \to R$ $f(x) = r x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$

f, g_j are posynomials: Σf

Decent direction d_k: $g(t) \triangleq f(x_k + t d_k) < f(x_k) = g(0)$

$d_k \to g'(0) = d_k^T \nabla f(x_k) \le 0$

$g'(0) = d_k^T \nabla f(x_k) < 0 \to d_k$

Lipschitz continuity: $\|f(x) - f(y)\| \le L \|x - y\|$

L-smooth : $\|\nabla f(x) - \nabla f(y)\| \le L\|x-y\|$, $\forall x,y$

$\lambda_{max}(\nabla^2 f(x)) \le L$ $|\lambda| \le L$ for all eigenvalues λ of $\nabla^2 f(x)$

Quadratic upper bound : $f(y) \le f(x) + \nabla f(x)^T(y-x) + \frac{L}{2}\|y-x\|^2$

Consequence of qudratic upper bound $f(x_{k+1}) \le f(x_k) - t(1-\frac{Lt}{2})\|\nabla f(x_k)\|^2$ for $0 < t \le \frac{1}{L}$:

Convergence analysis $f(x_k) - f(x^*) \le \frac{\|x_0 - x^*\|^2}{2tk}$ $f(x_k) - f(x_{k+1}) \ge \frac{t}{2}\|\nabla f(x_k)\|^2$

Strong convexity : $\tilde{f}(x) = f(x) - \frac{m}{2}\|x\|^2$ is convex

$\Longleftrightarrow$ $f(y) \ge f(x) + \nabla f(x)^T(y-x) + \frac{m}{2}\|y-x\|^2$ $\Longleftrightarrow$ $\lambda_{min}(\nabla^2 f(x)) \ge m$

$\frac{m}{2}\|x-y\|^2 \le f(y) - f(x) + \nabla f(x)^T(y-x) \le \frac{m}{2}\|x-y\|^2$

$f(x) - f(x^*) \le \frac{1}{2m}\|\nabla f(x)\|^2$

$f(x_k) - f(x^*) \le (1-mt)^k\|x_0 - x^*\|^2$

$\|x_k - x^*\|^2 \le (1-mt)^k\|x_0 - x^*\|^2$

Condition number $\kappa(Q) = \frac{\lambda_{max}(Q)}{\lambda_{min}(Q)}$ well-conditioned if small ill-condition if large

Exact line search

Backtracking line search

Newton

Stop here. Mid term.