1 Convex sets

1.1 Definition

Affine set :	$x_1, x_2 \in C \Rightarrow \theta_1 x_1 + \theta_2 x_2 \in C, \forall \theta_1, \theta_2 \in \mathbb{R}$
Convex set :	$x_1, x_2 \in C \Rightarrow \theta_1 x_1 + \theta_2 x_2 \in C, \forall \theta_1, \theta_2 \in \mathbb{R}_+, \theta_1 + \theta_2 = 1$
Cone set (not necessarly convex) :	$x \in C \Rightarrow \theta x \in C, \forall \theta, \in \mathbb{R}_+$
Convex cone :	$x_1, x_2 \in C \Rightarrow \theta_1 x_1 + \theta_2 x_2 \in C, \forall \theta_1, \theta_2 \in \mathbb{R}_+$

$a \in \mathbb{R}^n,\ b \in \mathbb{R},\ r \in \mathbb{R}_{++},\ P \in \mathcal{S}^n_{++}$

1.2 Simple Examples

$\{x \,\vert\, a^T x = b\}$	(Hyperplane affine)
$\{x \,\vert\, a^T x \leq b\}$	(Halfspace)
$\{x \,\vert\, (x-a)^T (x-a) \leq r^2\}$	(Euclidian ball)
$\{x \,\vert\, (x-a)^T P^{-1} (x-a) \leq r^2\}$	(Ellipsoid)
$\{x \,\vert\, \Vert x - a \Vert \leq r\}$	(Norm ball)
$\{(x,t) \,\vert\, \Vert x \Vert \leq t\}$	(Norm cone, convex cone)
$\{x \,\vert\, a_j^T x \leq b_j, 1 \leq j \leq m,\ c_j^T x = d_j, 1 \leq j \leq p\}$	(Polyhedron)
$\{x \,\vert\, x_i \geq 0, 1 \leq i \leq n\}$	(Non negative orthant, polyhedral cone)
$\{\theta^T x \,\vert\, \theta \succeq 0, 1^T \theta = 1\}$	(Simplex)
$\mathcal{S}^n_+$	(Semidefine matrices, convex cone)

1.3 Operation that preserve convexity

Intersection : $\mathcal{S}^n_+ = \cap_{z \in \mathbb{R}^n} \{X \in \mathcal{S}^n \,|\, z^T X z \geq 0\}$
Image/Inverse by affine function :
- **LMI** : $\{x \in \mathbb{R}^n \,|\, A(x) = x_1 A_1 + \cdots + x_n A_n \preceq B\}, A_i, B \in \mathcal{S}^n = f^{-1}(S_n), f(x) = B - A(x)$
- **Hyperbolic cone** : $\{x \,|\, x^T P x \leq (c^T x)^2, c^T x \geq 0\} = f(\{(z,t) \,|\, z^T z \leq t^2, t \geq 0\}, f(x) = (P^{\frac{1}{2}} x, c^T x)$

Image/Inverse by perspective : $P(z,t) = z/t,\ z \in \mathbb{R}^n, t \in \mathbb{R}_{++}$
Image/Inverse by linear fractional : $f(x) = \frac{Ax+b}{c^T x + d}$, $\mathbf{dom}\, f = \{x \,|\, c^T x + d > 0\}$

1.4 Results

Insersection with line : set is convex $\iff$ its intersection with any line is convex
Midpoint convex : C closed and midpoint convex $\Rightarrow$ C is convex
Separating hyperplane : $\mathcal{C}, \mathcal{D}$ disjoint convex set, are separated by a hyperplane
Supporting hyperplane in $x_0 \in \mathbf{bd}\, \mathcal{C}$: $\{x \,|\, a^T (x - x_0) = 0\}$ if $\forall x \in \mathcal{C}, a^T (x - x_0) \leq 0$.
- Hyperplane theorem : $\forall x_0 \in \mathbf{bd}\, C$, $\exists$ supporting hyperplane in x_0, converse if the set is closed, with non empty interior.

Proper cone : K convex, K closed, K has nonempty interior, K contains no lines (Ex : Non negative orthant, $\mathcal{S}^n_+$)
Dual cone : $K^* = \{y \,|\, x^T y \geq 0, \forall x \in K\}$, convex even if K is not

2 Convex functions

Definition : $\mathbf{dom}\, f$ convex and $\forall x, y \in \mathbf{dom}\, f, 0 \leq \theta \leq 1, f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y)$
Line segment : f convex $\iff \forall x, y \in \mathbf{dom}\, f,\ g(t) = f(x + ty)$ convex ($\mathbf{dom}\, g = \{t \,|\, x + ty \in \mathbf{dom}\, f\}$
First order conditions ($f \in \mathcal{C}^1$) : f convex $\iff$ $\mathbf{dom}\, f$ convex and $\forall x, y, f(y) \geq f(x) + \nabla f(x)^T (y - x)$
- f strictly convex $\iff$ $\mathbf{dom}\, f$ convex and $\forall x, y,\ x \neq y \Rightarrow f(y) > f(x) + \nabla f(x)^T (y - x)$

Second order conditions ($f \in \mathcal{C}^2$) : f convex $\iff$ $\mathbf{dom}\, f$ convex and $\forall x \in \mathbf{dom}\, f, \nabla^2 f(x) \succeq 0$
- $\mathbf{dom}\, f$ convex and $\forall x \in \mathbf{dom}\, f, \nabla^2 f(x) \succ 0 \Rightarrow f$ strictly convex (converse not true)
- $f(x) = x^T P x + q^T x + r$ is convex $\iff P \succeq 0$

2.1 Simple Examples

On $\mathbb{R}$

e^{ax}	on $\mathbb{R}$	convex	$a \in R$	(Exponential)
x^a	on $\mathbb{R}_{++}$	convex	$a \geq 1$ or $a \leq 0$	(Powers)
		concave	$0 \leq a \leq 1$	
$\|x\|^p$	on $\mathbb{R}$,	convex	$p \geq 1$	(Powers of abs value)
$\log(x)$	on $\mathbb{R}_{++}$	concave		(Logarithm)
$x\log(x)$	on $\mathbb{R}_+$	convex		(Negative entropy)

On $\mathbb{R}^n$

$\|\cdot\|$	on $\mathbb{R}^n$	convex	(Norm)	(Def)
$f(x) = \max\{x_1, \dots x_n\}$	on $\mathbb{R}^n$	convex	(Max)	(Def)
$f(x,y) = x^2/y$	on $\mathbb{R} \times \mathbb{R}_{++}$	convex	(Quadratic over linear)	$(Hess)$
$f(x) = \log(e^{x_1} + \dots + e^{x_n})$	on $\mathbb{R}^n$	convex	(Logsum exp (soft max))	$(Hess)$
$f(x) = (x_1 x_2 \dots x_n)^{\frac{1}{n}}$	on $\mathbb{R}^n_{++}$	concave	(Geometric mean)	$(Hess)$
$f(X) = \log\det X$	on $\mathcal{S}^n_{++}$	concave	(Log-determinant)	$(Line + Hess)$

Sublevel set : $C_\alpha = \{x \in \textbf{dom}\, f \,|\, f(x) \leq \alpha\}$
- f convex $\Rightarrow$ sublevels sets are convex. Sublevels sets are convex $\Rightarrow$ f quasiconvex.
- f concave $\Rightarrow$ supperlevels sets are convex.

Epigraph : $\textbf{epi}\, f = \{(x,t) \in \textbf{dom}\, f \times \mathbb{R} \,|\, f(x) \leq t\}$
- f convex $\iff$ $\textbf{epi}\, f$ convex set.
- f concave $\iff$ $\textbf{hypo}\, f$ convex set.
- $f(x,Y) = x^TY^{-1}x$ on $\mathbb{R}^n \times \mathcal{S}^n_{++}$ (Use epigraph and Schur complement)

2.2 Operation that preserve convexity

Non negative weightes sum : $f = w_1 f_1 + \dots + w_m f_m$ convex if f_i convex and $w_i \geq 0$

Affine composition : $g(x) = f(Ax+b)$ convex if f convex, $A \in \mathbb{R}^{n\times m}$, $b \in \mathbb{R}^m$

Pointwise maximum : $f(x) = \max\{f_1(x), \dots f_m(x)\}$ convex if f_i convex.

Pointwise supremum : if $\forall y \in A, x \to f(x,y)$ convex, $g(x) = \sup_{y\in A}\{f(x,y)$ convex.
- Distance to farthest point of C, $f(x) = \sup_{y \in C} normx - y$
- Maximum eigenvalue of symetric matrix, $f(X) = \sup\{y^TXy \,|\, \|y\|_2 = 1\}$
- Maximum singularvalue of a matrix, $f(X) = \sup\{u^TXv \,|\, \|u\|_2 = 1, \|v\|_2 = 1\}$
- Every convex function can be expressed as a supremum of affine functions

Composition : $h \in \mathbb{R}^k \to \mathbb{R}$, $g \in \mathbb{R}^n \to \mathbb{R}^k$
- $h \circ g$ convex if h convex, g_i are convex and $\tilde{h}$ nondecreasing in each argument (g K-convex and h K-nondecreasing, K is the non-negative orthant)
- $h \circ g$ convex if h convex, g_i are concave and $\tilde{h}$ nonincreasing in each argument
- g convex $\Rightarrow$ $exp(g)$ convex, g concave positive $\Rightarrow$ $\log(g(x))$ and $\frac{1}{g(x)}$ concave, g convex $\geq 0 \Rightarrow g(x)^p$ convex $(p \geq 1)$

Minimization : if f convex in (x,y) and C convex nonempty, $g(x) = \inf_{y\in C} f(x,y)$ convex
- Distance of a point to a convex set C, $f(x) = \inf_{y \in C} \|x-y\|$

Perspective : if $f : \mathbb{R}^n \to \mathbb{R}$ convex, $g(x,t) = tf(x/t)$ is convex
- $f : \mathbb{R}^m \to \mathbb{R}$, convex, $g(x) = (c^Tx + d)\frac{f(Ax+b)}{c^Tx+d}$ is convex

Conjuguate function : $f^*(y) = \sup_{x \in \textbf{dom}\, f}(y^Tx - f(x))$
- $f(x) = ax + b$, $\textbf{dom}\, f^* = \{a\}$, $f^*(a) = -b$
- $f(x) = -\log(x)$, $\textbf{dom}\, f^* = -\mathbb{R}_{++}$, $f^*(y) = -\log(-y) - 1$
- $f(x) = e^x$, $\textbf{dom}\, f^* = \mathbb{R}_+$, $f^*(y) = y\log(y) - y$
- $f(x) = x\log(x)$, $\textbf{dom}\, f^* = \mathbb{R}$, $f^*(y) = e^{y-1}$
- $f(x) = \frac{1}{x}$, $\textbf{dom}\, f^* = -\mathbb{R}_+$, $f^*(y) = -2\sqrt{-y}$
- $f(x) = \frac{1}{2}x^TQx$, $\textbf{dom}\, f^* = \mathbb{R}^n$, $f^*(y) = \frac{1}{2}y^TQ^{-1}y$
- $f(x) = \log\det X^{-1}$, $\textbf{dom}\, f^* = -\mathcal{S}^n_{++}$, $f^*(y) = \log\det -Y^{-1} - n$
- $f(x) = \|x\|$, $f^*(y) = 0$ if $\|y\|_* \leq 1$, ∞ otherwise

Quasiconvex function : $f : \mathbb{R}^n \to R$ quasiconvex if $\textbf{dom}\, f$ and all its sublevel set are convex
- Quasiconcave if $-f$ quasiconvex. Quasilinear if quasiconvex and quasiconcave.

2.3 Generalized inequalities

K-nondecreasing : $f : \mathbb{R}^n \to \mathbb{R}$ K-nondecreasing if $x \preceq_K y \Rightarrow f(x) \leq f(y)$
- K-increasing if $x \preceq_K y, x \neq y \Rightarrow f(x) < f(y)$

Convexity : $f : \mathbb{R}^n \to \mathbb{R}^m$ convex : $\forall x, y \in \textbf{dom}\, f, 0 \leq \theta \leq 1$, $f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta)f(y)$
- f convex w.r.t matrix inequality $\iff$ $\forall z \in \mathbb{R}^n, x \to z^Tf(x)z$ convex.

3 Optimization problems

Optimization problem in standard form :

$$\begin{array}{ll} \underset{x}{\text{minimize}} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0 \quad i = 1 \dots m \\ & h_i(x) = 0 \quad i = 1 \dots p \end{array}$$

x : optimization variable, $f_i(x) \leq 0$ inequality constraints, $h_i(x) = 0$ equality constraints, p^* optimal value, x^* optimal point.

Convex optimization problem :

$$\begin{array}{ll} \underset{x}{\text{minimize}} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0 \quad i = 1 \dots m \\ & a_i^T x = b_i \quad i = 1 \dots p \end{array}$$

$f_0, \dots f_m$ convex. Any locally optimal point is (globally) optimal.

3.1 Convex optimization problems

Linear optimization problem (LP) :

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^T x + d \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{array}$$

Quadratic program (QP) :

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \frac{1}{2} x^T P x + q^T x + r \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{array}$$

Quadraticly constraint quadratic program (QCQP) :

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \frac{1}{2} x^T P x + q^T x + r \\ \text{subject to} & \frac{1}{2} x^T P_i x + q_i^T x + r_i \preceq 0 \quad i = 1, \dots, m \\ & Ax = b \end{array}$$

Second order cone program (SOCP) :

$$\begin{array}{ll} \underset{x}{\text{minimize}} & f^T x \\ \text{subject to} & \|A_i x + b_i\| \leq c_i^T x + d_i \quad i = 1, \dots, m \\ & Fx = g \end{array}$$

Equivalent to QCQP is $c_i = 0, \forall i$.

3.2 Extensions

Geometric programming (GP) :

$$\begin{array}{ll} \underset{x}{\text{minimize}} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0 \quad i = 1 \dots m \\ & h_i(x) = 0 \quad i = 1 \dots p \end{array}$$

$f_0, \dots f_m$ are posynomial, $h_0, \dots h_p$ monomial.
(f monomial $\iff f(x) = c x_1^{a_1} \dots x_n^{a_n}$, $a_i \in R, c > 0$, posynomial is sum of monomials).

GP can be transformed to convex problems by change of variable $y_i = \log(x_i)$

Semidefinite programming (SDP) :

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^T x \\ \text{subject to} & x_1 F_1 + \dots + x_n F_n \preceq 0 \\ & Ax = b \end{array}$$

$F_1, \dots, F_n \in \mathcal{S}^k$, $A \in \mathbb{R}^{p \times n}$ (Similar to LP)

3.3 Examples

Chebyshev center (LP) : fit $\mathcal{B} = \{x_c + u \,|\, \|u\|_2 \leq r\}$ in $\mathcal{P} = \{x \,|\, a_i^T x \leq b_i, i = 1, \dots, m\}$

$$\begin{array}{ll} \underset{x}{\text{minimize}} & r \\ \text{subject to} & a_i^T x_c + r\|a_i\|_2 \leq b_i \quad i = 1, \dots, m \end{array}$$

Chebyshev center (LP) : fit $\mathcal{B} = \{x_c + u \,|\, \|u\|_2 \leq r\}$ in $\mathcal{P} = \{x \,|\, a_i^T x \leq b_i, i = 1, \dots, m\}$

$$\begin{array}{ll} \underset{x}{\text{minimize}} & r \\ \text{subject to} & a_i^T x_c + r\|a_i\|_2 \leq b_i \quad i = 1, \dots, m \end{array}$$

4 Duality

We consider problem in standard form.
Lagrangian : $L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^m \nu_i h_i(x)$. **dom** $L = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$.

— λ_i, ν_i are the Lagrange multipliers

Lagrange dual function : $g(\lambda, \nu) = \inf_{x \in \mathcal{D}} \mathcal{L}(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^m \nu_i h_i(x)$
Lower bound on p^* : $\forall \lambda \succeq 0,\ \forall \nu \in \mathbb{R}^p,\ g(\lambda, \nu) \leq p^*$
Lagrange dual problem :

$$\begin{array}{ll} \underset{\lambda, \nu}{\text{maximize}} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

— always a convex problem (concave maximization problem), even if the primal is not convex

Weak duality : $d^* \le p^*$, (always hold even when p^* and d^* infinite). $p^* - d^*$ is the duality gap

Strong duality : if $d^* = p^*$ holds, we have strong duality.

Slater's condition : for a convex problem, if $\exists x \in \textbf{relint}\, \mathcal{D},\ s.t. f_i(x) < 0,\ i = 1, \dots, m,\ Ax = b$ (strict feasibility), then strong duality holds (can be relaxed to $f_i(x) \le 0$ for affine constraints) and dual optimal value is attained.

Complementary slackness : suppose x^* primal optimal, (λ^*, ν^*) dual optimal, then x^* minimizes $\mathcal{L}(x, \lambda^*, \nu^*)$ over x and $\lambda_i^* f_i(x^*) = 0,\ i = 1, \dots, m$

KKT conditions : suppose f_i and h_i are differentiable, let x^* primal optimal, (λ^*, ν^*) dual optimal, then

— $\nabla f_0(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^m \nu_i^* \nabla h_i(x^*) = 0$

— $f_i(x^*) \le 0,\ i = 1, \dots, m$

— $h_i(x^*) = 0,\ i = 1, \dots, m$

— $\lambda_i^* \ge 0,\ i = 1, \dots, m$

— $\lambda_i^* f_i(x^*) = 0,\ i = 1, \dots, m$

KKT condition for convex pb : KKT condition are also sufficient if the problem is convex, e.g, if $(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$ satisfy the KKT conditions, then $\tilde{x}$ is primal optimal, $(\tilde{\lambda}, \tilde{\nu})$ is dual optimal and we have 0 duality gap.

KKT with Slaters condition : if Slater's condition holds, x is optimal i.i.f there are (λ, ν) s.t (x, λ, ν) satisfy the KKT conditions.

4.1 Examples

$$\begin{array}{ll} \min\limits_x & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \\ & \\ \max\limits_x & -b^T \mu \\ \text{subject to} & A^T \mu + c \succeq 0 \end{array}$$

$$\begin{array}{ll} \min\limits_x & c^T x \\ \text{subject to} & Ax \preceq b \\ & \\ \max\limits_x & -b^T \lambda \\ \text{subject to} & A^T \lambda + c = 0 \\ & \lambda \succeq 0 \end{array}$$

$$\begin{array}{ll} \min\limits_x & \|x\| \\ \text{subject to} & Ax = b \\ & \\ \max\limits_x & -b^T \mu \\ \text{subject to} & \|A^T \lambda\| \le 1 \end{array}$$

$$\begin{array}{ll} \min\limits_x & \log\det X^{-1} \\ \text{subject to} & a_i^T X a_i \le 1,\ i = 1, \dots, m \\ & \\ \max\limits_x & -\log\det(\sum_{i=1}^m \lambda_i a_i a_i^T) \\ & -1^T \lambda + n \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

$$\begin{array}{lll} \min\limits_x & \frac{1}{2} x^T P_0 x + q_0^T x + r_0 & \\ \text{subject to} & \frac{1}{2} x^T P_i x + q_i^T x + r_i \le 0 & i = 1, \dots, m \\ & & \\ \max\limits_x & -\frac{1}{2} q(\lambda)^T P(\lambda) q(\lambda) + r(\lambda) & \\ \text{subject to} & \lambda \succeq 0 & \end{array}$$

4.2 Extra stuff

Dual norm : $\|z\|_* = \sup\{z^T x \,|\, \|x\| \le 1\}$

— $\forall x, z \in \mathbb{R}^n,\ z^T x \le \|x\| \|z\|_*$

Eigenvalues and singular values inequalities : $\lambda_{\max}(A) = \sup_{x \ne 0} \frac{x^T A x}{x^T x}$, $\lambda_{\min}(A) = \inf_{x \ne 0} \frac{x^T A x}{x^T x}$ $(A \in \mathcal{S}^n)$, $\sigma_{\max}(A) = \sup_{x \ne 0, y \ne 0} \frac{x^T A y}{\|x\|_2 \|y\|_2}$

Schur complement : Let$X \in \mathcal{S}^n$, $X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$ where $A \in \mathcal{S}^k$, If $\det A \ne 0$, $S = C - B^T A^{-1} B$ is the Schur complement of A in X.

— $\det X = \det A \det S$

— $\inf_u \begin{bmatrix} u \\ v \end{bmatrix}^T \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = v^T S v$

— $X \succ 0 \iff A \succ 0$ and $S \succ 0$

— if $A \succ 0$ then $X \succeq 0 \iff S \succeq 0$

Taylor's approximation : $\hat{f}(x + v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f v$

Newton's method : $x_{n+1} \leftarrow x_n - \alpha_n (\nabla^2 f(x))^{-1} \nabla f(x)$

Some Gradients :

— $\nabla_x (a^T x + b) = a$

— $\nabla_x (\frac{1}{2} x^T A x) = \frac{1}{2}(A^T + A)x$

— $\nabla_x (\mathrm{Tr}\,(A^T X + b) = A$

— $\nabla_x (\det(X)) = \bar{X}$, $\bar{X}$ comatrix of X $(\bar{X} = \det(X) X^{-T})$

— $\nabla_x (\log\det(X)) = X^{-1}$

— $f(X) = X^{-1} \Rightarrow \nabla_x f(H) = -X^{-1} H X^{-1}$