1 Affine and Convex Sets

1.1 Line and Line Segments

Set of points of the form $y = \theta x_1 + (1-\theta)x_2$, where $\theta \in \mathbf{R}$

1.2 Affine sets

A set C is *affine* if the line through any two distinct points in C lies in C.
Affine combination is a point of a form $\sum_{i=1}^{k} \theta_i x_i$, where $\sum_{i=1}^{k} \theta_i = 1$.
Prove by induction that C is affine $\iff$ C contains all possible affine combinations.
If C is an affine set and $x_o \in C$, then $V = C - x_0$ is a subspace.

1.3 Affine hull

$\mathbf{aff} C = \{\sum_{i=1}^{k} \theta_i x_i | x_i \in C, \sum_{i=1}^{k} \theta_i = 1\}$ It is the smallest affine set that contains C. If S is any affine set such that $C \subseteq S$, then $\mathbf{aff}\ C \subseteq S$.

1.4 Relative interior

Affine dimension of C is $\dim \mathbf{aff}\ C$.
$\mathbf{relint}\ C = \{x \in C \mid B(x,r) \cap \mathbf{aff}\ C \subseteq C$ for some $r > 0\}$
Think of relative interior as the analogous version of normal interior, just as the universe is compressed from the original ambient set to the affine hull, with possibly a fall in dimensions. Assume Euclidean norm and Euclidean ball for understanding.

1.5 Convex sets

A set C is *convex* if the line segment between any two points in C lies in C. Line segment is a restriction of the line, with $\theta \in [0,1]$.
Convex combination takes the same form as the affine combination, with additional restriction of each $\theta_i \in [0,1]$.
Prove by induction that C is convex $\iff$ C contains all possible convex combinations.

1.6 Convex hull

$\mathbf{conv}\ C = \{\sum_{i=1}^{k} \theta_i x_i \mid x_i \in C, \sum_{i=1}^{k} \theta_i = 1, \theta_i \in [0,1]\}$
It is again the smallest convex set that contains C.If S is any convex set such that $C \subseteq S$, then $\mathbf{conv}\ C \subseteq S$.

1.7 Cones

A set C is a *cone* if $\forall x \in C$ and $\theta > 0$ we have $\theta x \in C$. C is convex cone if it satisfies both definitions.
Conic combination is a point of the form $\sum_{i=1}^{k} \theta_i x_i$, where $\theta_i \geq 0$.
Prove by induction that C is a **convex** cone $\iff$ C contains all possible conic combinations. Idea of *conic hull* follows similarly.

1.8 Examples

Empty set $\emptyset$, singleton, whole space $\mathbf{R}^n$ are affine. Line is affine, and is a convex cone if it passes through zero (making it a subspace). Non-trivial line segment is convex. Non-trivial ray is convex, and a convex cone if x_0 is 0. Any subspace is affine and a convex cone.

1.9 Hyperplanes and halfspaces

Set of the form $\{x \in \mathbf{R}^n | a^T x = b\}$ where $a \in \mathbf{R}^n, a \neq 0,$ and $b \in \mathbf{R}$. Natural extension of 3D plane seen in $\mathbf{R}^3$.
A hyperplane divides $\mathbf{R}^n$ into two *halfspaces*, which is a set of the form $\{x \in \mathbf{R}^n | a^T x \leq b\}$, where $a \neq 0$.

1.10 Some more examples

Euclidean ball in $\mathbf{R}^n$: $B(x_c, r) = \{x \in \mathbf{R}^n : |x - x_c|_2 \leq r\} = \{x_c + ru : |u|_2 \leq 1\}$ It is convex by homogeneity and triangle inequality of norm property.
Ellipsoids in $\mathbf{R}^n$ $\varepsilon = \{x \in \mathbf{R}^n : (x-x_c)^T P^{-1}(x - x_c \leq 1)\}$, where P is symmetric and positive definite. When $P = k^2 I$, the ellipsoid is reduced to a ball.
Norm ball, which is an Euclidean ball generalised to any norm, is convex. *Norm cone* associated with the norm $|\cdot|$ is $C = \{(x,t) \in \mathbf{R}^{n+1} : |x| \leq t\} \subseteq \mathbf{R}^{n+1}$, which is a convex cone.
Polyhedra is a finite intersection of halfspaces and hyperplanes, denoting inequality and equality constraints respectively in a linear programming problem.
Affinely independent means that for $k+1$ points $v_0, \ldots v_k \in \mathbf{R}^n$, then $v_i - v_0$ are linearly independent.
Simplex determined by those vectors is $\mathbf{conv}\{v_i\}$, which is a coordinate-free version of linear independence.
Set of symmetric n*n matrices $S^n = \{X \in \mathbf{R}^{n*n} : X = X^T\}$, set of symmetric positive semidefinite matrices $S_+^n = \{X \in S^n : X \succeq 0\}$, set of symmetric positive definite matrices $S_{++}^n = \{X \in\ S^n : X \succ 0\}$.

1.11 Operations preserving convexity

Finite/infinite intersections of convex sets: **Every** closed convex set is the intersection of all halfspaces containing it.
Affine function: function $f : \mathbf{R}^n \to \mathbf{R}^m$ with the form $f(x) = Ax + b$, where $A \in \mathbf{R}^{m \times n}$ and $B \in \mathbf{R}^m$ If $S \in \mathbf{R}^n$ is convex, then $f(S)$ is convex. If $P \in \mathbf{R}^m$ is convex, then $f^{-1}(P)$ is convex. Examples include scaling, translation, projection, sum, partial sum.
Linear fractional and perspective functions: $P : \mathbf{R}^{n+1} \to \mathbf{R}^n$, $\mathbf{dom}P = \mathbf{R}^n \times \mathbf{R}^{++}$, with the form $P(z,t) = z/t$. If $C \subseteq \mathbf{dom}P$ is convex, then $P(C)$ is convex. Linear fractional is formed by a composition of affine function and perspective function, f is a linear fraction if $f = P \circ g$, where g is an affine function.

1.12 Generalised order

Proper cone: a cone K which is convex, closed, solid(interior not empty), pointed(if $x \in K$ and $-x \in K$, then $x = 0$), then we can define $x \prec_K y \Leftrightarrow y - x \in K$, and $x \prec_K y \Leftrightarrow y - x \in \mathbf{int}K$. Some properties: preservation under addition, nonnegative scaling and limits, transitivity, reflexivity, anti-symmetric.

1.13 Separating hyperplane theorem

Theorem: suppose C and D are convex, nonempty and disjoint sets, $\exists a \neq 0$and$b$such that$a^T x \leq b \forall x \in Canda^T x >= b\ \forall x \in D$. If at least one of C and D is bounded, and both are closed, the *strict* inequalities hold. *Intuition*: Construct singleton/finite sets to satisfy strict separation condition. *Proof*: Fix two points in C and D that achieves the distance of two sets, then $a = d - c$, $b = \frac{1}{2}(|d|_2^2 - |c|_2^2)$. Prove by contradiction that it is true.

1.14 Supporting hyperplane

Suppose $C \in \mathbf{R}^n$, and $x_0 \in \mathbf{bd}C = \mathbf{cl}C \setminus \mathbf{int}C$, if $\exists a \neq 0$ such that $a^T x \leq a^T x_0\ \forall x \in C$, then the hyperplane $a^T x = a^T x_0$ is a *supporting hyperplane* to C at point x_0. For every nonempty convex set C and any $x_0 \in \mathbf{bd}C$, such a plane exists.

1.15 Dual cone

Let K be a cone. We define $K^* = \{y : x^T y \geq 0 \forall x \in K\}$ as the *dual cone* of K, and the dual cone is always a convex cone. Geometrically, dual cone contains vectors that make acute angle with all vectors in the original cone.

2 Convex functions

2.1 First-order conditions

If f is differentiable, then f is convex iff $\mathbf{dom}f$ is convex and $f(y) \geq f(x) + \nabla f(x)^T (y-x)\ \forall x, y \in \mathbf{dom}f$. Note that zero gradient in this case implies a global minimum, and *the* global minimum if f is strictly convex. Strict convexity can be implied by strict inequality.

2.2 Second-order conditions

Hessian matrix: $H_{ij} = \frac{\delta f}{\delta x_i \delta x_j}$. f is convex iff $\mathbf{dom}f$ is convex and its Hessian is positive semidefinite.

2.3 Examples

Exponential: e^{ax} is convex on $\mathbf{R} \forall a \in \mathbf{R}$.
Powers: x^a is convex on $\mathbf{R}^{++}$ when $a \geq 1$or$a \leq 0$, and concave for $a \in [0,1]$.
Powers of absolute value: $|x|^p$, for $p \geq 1$, is convex on $\mathbf{R}$.
Logarithm: $\log x$ is concave on $\mathbf{R}^{++}$.
Negative entropy: $x \log x$ is strictly convex on $\mathbf{R}^{++}$.
Other examples include norms, max, quadratic-over-linear, log-sum-exp, geometric mean, log-determinant.

2.4 Sublevel sets

The α-sublevel set of a function $f : \mathbf{R}^n \to \mathbf{R}$ is defined as $C_\alpha = \{x \in \mathbf{dom}f : f(x) \leq \alpha\}$. It is convex if f is convex for any value of α. Analogous α-superlevel set by $\{x \in \mathbf{dom}f : f(x) \geq \alpha\}$ is convex if f is concave. Converse is not true. Consider $-e^x$.

2.5 Epigraph

$\mathbf{graph}(f) = \{(x, f(x)) : x \in \mathbf{dom}f\}$
$\mathbf{epi}(f) = \{(x,t) : x \in \mathbf{dom}f, t \geq f(x)\}$
$\mathbf{hypo}(f) = \{(x,t) : x \in \mathbf{dom}f, t \leq f(x)\}$
f is convex iff $\mathbf{epi}(f)$ is convex, and is concave iff $\mathbf{hypo}(f)$ is convex.

2.6 Operations preserving convexity

Conic combinations of convex functions
Composition with affine functions
Pointwise max/sup: geometrically, the epigraph of the max/sup function is the intersection of epigraphs of all component functions, which are all convex. **Intuition**: establish convexity of a function by construction from max/sup of other more obvious convex functions.
Composition: Assume **twice differentiability**, check cases to ensure the second derivative is non-negative. **Some compositions**: If g is convex, then $\exp g(x)$ is convex. If g is concave and positive, then $\log g(x)$ is concave. If g is concave and positive, then $\frac{1}{g(x)}$ is convex. If g is convex and non-negative and $p \geq 1$, then $g(x)^p$ is convex. If g is convex then $-\log(-g(x))$ is convex on $\{x \in \mathbf{dom}g : g(x) < 0\}$.
Minimization: if f is convex in (x,y), and C is convex and nonempty, then $g(x) = \inf_{y \in C} f(x,y)$ is convex. $\mathbf{dom}g = \{x : \exists y \in C, (x,y) \in \mathbf{dom}f\}$

3 Conjugate functions

3.1 Definitions

Let f: $\mathbf{R}^n \to \mathbf{R}$, then *conjugate* of f, which is f^*: $\mathbf{R}^n \to \mathbf{R}$ is defined as $f^*(y) = \sup_{x \in \mathbf{dom} f}(y^T x - f(x))$. The domain of f^* consists of $y \in \mathbf{R}^n$ where the supremum is finite. We can understand it as the maximum gap between the linear function yx and $f(x)$. If f is differentiable, this occurs when $f'(x) = y$.
f^* is convex in y as it is the pointwise supremum of a family of convex(indeed affine) functions of y.

3.2 Examples

Affine functions: $f(x) = ax + b$, $x \in \mathbf{R}$, then $f^*(y) = \sup_{x \in \mathbf{R}}\{xy - ax - b\} = \sup_{x \in \mathbf{R}}\{x(y-a)\} - b$, hence $f^*(y) = -b$ when $y = a$ and $+\infty$ otherwise.
Negative logarithm: $f(x) = -\log(x), x \in \mathbf{R}^{++}$. $f^*(y) = \sup_{x \in \mathbf{R}^{++}}\{xy + \log(x)\}$, hence $f^*(y) = +\infty$ when $y \geq 0$ and $f^*(y) = -\log(-y) - 1$ when $y < 0$ by differentiation. **Intuition**: divide cases for y, draw graphs to see the sum and find potential extremum by calculus.
Exponential: $f(x) = e^x$, $x \in \mathbf{R}$, then $f^*(y) = \sup_{x \in \mathbf{R}}\{xy - e^x\}$, hence $f^*(y) = +\infty$ when $y < 0$, and $f^*(y) = y\log(y) - y$ when $y > 0$ and $f^*(0) = 0$.
Inverse: $f(x) = \frac{1}{x}$, $x \in \mathbf{R}^{++}$, then $f^*(y) = \sup_{x \in \mathbf{R}^{++}}\{xy - \frac{1}{x}\}$, hence $f^*(y) = +\infty$ when $y > 0$, $f^*(y) = -2(-y)^{\frac{1}{2}}$ when $y < 0$, and $f^*(y) = 0$ when $y = 0$.
Log-sum-exp function: $f(x) = \log(\Sigma_{i=1}^n e^{x_i})$, $x \in \mathbf{R}^n$, then $\mathrm{dom} f^* = \{y \in \mathbf{R}^n : y_i \geq 0, \Sigma_{i=1}^n y_i = 1\}$, and $f^*(y) = \Sigma_{i=1}^n y_i \log(y_i)$

3.3 Properties

By definition, we have $f^*(y) \geq x^T y - f(x) \to f^*(y) + f(x) \geq x^T y\ \forall x \in \mathbf{dom}f$, which is Fenchel's inequality.
Scaling and composition with affine function: For $a > 0$ and $b \in \mathbf{R}$, the conjugate of $g(x) = af(x) + b$ is $g^*(y) = af^*(\frac{y}{a}) - b$, and $\mathbf{dom}g^* = \{y \in \mathbf{R}^n : \frac{y}{a} \in \mathbf{dom}f^*\}$.
Sum of independent functions: $f(u,v) = f_1(u) + f_2(v)$, then $f^* = f_1^* + f_2^*$.
Conjugate of conjugate: Let $f : \mathbf{R}^n \to \mathbf{R}$, then (i)$f(x) \geq f^{**}(x) \forall x \in \mathbf{R}^n$ (ii)If f is closed($\mathbf{epi}f$ is closed) and convex, then $f^{**} = f$

4 Quasiconvex functions

4.1 Definition

$f : \mathbf{R}^n \to \mathbf{R}$ is *quasiconvex* if $\mathbf{dom}f$ is convex and the sublevel sets $S_\alpha = \{x \in domf : f(x) \leq \alpha\}$ are convex $\forall \alpha \in \mathbf{R}$.$f$ is *quasiconcave* if $-f$ is quasiconvex. f is *quasilinear* if f is simultaneously quasiconvex and quasiconcave.

4.2 Examples

Logarithm: $f = \log(x)$, $x \in \mathbf{R}^{++}$ is concave but quasilinear.
Square root of absolute: $f = \sqrt{|x|}$, $x \in \mathbf{R}$ is quasiconvex.
Product: $f(x_1, x_2) = x_1 x_2$, $\mathbf{dom} f = \mathbf{R}^2_+$ is quasiconcave.
Max-nonzero-indices: $f(x) = \max\{i \in [n] : x_i \neq 0\}$ is quasiconvex.

4.3 Properties

Theorem: f is quasiconvex iff $\mathbf{dom} f$ is convex and $\forall x, y \in \mathbf{dom} f, \forall \theta \in [0,1]$, $f[\theta x + (1-\theta)y] \leq \max\{f(x), f(y)\}$.
Theorem: The following statements are equivalent: (i)S_α is convex $\forall \alpha \in \mathbf{R}$ (ii)$\forall x, y \in \mathbf{dom} f, \forall \theta \in [0,1]$, $f[\theta x + (1-\theta)y] \leq \max\{f(x), f(y)\}$.
Theorem: If $f : \mathbf{R} \to \mathbf{R}$ is continuous, it is quasiconvex iff at least one of the following is true: (i)it is non-decreasing. (ii)it is non-increasing. (iii)$\exists c \in \mathbf{dom} f$ such that $\forall t \leq c$, f is non-increasing and $\forall t \geq c$, f is non-decreasing.

4.4 Differentiable quasiconvex functions

Suppose $f : \mathbf{R}^n \to \mathbf{R}$ is differentiable.
First-order condition: f is quasiconvex iff $\mathbf{dom} f$ is convex and $\forall x, y \in \mathbf{dom} f$, $f(y) < f(x) \implies \nabla f(x)^T (y - x) \leq 0$.
Second-order condition: Suppose f is twice-differentiable, then f is quasiconvex iff $\forall y \in \mathbf{R}^n, \nabla f(x)^T y = 0 \implies y^T \nabla^2 f(x) y \geq 0$.
If $f_1 \dots f_n$ are quasiconvex, and $w_i \geq 0$, then $f = \max\{w_1 f_1, \dots, w_n f_n\}$ is quasiconvex.
If f is quasiconvex in (x, y), and C is a convex set, then $g(x) = \inf_{y \in C}\{f(x, y)\}$ is quasiconvex.

5 Optimization problems

5.1 Standard form

$$\min_x f_0(x)$$

s.t.$f_i(x) \leq 0 \; \forall i = 1, \dots, m$
$h_i(x) = 0 \; \forall i = 1, \dots, p$ f_0 is the *objective* function. f_i is the i^{th} inequality constraint, and h_i is the i^{th} equality constraint. **Domain** of the problem is defined as $D = (\bigcap \mathbf{dom} f_i) \cap (\bigcap \mathbf{dom} h_i)$

5.2 Optimal value

We denote feasible set $C = \{x : f_i(x) \leq 0 \; \forall i, h_j(x) = 0 \; \forall j\}$.
We denote the optimal value $p^* = \inf\{f_0(x) : x \in C\}$. If $C = \emptyset$, then the problem is *infeasible*, we define $p^* = +\infty$. If $\exists (x_k)_{k=1}^{+\infty}$ such that x_k is feasible and $f_0(x_k) \to -\infty$ as $k \to +\infty$, then the problem is *unbounded below*, $p^* = -\infty$

5.3 Set of optimal points

Suppose $D \subseteq \mathbf{R}^n$.
$X_{opt} = \{x \in D : x \in C, f_0(x) = p^*\}$. If $X_{opt} \neq \emptyset$, then the optimal value is *attained*. If $X_{opt} = \emptyset$, (assume problem is feasible), we could distinguish between two cases: problem is unbounded below, or the infimum value exists but is not attainable. If we minimise a continuous function over a compact set, then the optimal solution is attainable. A feasible x is ϵ-optimal if $f_0(x) \leq p^* + \epsilon$.

5.4 Local optimality

x is *locally optimal* if $\exists R > 0$ such that $f_0(x) = \inf\{f_0(z) : z \in C, |z - x|_2 \leq R\}$, geometrically it means that x minimizes f_0 over its R-neighborhood. If x is feasible and $f_i(x) = 0$, we say the i^{th} inequality constraint is active, otherwise inactive.

5.5 Feasibility

We can formulate an optimization problem in a way that we are to optimize a constant value, say 0, over a family of constraints. If $p^* \neq +\infty$, we find x that satisfies all constraints.

5.6 Standardise problems

$x_i \leq u_i \to x_i - u_i \leq 0$
$x_i \geq li \to -x_i + l_i \leq 0$

5.7 Change of variables

If $\exists \phi : \mathbf{R}^n \to \mathbf{R}^n$ and is into, we define $x = \phi(z)$, and $f_i'(z) = f_i(\phi(z))$, $h_i'(z) = h_i(\phi(z))$. If x solves the original optimization problem, then z solves the transformed optimization problem.

5.8 Change of objective and constraints

Suppose $\psi_0 : \mathbf{R} \to \mathbf{R}$ is monotone increasing. $\psi_i : \mathbf{R} \to \mathbf{R}, i = 1, \dots, m$ satisfies that $\psi_i(u) \leq 0 \iff u \leq 0$, and $\varphi_j : \mathbf{R} \to \mathbf{R}, j = 1, \dots, p$ satisfies that $\varphi_j(u) = 0$ when $u = 0$, then we define $f_i'(x) = \psi_i(f_i(x))$, $i = 0, 1, \dots, m$, $h_j'(x) = \varphi_j(h_j(x))$, $j = 1, \dots, p$. This transformed problem is equivalent to the original problem.

5.9 Slack variables

We can change inequality constraints into equality constraints by introducing non-negative slack variables.
$f_i(x) \leq 0 \iff \exists \, s_i \geq 0$ such that $f_i(x) + s_i = 0$.
x is optimal for original problem $\iff (x, s)$ is optimal for transformed problem, and $s_i = -f_i(x)$.

5.10 Epigraph form

$$\min_{x,t} t$$

s.t. $t \geq f_0(x)$
$f_i(x) \leq 0 \; \forall i = 1, \dots, m$
$h_i(x) = 0 \; \forall i = 1, \dots, p$
It is equivalent to the original problem.
(x, t) is optimal for the transformed problem $\iff x$ is optimal for the original problem and $t = f_0(x)$

6 Convex optimization

6.1 Standard form

The form follows from the general optimization problem, with additional requirements:

1. f_0 must be convex(if quasiconvex, then the problem is quasiconvex optimization)
2. f_i must be convex.
3. h_i must be affine, that is, $h_i(x) = a_i^T x - b_i$

6.2 Optimality

For convex optimization problems, we have that any local minimum is a global minimum. The optimal set is convex. If the objective is strictly convex, then the optimal set contains at most one point.

6.3 Optimality condition for differentiable f_0

Theorem: Let C be the feasible set for a convex problem. x is optimal $\iff x \in C$ and $\nabla f_0(x)^T (y - x) \geq 0 \; \forall y \in C$. Geometrically, $-\nabla f_0(x)$ defines a supporting hyperplane to the feasible set C, $-\nabla f_0(x)$ makes an obtuse angle with $y - x$. If problem is unconstrained, then the optimal condition reduces to $\nabla f_0(x) = 0$.

6.4 Quadratic optimization problem

$\min_{x \in \mathbf{R}^n} \frac{1}{2} x^T P x + q^T x + r$ such that $Gx \leq h$, $Ax = b$, where $P \in S^n_+, G \in \mathbf{R}^{m*n}, A \in \mathbf{R}^{p*n}, q \in \mathbf{R}^n, h \in \mathbf{R}^m, b \in \mathbf{R}^p$
If P is positive semidefinite, the problem is convex. If P is positive definite, the problem is strictly convex.

6.5 Quadratic constrained quadratic program

$\min_x \frac{1}{2} x^T P_0 x + q_0^T x + r_0$ such that $\frac{1}{2} x^T P_i x + q_i^T x + r_i \leq 0 \; \forall i = 1, \dots, m$, $Ax = b$, where $P_i \in S^n_+ \to f_i$ are convex functions.
If we take $P_i = 0$, we recover a linear program. Therefore, LP is a subset of QCQP.

6.6 Second order cone programming

Second order cone $K = \{(x, t) \in \mathbf{R}^{n+1} : t \geq 0, |x|_2 \leq t\}$
$\min_{x \in \mathbf{R}^n} f^T x$
such that $|A_i x + b_i|_2 \leq c_i^T x + d_i, i \in [m], Fx = g$, where $A_i \in \mathbf{R}^{n_i * n}, b_i \in \mathbf{R}^{n_i}, c_i \in \mathbf{R}^n, d_i \in \mathbf{R}, F \in \mathbf{R}^{p*n}, G \in \mathbf{R}^p$.
We say that $|A_i x + b_i|_2 \leq c_i^T x + d_i$ is a second order cone constraint, since $(Ax + b, c^T x + d) \in K$.
If we take $c_i = 0$, we recover quadratic constraint. If we take $A_i = 0$, we recover affine constraint. Therefore, QCQP is a subset of SOCP.

6.7 Geometric programming

A function $f : \mathbf{R}^n \to \mathbf{R}$ with $\mathbf{dom} f = \mathbf{R}^n_{++}$ defined as $f(x) = c x_1^{a_1} \dots x_n^{a_n}$ where $c > 0, a_i \in \mathbf{R}$ is a *monomial*.
A function $f : \mathbf{R}^n \to \mathbf{R}$ with $\mathbf{dom} f = \mathbf{R}^n_{++}$ defined as $f(x) = \Sigma_{i=1}^k f_i(x)$, where f_i are monomials, is a *posynomial*.
$\min_x f_0(x)$ such that $f_i(x) \leq 1 \; \forall i \in [m], h_i(x) = 1 \; \forall i \in [p]$, where f_i are posynomials and h_i are monomials.
Note that GP is not guaranteed convex. However, consider substitution $y_i = \log(x_i) \iff x_i = e^{y_i}$ (Note that this change of variable makes sense as $x_i > 0$), then the objective function and the constraints can be rewritten as a sum of exponential. Optimise over the logarithms of the transformed objectives and inequality constraints, then we have changed the objective and inequality constraints into log-sum-exp functions, which are convex. GP thus is transformed into a convex optimization.

6.8 Generalized inequalities

Remember that a proper cone K defines a partial order.
Definition: $f : \mathbf{R}^n \to \mathbf{R}^m$ is K-convex if $\forall x, y \in \mathbf{dom} f, \forall \theta \in [0,1]$, $f(\theta x + (1 - \theta y)) \leq_K \theta f(x) + (1 - \theta) f(y)$.

6.9 Semidefinite programming

We consider an optimization problem with generalized inequality constraints.
$\min_x f_0(x)$ such that $f_i(x) \leq_{K_i} 0 \; \forall i \in [m], Ax = b$
When $K = S^n_+$, the associated conic form problem is called a semidefinite program.
$\min_x c^T x$ such that $x_1 F_1 + \dots + x_n F_n + G \leq_K 0, Ax = b$, where $G, F_i \in S^n_+, A \in \mathbf{R}^{p*n}$.
If G, F_i are diagonal, we recover an LP.
Consider the matrix $X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$, where $\det(A) \neq 0$, then the *Schur complement of A in X* $S = C - B^T A^{-1} B$.
We have Schur complement lemma: if $A > 0$, then $X \geq 0 \iff S \geq 0$.
Therefore, consider matrix constraint $\begin{bmatrix} tI & x \\ x^T & t \end{bmatrix} \geq 0$, by Schur complement lemma, it is equivalent to $t - x^T (tI)^{-1} x \geq 0 \iff t \geq t^{-1} x^T x \iff |x|_2 \leq t$, therefore we recover the SOCP inequality constraint. Therefore, SOCP is a subset of SDP.