1 Sets

1.1 Generalized inequalities

- A proper cone K is used to define the *generalized inequality* in a space A, where $K \subset A$.
- $\mathbf{x} \preceq \mathbf{y} \iff \mathbf{y} - \mathbf{x} \in K$ for $\mathbf{x}, \mathbf{y} \in A$ (generalized inequality).
- $\mathbf{x} \prec \mathbf{y} \iff \mathbf{y} - \mathbf{x} \in \operatorname{int} K$ for $\mathbf{x}, \mathbf{y} \in A$ (strict generalized inequality).
- There are two cases where K and A are understood from context and the subscript K is dropped out:
  - When $K = \mathbb{R}^n_+$ (the nonnegative orthant) and $A = \mathbb{R}^n$. In this case, $\mathbf{x} \preceq \mathbf{y}$ means that $x_i \le y_i$.
  - When $K = \mathbb{S}^n_+$ and $A = \mathbb{S}^n$, or $K = \mathbb{S}^n_{++}$ and $A = \mathbb{S}^n$, where $\mathbb{S}^n$ denotes the set of symmetric $n \times n$ matrices, $\mathbb{S}^n_+$ is the space of the positive semidefinite matrices, and $\mathbb{S}^n_{++}$ is the space of the positive definite matrices. $\mathbb{S}^n_+$ is a proper cone in $\mathbb{S}^n$ (??). In this case, the generalized inequality $\mathbf{Y} \succeq \mathbf{X}$ means that $\mathbf{Y} - \mathbf{X}$ is a positive semidefinite matrix belonging to the positive semidefinite cone $\mathbb{S}^n_+$ in the subspace of symmetric matrices $\mathbb{S}^n$. It is usual to denote $\mathbf{X} \succ \mathbf{0}$ and $\mathbf{X} \succeq \mathbf{0}$ to mean than $\mathbf{X}$ is a positive definite and semidefinite matrix, respectively, where $\mathbf{0} \in \mathbb{R}^{n \times n}$ is a zero matrix.
- Another common usage is when $K = \left\{\mathbf{c} \in \mathbb{R}^n \mid c_1 + c_2 t + \cdots + c_n t^{n-1} \ge 0, \text{ for } 0 \le t \le 1\right\}$ and $A = \mathbb{R}^n$. In this case, $\mathbf{x} \preceq_K \mathbf{y}$ means that $x_1 + x_2 t + \cdots + x_n t^{n-1} \le y_1 + y_2 t + \cdots + y_n t^{n-1}$.
- The generalized inequality has the following properties:
  - If $\mathbf{x} \preceq_K \mathbf{y}$ and $\mathbf{u} \preceq_K \mathbf{v}$, then $\mathbf{x} + \mathbf{u} \preceq_k \mathbf{y} + \mathbf{v}$ (preserve under addition).
  - If $\mathbf{x} \preceq_K \mathbf{y}$ and $\mathbf{y} \preceq_K \mathbf{z}$, then $\mathbf{x} \preceq_K \mathbf{z}$ (transitivity).
  - If $\mathbf{x} \preceq_K \mathbf{y}$, then $\alpha\mathbf{x} \preceq_K \mathbf{y}$ for $\alpha \ge 0$ (preserve under nonnegative scaling).
  - $\mathbf{x} \preceq_K \mathbf{x}$ (reflexivity).
  - If $\mathbf{x} \preceq_K \mathbf{y}$ and $\mathbf{y} \preceq_K \mathbf{x}$, then $\mathbf{x} = \mathbf{y}$ (antisymmetric).
  - If $\mathbf{x}_i \preceq_K \mathbf{y}_i$, for $i = 1, 2, \ldots$, and $\mathbf{x}_i \to \mathbf{x}$ and $\mathbf{y}_i \to \mathbf{y}$ as $i \to \infty$, then $\mathbf{x} \preceq_K \mathbf{y}$.
- It is called partial ordering because $\mathbf{x} \npreceq_K \mathbf{y}$ and $\mathbf{y} \npreceq_K \mathbf{x}$ for many $\mathbf{x}, \mathbf{y} \in A$. When it happens, we say that $\mathbf{x}$ and $\mathbf{y}$ are not comparable (this case does not happen in ordinary inequality, $<$ and $>$).

1.2 Minimum (maximum)

- The minimum (maximum) element of a set S is always defined with respect to the proper cone K.
- $\mathbf{x} \in S$ is the *minimum* element of the set S with respect to the proper cone K if $\mathbf{x} \preceq_K \mathbf{y}$, $\forall\, \mathbf{y} \in S$ (for *maximum*, $\mathbf{x} \succeq_K \mathbf{y}$, $\forall\, \mathbf{y} \in S$).
- It means that $S \subseteq \mathbf{x} + K$ (for the maximum, $S \subseteq \mathbf{x} - K$), where $\mathbf{x} + K$ denotes the set K shifted from the origin by $\mathbf{x}$. Note that any point in $K + \mathbf{x}$ is comparable with $\mathbf{x}$ and is greater or equal to $\mathbf{x}$ in the generalized inequality sense.
- The set S does not necessarily have a minimum (maximum), but the minimum (maximum) is unique if it does.

1.3 Minimal (maximal)

- The minimal (maximal) element of a set S is always defined with respect to the proper cone K.
- $\mathbf{x} \in S$ is the *minimal* element of S with respect to the proper cone K if $\mathbf{y} \preceq_K \mathbf{x}$ only when $\mathbf{y} = \mathbf{x}$ (for the *maximal*, $\mathbf{y} \succeq_K \mathbf{x}$ only when $\mathbf{y} = \mathbf{x}$).
- It means that $(\mathbf{x} - K) \cap S = \{\mathbf{x}\}$ for minimal (for the maximal $(\mathbf{x} + K) \cap S = \{\mathbf{x}\}$), where $\mathbf{x} - K$ denotes the reflected set K shift by $\mathbf{x}$.
- Any point in $\mathbf{x} - K$ is comparable with $\mathbf{x}$ and is less than or equal to $\mathbf{x}$ in the generalized inequality mean.
- The set S can have many minimal (maximal) elements.

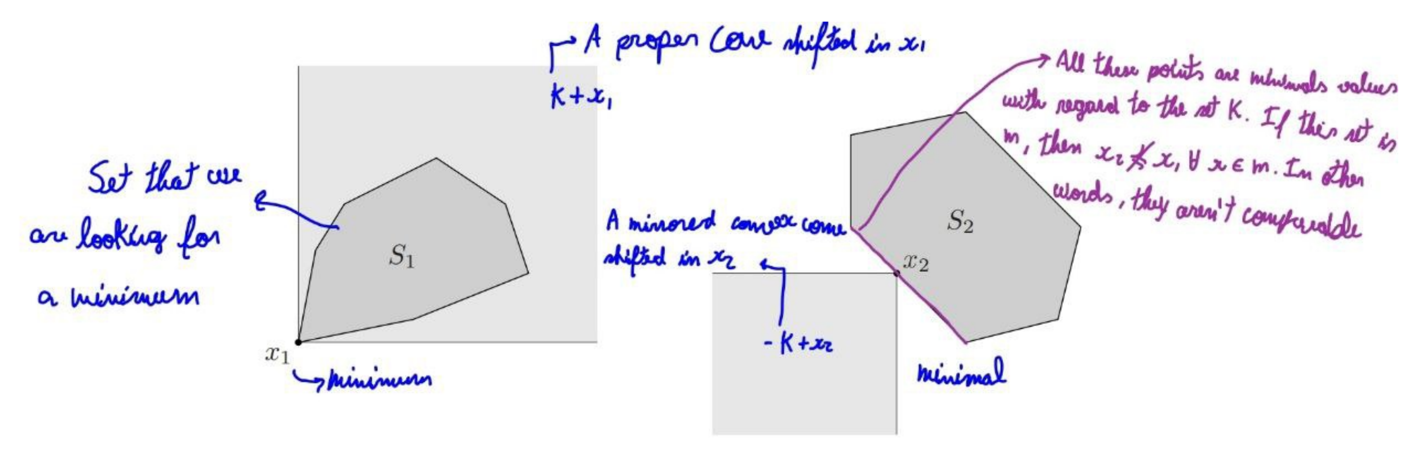


Figure 2.17 *Left.* The set S_1 has a minimum element x_1 with respect to componentwise inequality in $\mathbf{R}^2$. The set $x_1 + K$ is shaded lightly; x_1 is the minimum element of S_1 since $S_1 \subseteq x_1 + K$. *Right.* The point x_2 is a minimal point of S_2. The set $x_2 - K$ is shown lightly shaded. The point x_2 is minimal because $x_2 - K$ and S_2 intersect only at x_2.

1.4 Table of the known sets

Convex sets	
Set	**Comments**
Convex hull: • $\operatorname{conv} C = \left\{\sum_{i=1}^k \theta_i \mathbf{x}_i \mid \mathbf{x}_i \in C \text{ for } i = 1, \cdots, k, \mathbf{0} \preceq \boldsymbol{\theta} \preceq \mathbf{1}, \mathbf{1}^\top \boldsymbol{\theta} = 1\right\}$	• $\operatorname{conv} C$ is the smallest convex set that contains C. • $\operatorname{conv} C$ is a finite set as long as C is also finite.
Affine hull: • $\operatorname{aff} C = \left\{\sum_{i=1}^k \theta_i \mathbf{x}_i \mid \mathbf{x}_i \in C \text{ for } i = 1, \cdots, k, \mathbf{1}^\top \boldsymbol{\theta} = 1\right\}$	• $\operatorname{aff} C$ is the smallest affine set that contains C. • $\operatorname{aff} C$ is always an infinite set. If $\operatorname{aff} C$ contains the origin, it is also a subspace. • Different from the convex set, θ_i is not restricted between 0 and 1
Conic hull: • $A = \left\{\sum_{i=1}^k \theta_i \mathbf{x}_i \mid \mathbf{x}_i \in C, \theta_i \ge 0 \text{ for } i = 1, \cdots, k\right\}$	• A is the smallest convex conic that contains C. • Different from the convex and affine sets, θ_i does not need to sum up 1.
Ray: • $\mathcal{R} = \{\mathbf{x}_0 + \theta\mathbf{v} \mid \theta \ge 0\}$	• The ray is an infinite set that begins in $\mathbf{x}_0$ and extends infinitely in direction of $\mathbf{v}$. In other words, it has a beginning, but it has no end. • The ray becomes a convex cone if $\mathbf{x}_0 = \mathbf{0}$.
Hyperplane: • $\mathcal{H} = \left\{\mathbf{x} \mid \mathbf{a}^\top \mathbf{x} = b\right\}$ • $\mathcal{H} = \left\{\mathbf{x} \mid \mathbf{a}^\top (\mathbf{x} - \mathbf{x}_0) = \mathbf{0}\right\}$ • $\mathcal{H} = \mathbf{x}_0 + a^\perp$	• It is an infinite set $\mathbb{R}^{n-1} \subset \mathbb{R}^n$ that divides the space into two halfspaces. • The inner product between $\mathbf{a}$ and any vector in $\mathcal{H}$ yields the constant value b. • $a^\perp = \left\{\mathbf{v} \mid \mathbf{a}^\top \mathbf{v} = 0\right\}$ is the infinite set of vectors perpendicular to $\mathbf{a}$. It passes through the origin. • $a^\perp$ is offset from the origin by $\mathbf{x}_0$, which is any vector in $\mathcal{H}$.
Halfspaces: • $\mathcal{H}_- = \left\{\mathbf{x} \mid \mathbf{a}^\top \mathbf{x} \le b\right\}$ • $\mathcal{H}_+ = \left\{\mathbf{x} \mid \mathbf{a}^\top \mathbf{x} \ge b\right\}$	• They are infinite sets of the parts divided by $\mathcal{H}$.
Euclidean ball: • $B(\mathbf{x}_c, r) = \{\mathbf{x} \mid \\|\mathbf{x} - \mathbf{x}_c\\| \le r\}$ • $B(\mathbf{x}_c, r) = \left\{\mathbf{x} \mid (\mathbf{x} - \mathbf{x}_c)^\top (\mathbf{x} - \mathbf{x}_c) \le r^2\right\}$ • $B(\mathbf{x}_c, r) = \{\mathbf{x}_c + r\\|\mathbf{u}\\| \mid \\|\mathbf{u}\\| \le 1\}$	• $B(\mathbf{x}_c, r)$ is a finite set as long as $r < \infty$. • $\mathbf{x}_c$ is the center of the ball. • r is its radius.
Ellipsoid: • $\mathcal{E} = \left\{\mathbf{x} \mid (\mathbf{x} - \mathbf{x}_c)^\top \mathbf{P}^{-1} (\mathbf{x} - \mathbf{x}_c) \le 1\right\}$ • $\mathcal{E} = \left\{\mathbf{x}_c + \mathbf{P}^{1/2}\mathbf{u} \mid \\|\mathbf{u}\\| \le 1\right\}$	• $\mathcal{E}$ is a finite set as long as $\mathbf{P}$ is a finite matrix. • $\mathbf{P}$ is symmetric and positive definite, that is, $\mathbf{P} = \mathbf{P}^\top \succ \mathbf{0}$. It determines how far the ellipsoid extends in every direction from $\mathbf{x}_c$. • $\mathbf{x}_c$ is the center of the ellipsoid. • The lengths of the semi-axes are given by $\sqrt{\lambda_i}$. • When $\mathbf{P}^{1/2} \succeq \mathbf{0}$ but singular, we say that $\mathcal{E}$ is a degenerated ellipsoid (degenerated ellipsoids are also convex).
Norm cone: • $C = \left\{(x_1, x_2, \cdots, x_n, t) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathbb{R}^n, \\|\mathbf{x}\\|_p \le t\right\} \subseteq \mathbb{R}^{n+1}$	• Although it is named "Norm cone", it is a set, not a scalar. • The cone norm increases the dimension of $\mathbf{x}$ in 1. • For $p = 2$, it is called the second-order cone, quadratic cone, Lorentz cone or ice-cream cone.
Proper cone: $K \subset \mathbb{R}^n$ is a proper cone when it has the following properties • K is a convex cone, i.e., $\alpha K \equiv K, \alpha > 0$. • K is closed. • K is solid. • K is pointed, i.e., $-K \cap K = \{\mathbf{0}\}$.	• When $K = \mathbb{R}_+$ and $S = \mathbb{R}$ (ordinary inequality), the minimal is equal to the minimum and the maximal is equal to the maximum. • When we say that a scalar-valued function $f : \mathbb{R}^n \to \mathbb{R}$ is nondecreasing, it means that whenever $\mathbf{u} \preceq \mathbf{v}$, we have $\tilde{h}(\mathbf{u}) \le \tilde{h}(\mathbf{v})$. Similar results hold for decreasing, increasing, and nonincreasing scalar functions.
Subspace (cone set?) of the symmetric matrices: • $\mathbb{S}^n = \left\{\mathbf{X} \in \mathbb{R}^{n \times n} \mid \mathbf{X} = \mathbf{X}^\top\right\}$	• The positive semidefinite cone is given by $\mathbb{S}^n_+ = \{\mathbf{X} \in \mathbb{R}^{n \times n} \mid \mathbf{X} \succeq \mathbf{0}\} \subset \mathbb{S}^n$. This is the proper cone used to define the generalized inequalities between matrices, e.g., $\mathbf{A} \preceq \mathbf{B}$. • The positive definite cone is given by $\mathbb{S}^n_{++} = \{\mathbf{X} \in \mathbb{R}^{n \times n} \mid \mathbf{X} \succ \mathbf{0}\} \subset \mathbb{S}^n_+$. This is the proper cone used to define the generalized inequalities between matrices, e.g., $\mathbf{A} \prec \mathbf{B}$.
Dual cone: • $K^* = \left\{\mathbf{y} \mid \mathbf{x}^\top \mathbf{y} \ge 0, \ \forall\, \mathbf{x} \in K\right\}$	• K^* is a cone, and it is convex even when the original cone K is nonconvex. • K^* has the following properties: ▹ K^* is closed and convex. ▹ $K_1 \subseteq K_2$ implies $K_1^* \subseteq K_2^*$. ▹ If K has a nonempty interior, then K^* is pointed. ▹ If the closure of K is pointed then K^* has a nonempty interior. ▹ K^{**} is the closure of the convex hull of K. Hence, if K is convex and closed, $K^{**} = K$.
Polyhedra: • $\mathcal{P} = \left\{\mathbf{x} \mid \mathbf{a}_j^\top \mathbf{x} \le b_j, j = 1, \ldots, m, \mathbf{a}_j^\top \mathbf{x} = d_j, j = 1, \cdots, p\right\}$ • $\mathcal{P} = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} \preceq \mathbf{b}, \mathbf{C}\mathbf{x} = \mathbf{d}\}$, where $\mathbf{A} = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_m\end{bmatrix}^\top$ and $\mathbf{C} = \begin{bmatrix}\mathbf{c}_1 & \mathbf{c}_2 & \ldots & \mathbf{c}_m\end{bmatrix}^\top$	• The polyhedron may or may not be an infinite set. • Polyhedron is the result of the intersection of m halfspaces and p hyperplanes. • Subspaces, hyperplanes, lines, rays line segments, and halfspaces are all special cases of polyhedra. • The *nonnegative orthant*, $\mathbb{R}^n_+ = \{\mathbf{x} \in \mathbb{R}^n \mid x_i \le 0 \text{ for } i = 1, \ldots n\} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{I}\mathbf{x} \succeq \mathbf{0}\}$, is a special polyhedron.
Simplex: • $S = \operatorname{conv}\ \{\mathbf{v}_m\}_{m=0}^k = \left\{\sum_{i=0}^k \theta_i \mathbf{v}_i \mid \mathbf{0} \preceq \boldsymbol{\theta} \preceq \mathbf{1}, \mathbf{1}^\top \boldsymbol{\theta} = 1\right\}$ • $S = \{\mathbf{x} \mid \mathbf{x} = \mathbf{v}_0 + \mathbf{V}\boldsymbol{\theta}\}$, where $\mathbf{V} = \begin{bmatrix}\mathbf{v}_1 - \mathbf{v}_0 & \ldots & \mathbf{v}_n - \mathbf{v}_0\end{bmatrix} \in \mathbb{R}^{n \times k}$ • $S = \{\mathbf{x} \mid \underbrace{\mathbf{A}_1\mathbf{x} \preceq \mathbf{A}_1\mathbf{v}_0, \mathbf{1}^\top \mathbf{A}_1 \mathbf{x} \le 1 + \mathbf{1}^\top \mathbf{A}_1 \mathbf{v}_0}_{\text{Linear inequalities in } x}, \underbrace{\mathbf{A}_2\mathbf{x} = \mathbf{A}_2\mathbf{v}_0}_{\text{Linear equalities in } x}\}$ (Polyhedra form), where $\mathbf{A} = \begin{bmatrix}\mathbf{A}_1 \\ \mathbf{A}_2\end{bmatrix}$ and $\mathbf{AV} = \begin{bmatrix}\mathbf{I}_{k \times k} \\ \mathbf{0}_{n-k \times n-k}\end{bmatrix}$	• Simplexes are a subfamily of the polyhedra set. • Also called k-dimensional Simplex in $\mathbb{R}^n$. • The set $\{\mathbf{v}_m\}_{m=0}^k$ is a affinely independent, which means $\{\mathbf{v}_1 - \mathbf{v}_0, \ldots, \mathbf{v}_k - \mathbf{v}_0\}$ are linearly independent. • $\mathbf{V} \in \mathbb{R}^{n \times k}$ is a full-rank tall matrix, i.e., $\operatorname{rank}(\mathbf{V}) = k$. All its column vectors are independent. The matrix $\mathbf{A}$ is its left pseudoinverse.
α-sublevel set: • $C_\alpha = \{\mathbf{x} \in \operatorname{dom}(f) \mid f(\mathbf{x}) \le \alpha\}$ (regarding convexity), where $f : \mathbb{R}^n \to \mathbb{R}$ • $C_\alpha = \{\mathbf{x} \in \operatorname{dom}(f) \mid f(\mathbf{x}) \ge \alpha\}$ (regarding concavity), where $f : \mathbb{R}^n \to \mathbb{R}$	• If f is a convex (concave) function, then sublevel sets of f are convexes (concaves) for any $\alpha \in \mathbb{R}$. • The converse is not true: a function can have all its sublevel set convex and not be a convex function. • $C_\alpha \subseteq \operatorname{dom}(f)$

1.5 Operations on set and their implications regarding curvature

Operation	Curvature
Union $C = A \cup B$ • $C = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \in A \text{ or } \mathbf{x} \in B\}$.	It is neither convex nor concave in most of the cases
Intersection: $C = A \cap B$ • $C = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \in A, \mathbf{x} \in B\}$.	It is convex (concave) as long as A and B are convexes (concaves)
Minkowski sum: $C = A + B$ • $C = \{\mathbf{x} + \mathbf{y} \in \mathbb{R}^n \mid \mathbf{x} \in A, \mathbf{y} \in B\}$.	It is convex (concave) as long as A and B are convexes (concaves)
Offset: $C = A + k$ • $C = \{\mathbf{x} + k \in \mathbb{R}^n \mid \mathbf{x} \in A, k \in \mathbb{R}\}$.	It is convex (concave) as long as A and B are convexes (concaves)
Cartesian product: $C = A \times B$ • $C = \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \in A, \mathbf{y} \in \mathbb{B}\}$.	It is convex (concave) as long as A and B are convexes (concaves)

2 Fuctions

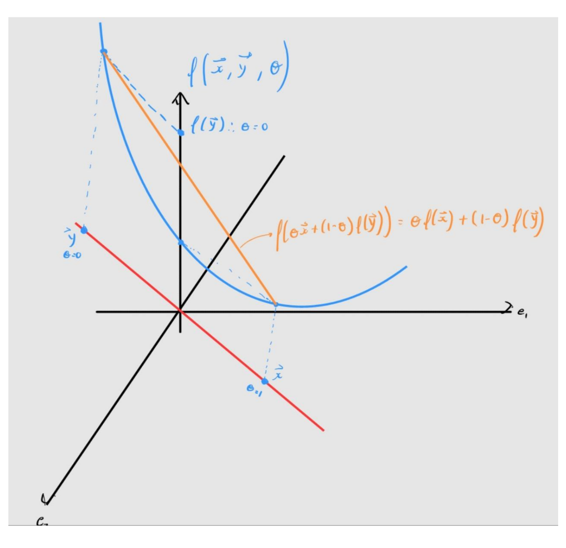
2.1 Categories of functions regarding its curvature for CVX

- On CVX, for functions with multiple arguments (a vector as input), the curvature categories are always considered jointly [5].
- The CVX optimization package, and, apparently, its derivatives (CVXPY, Convex.jl, CVXR...) categorize the functions as follows [5]:

2.1.1 Convex

$$f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \le \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}),\ \forall\, \mathbf{x}, \mathbf{y} \in \operatorname{dom}(f), 0 \le \theta \le 1 \tag{1}$$

- $f : \operatorname{dom}(f) \to \mathbb{R}$, where $\operatorname{dom}(f) \subseteq \mathbb{R}^n$.
- The Eq.(1) implies that $\operatorname{dom}(f)$ is a convex set, that is, all points for any line segment within $\operatorname{dom}(f)$ belong to it.
- The Eq.(1) implies that any line segment within $\operatorname{dom}(f)$ gives a convex graph (bowl-shaped).
- Graphically, any line segment between $(\mathbf{x}, f(\mathbf{x}))$ and $(\mathbf{y}, f(\mathbf{y}))$ lies always above the graph f. If the line touches the graph but does not cross it, then the function is strictly convex.

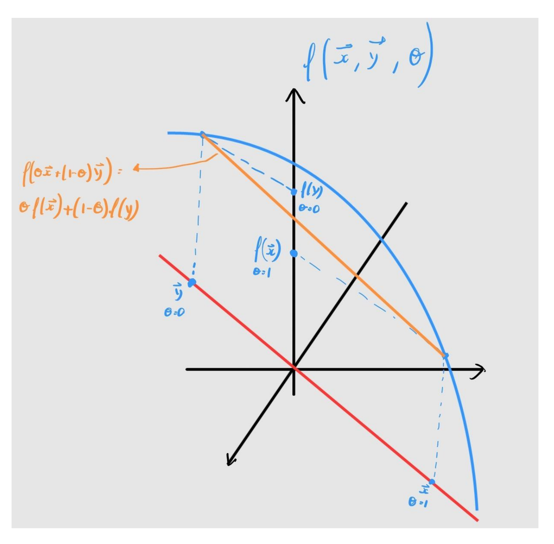


- It is guaranteed that $\exists!\, \mathbf{x}^\star \in \mathbb{R}^n \mid f(\mathbf{x}^\star) \le f(\mathbf{y})\ \forall\, \mathbf{y} \in \operatorname{dom}(f)$, and $\nabla f(\mathbf{y}) = \mathbf{0}$ iff $\mathbf{y} = \mathbf{x}^\star$. This $\mathbf{x}^\star$ is the global minimum.
- If f is (strictly convex) convex, then $-f$ is (strictly concave) concave.

2.1.2 Concave

$$f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \ge \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}),\ \forall\, \mathbf{x}, \mathbf{y} \in \operatorname{dom}(f), 0 \le \theta \le 1 \tag{2}$$

- $f : \operatorname{dom}(f) \to \mathbb{R}$, where $\operatorname{dom}(f) \subseteq \mathbb{R}^n$.
- The Eq.(2) implies that $\operatorname{dom}(f)$ is a convex set, that is, all points for any line segment within $\operatorname{dom}(f)$ belong to it.
- The Eq.(2) implies that any line segment within $\operatorname{dom}(f)$ gives a concave graph (hyperhyperbola-shaped).
- Graphically, any line segment between $(\mathbf{x}, f(\mathbf{x}))$ and $(\mathbf{y}, f(\mathbf{y}))$ lies always below the graph f. If the line touches the graph but does not cross it, then the function is strictly concave.

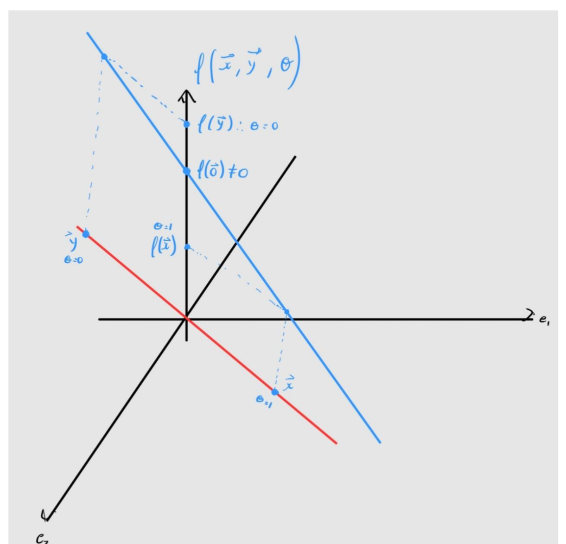


- It is guaranteed that $\exists!\, \mathbf{x}^\star \in \mathbb{R}^n \mid f(\mathbf{x}^\star) \le f(\mathbf{y})\ \forall\, \mathbf{y} \in \operatorname{dom}(f)$, and $\nabla f(\mathbf{y}) = \mathbf{0}$ iff $\mathbf{y} = \mathbf{x}^\star$. This $\mathbf{x}^\star$ is the global maximum.
- If f is (strictly concave) concave, then $-f$ is (strictly convex) convex.

2.1.3 Affine

$$f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) = \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}),\ \forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \theta \in \mathbb{R} \tag{3}$$

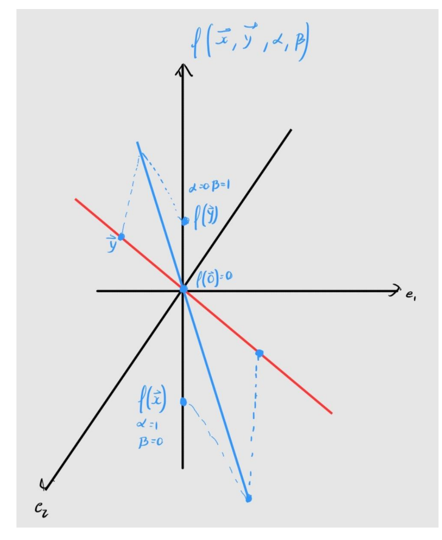
- $f : \mathbb{R}^n \to \mathbb{R} \therefore \operatorname{dom}(f) = \mathbb{R}^n$.
- $\operatorname{dom}(f)$ must be infinite since θ is not restricted to an interval.
- The affine function has the following characteristic



- For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, f yields a line with the variation of θ.
- The affine function is a broader category that encompasses the class of linear functions. The main difference is that linear functions must have its origin fixed after the transformation, whereas affine functions do not necessarily have it (when not, this makes the affine function nonlinear). Mathematically, the linear function shall obey the following relation

$$f(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y}),\ \forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \alpha, \beta \in \mathbb{R}. \tag{4}$$

  When $\alpha = \beta = 0$, $f(\mathbf{0}) = 0$. It leads to the following graph



- We can think of an affine function as a linear transformation plus a shift from the origin.
- Affine functions are both convex and concave.

2.1.4 Constant

$$f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) = k,\ \forall\, \mathbf{x}, \mathbf{y} \in \operatorname{dom}(f), \theta \in \mathbb{R} \tag{5}$$

- $f : \mathbb{R}^n \to \mathbb{R} \therefore \operatorname{dom}(f) = \mathbb{R}^n$.
- $\operatorname{dom}(f)$ must be infinite since θ is not restricted to an interval.
- $k \in \mathbb{R}$ is a constant.
- It is a special case of affine function.
- A constant function is convex and concave, simultaneously.

2.1.5 Unkown

- Nonconvex and nonconcave functions do not satisfy the convexity or concavity rule and are categorized as unknown curvature.
- Possibly, convex and/or concave functions can also be categorized as unknown if it does not follow the DCP ruleset.

2.2 Categories of functions regarding its optimization variables

$\mathbf{x} \in \mathbb{R}^n$	Continuous optimization
$\mathbf{x} \in \mathbb{Z}^n$	Integer optimization
$x_1, x_2, \ldots, x_k \in \mathbb{R}$ and $x_{k+1}, \ldots, x_n \in \mathbb{Z}$	Mixed-optimization

2.3 Table of known functions

Functions and their implications regarding curvatuve

Function	Curvature and monoticity	Comments
Matrix functions $f : \mathbb{R}^n \to \mathbb{R}^m$ • $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{b} \in \mathbb{R}^m, \mathbf{x} \in \mathbb{R}^n$	• Affine. • If $\mathbf{b} = \mathbf{0}$, then $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$ is a linear function.	• A special case of the linear function is when $\mathbf{A} = \mathbf{c}^\mathsf{T}$. In this case, we have $f(\mathbf{x}) = \mathbf{c}^\mathsf{T}\mathbf{x}$, which is the inner product between the vector $\mathbf{c}$ and $\mathbf{x}$. • The inverse image of C, $f^{-1}(C) = \{\mathbf{x} \mid f(\mathbf{x}) \in C\}$, is also convex. • The *linear matrix inequality* (LMI), $\mathbf{A}(\mathbf{x}) = x_1\mathbf{A}_1 + \cdots + x_n\mathbf{A}_n \preceq \mathbf{B}$, is a special case of sums of matrix functions. In other words, $f(S) = \{\mathbf{x} \mid \mathbf{A}(\mathbf{x}) \preceq \mathbf{B}\}$ is a convex set if S is convex. Many optimization problems can be formulated as LMI problems and solved optimally.
Exponential function $f : \mathbb{R} \to \mathbb{R}$ • $f(x) = e^{ax} \in \mathbb{R}$, where $a \in \mathbb{R}$	Convex.	
Quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ • $f(\mathbf{x}) = a\mathbf{x}^\mathsf{T}\mathbf{P}\mathbf{x} + \mathbf{p}^\mathsf{T}\mathbf{x} + r \in \mathbb{R}$, where $\mathbf{x}, \mathbf{p} \in \mathbb{R}^n, \mathbf{P} \in \mathbb{R}^{n\times n}$, and $a, b \in \mathbb{R}$	It depends on the matrix $\mathbf{P}$: • f is convex iff $\mathbf{P} \succeq \mathbf{0}$. • f is strictly convex iff $\mathbf{P} \succ \mathbf{0}$. • f is concave iff $\mathbf{P} \preceq \mathbf{0}$. • f is strictly concave iff $\mathbf{P} \prec \mathbf{0}$.	
Power function $f : \mathbb{R}_{++} \to \mathbb{R}$ • $f(x) = x^a$	It depends on a • f is convex iff $a \ge 1$ or $a \le 0$. • f is concave iff $0 \le a \le 1$.	• Note that it is guaranteed to be convex or concave iff the base power is solely x. For instance, $(x+1)^2$ is convex, but $(x-1)^2$ is nonconvex and nonconcave.
Power of absolute value: $f : \mathbb{R} \to \mathbb{R}$ • $f(x) = \lvert x\rvert^p$, where $p \le 1$.	Convex.	
Minkowski distance, p-norm function, or l_p norm function: $f : \mathbb{R}^n \to \mathbb{R}$ • $f(\mathbf{x}) = \lVert\mathbf{x}\rVert_p$, where $p \in \mathbb{N}_{++}$.	Convex.	• It can be proved by triangular inequality.
Maximum element: $f : \mathbb{R}^n \to \mathbb{R}$ • $f(\mathbf{x}) = \max\{x_1, \ldots, x_n\}$.	Convex.	
Maximum function (pointwise maximum): $f : \mathbb{R}^n \to \mathbb{R}$ • $f(\mathbf{x}) = \max\{f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})\}$.	f is convex (concave) if $f_1, \ldots, f_n$ are convex (concave) functions.	• Its domain $\operatorname{dom}(f) = \bigcap_{i=1}^{n} \operatorname{dom}(f_i)$ is also convex.
Minimum function (pointwise minimum): $f : \mathbb{R}^n \to \mathbb{R}$ • $f(\mathbf{x}) = \min\{f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})\}$.	Nonconvex and nonconcave in most of the cases.	
Pointwise infimum: • $f(\mathbf{x}) = \inf\limits_{\mathbf{y}\in\mathcal{A}} g(\mathbf{x}, \mathbf{y})$.	f is concave if g is concave for each $\mathbf{y} \in \mathcal{A}$.	• For each value of x, we have an infinite set of points $g(x, y)\rvert_{y\in\mathcal{A}}$. The value $f(x)$ will be the greatest value in the codomain of f that is less than or equal this set. • $\operatorname{dom}(f) = \left\{x \mid (x, y) \in \operatorname{dom}(g)\ \forall\, y \in \mathcal{A}, \inf\limits_{y\in\mathcal{A}} g(x, y) > -\infty\right\}$.
Pointwise supremum: • $f(\mathbf{x}) = \sup\limits_{\mathbf{y}\in\mathcal{A}} g(\mathbf{x}, \mathbf{y})$.	f is convex if g is convex for each $\mathbf{y} \in \mathcal{A}$.	• For each value of x, we have an infinite set of points $g(x, y)\rvert_{y\in\mathcal{A}}$. The value $f(x)$ will be the least value in the codomain of f that is greater than or equal this set. • $\operatorname{dom}(f) = \left\{x \mid (x, y) \in \operatorname{dom}(g)\ \forall\, y \in \mathcal{A}, \sup\limits_{y\in\mathcal{A}} g(x, y) < \infty\right\}$. • In terms of epigraphs, the pointwise supremum of the infinite set of functions $g(x, y)\rvert_{y\in\mathcal{A}}$ corresponds to the intersection of the following epigraphs: $\operatorname{epi} f = \bigcap\limits_{y\in\mathcal{A}} \operatorname{epi} g(\cdot, y)$
Logarithm function: $f : \mathbb{R}_{++} \to \mathbb{R}$ • $f(x) = \log x$	Concave and nondecreasing.	
Negative entropy function: $f : \mathbb{R}_{+} \to \mathbb{R}$ • $f(x) = x\log x$	Convex.	• When it is defined $f(x)\rvert_{x=0} = 0$, $\operatorname{dom}(f) = \mathbb{R}$.
Log-sum-exp function: $f : \mathbb{R}^n \to \mathbb{R}$ • $f(\mathbf{x}) = \log(e^{x_1} + \cdots + e^{x_n})$	Convex.	• This function is interpreted as the approximation of the maximum element function, since $\max\{x_1, \ldots, x_n\} \le f(\mathbf{x}) \le \max\{x_1, \ldots, x_n\} + \log n$
Geometric mean function $f : \mathbb{R}^n \to \mathbb{R}$ • $f(\mathbf{x}) = \left(\Pi_{i=1}^n x_i\right)^{1/n}$	Convex.	
Log-determinant function $f : \mathbb{S}^n_{++} \to \mathbb{R}$ • $f(\mathbf{X}) = \log\lvert\mathbf{X}\rvert$	Convex.	• $\mathbf{X}$ is positive semidefinite, i.e., $\mathbf{X} \succ \mathbf{0} \therefore \mathbf{X} \in \mathbb{S}^n_{++}$.
Composite function $f = h \circ g : \mathbb{R}^n \to \mathbb{R}$ • $f = g \circ h$, i.e., $f(\mathbf{x}) = (h \circ g)(\mathbf{x}) = h(g(\mathbf{x}))$, where: ▹ $g : \mathbb{R}^n \to \mathbb{R}^k$. ▹ $h : \mathbb{R}^k \to \mathbb{R}$. ▹ $\operatorname{dom}(f) = \{\mathbf{x} \in \operatorname{dom}(g) \mid g(\mathbf{x}) \in \operatorname{dom}(h)\}$.	• Scalar composition: the following statements hold for $k = 1$ and $n \ge 1$, i.e., $h : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$: ▹ f is convex if h is convex, $\tilde{h}$ is nondecreasing, and g is convex. In this case, $\operatorname{dom}(h)$ is either $(-\infty, a]$ or $(-\infty, a)$. ▹ f is convex if h is convex, $\tilde{h}$ is nonincreasing, and g is concave. In this case, $\operatorname{dom}(h)$ is either $[a, \infty)$ or (a, ∞). ▹ f is concave if h is concave, $\tilde{h}$ is nondecreasing, and g is concave. ▹ f is concave if h is concave, $\tilde{h}$ is nonincreasing, and g is convex. • Vector composition: the following statements hold for $k \ge 1$ and $n \ge 1$, i.e., $h : \mathbb{R}^k \to \mathbb{R}$ and $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^k$. Hence, $g(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_k(\mathbf{x}))$ is a vector-valued function (or simply, vector function), where $g_i : \mathbb{R}^k \to \mathbb{R}$ for $1 \le i \le k$. ▹ f is convex if h is is convex, $\tilde{h}$ is nondecreasing in each argument of $\mathbf{x}$, and $\{g_i\}_{i=1}^k$ is a set of convex functions. ▹ f is convex if h is is convex, $\tilde{h}$ is nonincreasing in each argument of $\mathbf{x}$, and $\{g_i\}_{i=1}^k$ is a set of concave functions. ▹ f is concave if h is is concave, $\tilde{h}$ is nondecreasing in each argument of $\mathbf{x}$, and $\{g_i\}_{i=1}^k$ is a set of concave functions. Where $\tilde{h}$ is the extended-value extension of the function h, which assigns the value ∞ $(-\infty)$ to the point not in $\operatorname{dom}(h)$ for h convex (concave).	• The composition function allows us to see a large class of functions as convex (or concave). • For scale composition, the remarkable ones are: ▹ If g is convex then $f(x) = h(g(\mathbf{x})) = \exp g(\mathbf{x})$ is convex. ▹ If g is concave and $\operatorname{dom}(g) \subseteq \mathbb{R}_{++}$, then $f(\mathbf{x}) = h(g(\mathbf{x})) = \log g(\mathbf{x})$ is concave. ▹ If g is concave and $\operatorname{dom}(g) \subseteq \mathbb{R}_{++}$, then $f(\mathbf{x}) = h(g(\mathbf{x})) = 1/g(\mathbf{x})$ is convex. ▹ If g is convex and $\operatorname{dom}(g) \subseteq \mathbb{R}_{+}$, then $f(\mathbf{x}) = h(g(\mathbf{x})) = g^p(\mathbf{x})$ is convex, where $p \ge 1$. ▹ If g is convex then $f(\mathbf{x}) = h(g(\mathbf{x})) = -\log(-g(x))$ is convex, where $\operatorname{dom}(f) = \{\mathbf{x} \mid g(\mathbf{x}) < 0\}$. • For vector composition, we have the following examples: ▹ If g is an affine function, then $f = h \circ g$ is convex (concave) if h is convex (concave). ▹ Let $h(\mathbf{x}) = x_{[1]} + \cdots + x_{[r]}$ be the sum of the r largest components of $\mathbf{x} \in \mathbb{R}^k$. If $g_1, g_2, \ldots, g_k$ are convex, where $\operatorname{dom}(g_i) = \mathbb{R}^n$, then $f = h \circ g$, which is the pointwise sum of the largest g_i's, is convex. ▹ $f = h \circ g$ is a convex function when $h(\mathbf{x}) = \log\left(\sum_{i=1}^k e^{x_i}\right)$ and $g_1, g_2, \ldots, g_k$ are convex functions. ▹ For $0 < p \le 1$, the function $h(\mathbf{x}) = \left(\sum_{i=1}^k x_i^p\right)^{1/p}$, where $\operatorname{dom}(h) = \mathbb{R}^n_+$, is concave. If $g_1, g_2, \ldots, g_k$ are concaves (convexes) and nonnegatives, then $f = h \circ g$ is concave (convex).
Nonnegative weighted sum: $f : \operatorname{dom}(f) \to \mathbb{R}$ • $f(\mathbf{x}) = \sum_{i=1}^m w_i f_i(\mathbf{x})$, where $w \ge 0$.	• If $f_1, f_2, \ldots, f_m$ are convex or concave functions, then f is a convex or concave function, respectively. • If $f_1, f_2, \ldots, f_m$ are strictly convex or concave functions, then f is a strictly convex or concave function, respectively.	• Special cases are when – $f = wf$ (a nonnegative scaling) – $f = f_1 + f_2$ (sum).
Addition/subtraction by a constant: $f : \operatorname{dom}(f) \to \mathbb{R}$: • $f(\mathbf{x}) = g(\mathbf{x}) + k$, where $k \in \mathbb{R}$ is a constant and $g : \mathbb{R}^n \to \mathbb{R}$.	• If g is convex (concave), then f is convex (concave)	
Integral function $f : \mathbb{R}^n \to \mathbb{R}$: • $f(\mathbf{x}) = \int_{\mathcal{A}} w(\mathbf{y}) g(\mathbf{x}, \mathbf{y})\, d\mathbf{y}$, where $\mathbf{y} \in \mathcal{A} \subseteq \mathbb{R}^m$, and $w : \mathbb{R}^m \to \mathbb{R}$.	If g is convex in $\mathbf{x}$ for each $\mathbf{y} \in \mathcal{A}$ and if $w(\mathbf{y}) \ge 0$, $\forall\, \mathbf{y} \in \mathcal{A}$, then f is convex (provided the integral exists).	
Perspective function $f : \mathbb{R}^n \times \mathbb{R}_{++} \to \mathbb{R}^n$ • $f(\mathbf{x}, t) = \mathbf{x}/t$, where $\mathbf{x} \in \mathbb{R}^n, t \in \mathbb{R}$.	Yes, if $S \subseteq \operatorname{dom}(f)$ is a convex set, then its image, $f(S) = \{f(\mathbf{x}) \rvert \mathbf{x} \in S\} \subseteq \mathbb{R}^n$, is also convex.	• The perspective function decreases the dimension of the function domain since $\dim(\operatorname{dom}(f)) = n + 1$. • Its effect is similar to the camera zoom. • The inverse image is also convex, that is, if $C \subseteq \mathbb{R}^n$ is convex, then $f^{-1}(C) = \left\{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid \mathbf{x}/t \in C, t > 0\right\}$ is also convex. • A special case is when $n = 1$, which is called *quadratic-over-linear function*.
Projective (or linear-fractional) function, $f : \mathbb{R}^n \to \mathbb{R}^m$ • $f = p \circ g$, i.e., $f(\mathbf{x}) = (p \circ g)(\mathbf{x}) = p(g(\mathbf{x}))$, where ▹ $g : \mathbb{R}^n \to \mathbb{R}^{m+1}$ is an affine function given by $g(\mathbf{x}) = \begin{bmatrix}\mathbf{A}\\ \mathbf{c}^\mathsf{T}\end{bmatrix}\mathbf{x} + \begin{bmatrix}\mathbf{b}\\ d\end{bmatrix}$, being $\mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{b} \in \mathbb{R}^m, \mathbf{c} \in \mathbb{R}^n$, and $d \in \mathbb{R}$. ▹ $p : \mathbb{R}^{m+1} \to \mathbb{R}^m$ is the perspective function. • $f(\mathbf{x}) = \mathcal{P}^{-1}(\mathbf{Q}\mathcal{P}(\mathbf{x}))$ ▹ $\mathcal{P}(\mathbf{x}) = \{(t\mathbf{x}, t) \mid t \ge 0\} \subset \mathbb{R}^{n+1}$ ▹ $\mathbf{Q} = \begin{bmatrix}\mathbf{A} & \mathbf{b}\\ \mathbf{c}^\mathsf{T} & d\end{bmatrix} \in \mathbb{R}^{(m+1)\times(n+1)}$	Yes, if $S \subseteq \operatorname{dom}(f)$ is a convex set, then its image, $f(S) = \{f(\mathbf{x}) \rvert \mathbf{x} \in S\} \subseteq \mathbb{R}^n$, is also convex.	• The linear and affine functions are special cases of the linear-fractional function. • $\operatorname{dom}(f) = \left\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{c}^\mathsf{T}\mathbf{x} + d > 0\right\}$ • $\mathcal{P}(\mathbf{x}) \subset \mathbb{R}^{n+1}$ is a ray set that begins at the origin and its last component takes only positive values. For each $\mathbf{x} \in \operatorname{dom}(f)$, it is associated a ray set in $\mathbb{R}^{n+1}$ in this form. This (projective) correspondence between all points in $\operatorname{dom}(f)$ and their respective sets $\mathcal{P}$ is a biunivocal mapping. • The linear transformation $\mathbf{Q}$ acts on these rays, forming another set of rays. • Finally we take the inverse projective transformation to recover $f(\mathbf{x})$.
Epigraph: • $\operatorname{epi} f = \{(\mathbf{x}, t) \mid \mathbf{x} \in \operatorname{dom}(f), t \ge f(\mathbf{x})\}$	• The function f is convex iff its epigraph is convex.	• Visually, it is the graph above the $(\mathbf{x}, f(\mathbf{x}))$ curve.
Hypograph: • $\operatorname{hypo} f = \{(\mathbf{x}, t) \mid \mathbf{x} \in \operatorname{dom}(f), t \ge f(\mathbf{x})\}$	• The function f is concave iff its hypograph is convex.	• Visually, it is the graph below the $(\mathbf{x}, f(\mathbf{x}))$ curve.

3 Convexity

3.1 First-order condition of convexity

$$f(\mathbf{y}) \ge f(\mathbf{x}) + \nabla f(\mathbf{x})^\mathsf{T}(\mathbf{y} - \mathbf{x}), \forall\, \mathbf{x}, \mathbf{y} \in \operatorname{dom}(f), \mathbf{x} \neq \mathbf{y} \tag{6}$$

- $\nabla f(\mathbf{x})$: gradient vector.
- This inequation says that the first-order Taylor approximation is a *underestimator* for convex functions.
- The first-order condition requires that f is differentiable.

3.2 Second-order condition of convexity

$$\mathbf{H} \succeq \mathbf{0} \tag{7}$$

- In other words, the Hessian matrix $\mathbf{H}$ is a positive semidefinite matrix.
- The graphic of the curvature has a positive (upward) curvature at $\mathbf{x}$.
- If $\mathbf{H} \succ \mathbf{0}, \forall\, \mathbf{x} \in \operatorname{dom}(f)$, then f is strictly convex. But if f is strictly convex, not necessarily $\mathbf{H} \succ \mathbf{0}, \forall\, \mathbf{x} \in \operatorname{dom}(f)$. Therefore, strict convexity can only be partially characterized.

4 CVX and Disciplined Convex Programming (DCP)

4.1 Introduction

- CVX is a Matlab package for constructing and solving Disciplined Convex Programs (DCP's).
- Disciplined convex programming is a methodology for constructing convex optimization problems proposed by Michael Grant, Stephen Boyd, and Yinyu Ye.
- The CVX package is also implemented in other programming languages:
  - Julia: Convex.jl.
  - R: CVXR.
  - Python: CVXPY.
- What distinguishes disciplined convex programming from more general convex programming is the rules, called *DCP ruleset*, that govern the construction of the expressions used in objective functions and constraints [5].
- Problems that violate the ruleset are rejected—even when the problem is convex. That is not to say that such problems cannot be solved using DCP; they just need to be rewritten in a way that conforms to the DCP ruleset.
- For matrix and array expressions, these rules are applied on an elementwise basis.
- CVX is *not* meant to be a tool for checking whether your problem is convex.

4.2 *No-product rule* and the scalar quadratic form expection

- CVX generally forbids products between nonconstant expressions, e.g., $x * x$ (assuming x is a scalar variable). We call this the *no-product rule*, and paying close attention to it will go a long way to ensuring that the expressions you construct are valid [5].
  - For example, the expression `x*sqrt(x)` happens to be a convex function of `x`, but its convexity cannot be verified using the CVX ruleset, and so it is rejected.
  - It can be expressed as `pow_p(x,3/2)` though, where `pow_p(·)` is a function from the atom library that substitutes power expressions.
- For practical reasons, we have chosen to make an exception to the ruleset to allow for the recognition of certain specific quadratic forms that map directly to certain convex quadratic functions (or their concave negatives) in the CVX atom library:
  - `x.*x` is mapped to the function `square(x)` from the CVS atom library, where $\mathbf{x} \in \mathbb{R}^n$.
  - `conj(x).*x` is mapped to the function `square_abs(x)` from the CVS atom library, where $\mathbf{x} \in \mathbb{C}^n$.
  - `x'.*x` is mapped to the function `square_abs(x)` from the CVS atom library, where $\mathbf{x} \in \mathbb{C}^n$ and `x'` is the complex conjugate.
  - $(\mathbf{Ax}+\mathbf{a})'*\mathbf{Q}*(\mathbf{Ax}+\mathbf{b})$ is mapped to the function $\texttt{quad\_form}(\mathbf{x}, \mathbf{Q})$ from the CVS atom library, where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{Q} \in \mathbb{S}^n$ (is it symmetric?), and `x'` is the complex conjugate. Note that `a` is not necessarily equal to `b`, as it is in the quadratic form.

  CVX detects the quadratic expressions such as those on the left above, and determines whether or not they are convex or concave; and if so, translates them to an equivalent function call from the atom library.
- It will *not* check, for example, sums of products of affine expressions. For example, $\texttt{x\^{}2} + 2 * \texttt{x} * \texttt{y} + \texttt{y\^{}2}$, where $\texttt{x}, \texttt{y} \in \mathbb{R}$, will cause an error on CVX, because the second term is neither convex nor concave. However, the alternative expressions $(x+y)\text{\^{}}2$ and $(x+y)*(x+y)$ are compatible to CVX.
- The quadratic form, however, can (and must) be avoided since there exist equivalent expressions.
  - For instance, $\texttt{sum}((\mathbf{A}*\mathbf{x}-\mathbf{b}).^2) \texttt{ <= 1}$ can the rewritten to the equivalent expression by using the Euclidean norm: $\texttt{norm}(\texttt{A}*\texttt{x}-\texttt{b}) \texttt{ <= 1}$, which is more efficient than that former [5].

4.3 General approaches of convex analyses

- Assume that the objective function of convex and proceed.
  - It may lead to errors.
- Verify whether the problem is convex or not
  - The basic approach is the first- and second-order conditions.
  - It usually leads to complicated analysis.
- Construct the problem as convex from the DCP ruleset and a "atom library", which is a set of basic functions that preserve convexity/concavity.
  - It is restricted to the atom library and DCP ruleset, but the convexity verification is automatic.
  - It usually involves adding auxiliary variables and reformulating the original optimization problem in order to get an expression that obeys the CDP ruleset [2].
  - The manipulation of the original problem by using operations that preserve the convexity/concavity is called convex calculus[1].
  - The reformulation usually leads to a new optimization problem that is not equal to the original one. However, they are equivalents, that is, if your find the solve the reformulated problem, then you also find the solution to the original problem.

4.4 CVX and convexity

- CVX does not consider a function to be convex or concave if it is so only over a portion of its domain, even if the argument is constrained to lie in one of these portions.
  - As an example, consider the function `1/x`. This function is convex for `x>0`, and concave for `x<0`. But you can never write `1/x` on CVX (unless `x` is constant), even if you have imposed a constraint such as `x>=1`, which restricts `x` to lie in the convex portion of function `1/x`.
  - You can use the CVX function `inv_pos(x)` (`invpos(x)` on Convex.jl), defined as `1/x` for `x>0` and ∞ otherwise, for the convex portion of `1/x`. CVX recognizes this function as convex and nonincreasing.
- Some computational functions are convex, concave, or affine only for a subset of its arguments[5].
  - For example, the function `norm(x,p)` where $p \geq 1$ is convex only in its first argument. Whenever this function is used in a CVX specification, then, the remaining arguments must be constant (these kinds of input values are called *parameters*), or CVX will issue an error message.
  - Such arguments correspond to a function's parameters in mathematical terminology; e.g.,

    $$f_p(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}, f_p(\mathbf{x}) \triangleq \|\mathbf{x}\|_p$$

    So it seems fitting that we should refer to such arguments as parameters in this context as well.
  - Henceforth, whenever we speak of a CVX function as being convex, concave, or affine, we will assume that its parameters are known and have been given appropriate, constant values.

4.5 Ruleset [5]

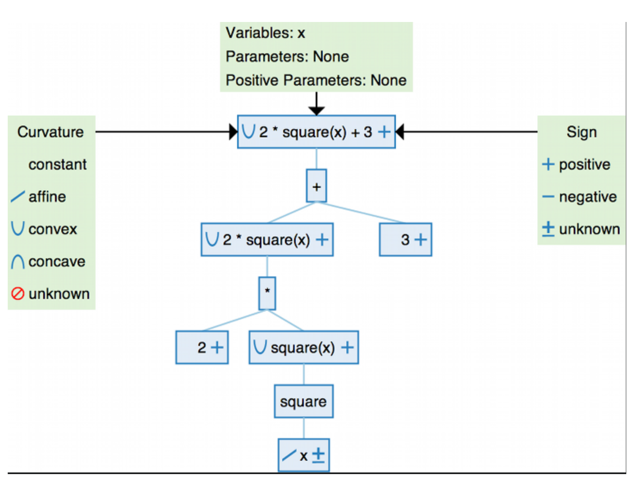
- A valid constant expression is
  - Any well-formed Matlab expression that evaluates to a finite value.
- A valid affine expression is
  - A valid constant expression;
  - A declared variable;
  - A valid call to a function in the atom library with an affine result;
  - The sum or difference of affine expressions;
  - The product of an affine expression and a constant.
- A valid convex expression is
  - A valid constant or affine expression;
  - A valid call to a function in the atom library with a convex result;
  - An affine scalar raised to a constant power $p \geq 1$, $p \notin \{3, 5, 7, 9, \dots\}$;
  - A convex scalar quadratic form;
  - The sum of two or more convex expressions;
  - The difference between a convex expression and a concave expression;
  - The product of a convex expression and a nonnegative constant;
  - The product of a concave expression and a nonpositive constant;
  - The negation of a concave expression.
- A valid concave expression is
  - A valid constant or affine expression;
  - A valid call to a function in the atom library with a concave result;
  - A concave scalar raised to a power $p \in (0, 1)$;
  - A concave scalar quadratic form;
  - The sum of two or more concave expressions;
  - The difference between a concave expression and a convex expression;
  - The product of a concave expression and a nonnegative constant;
  - The product of a convex expression and a nonpositive constant;
  - The negation of a convex expression.

4.6 Construction examples of DCP-compliant expressions

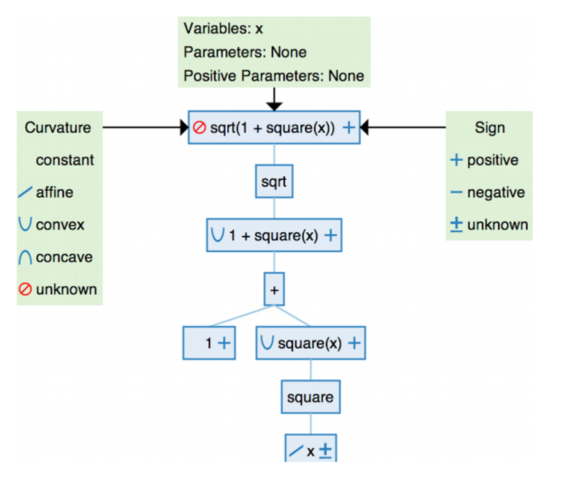
- When constructing a DCP-compliant expression, one must pay attention to three aspects of the function:
  - The range sign in the codomain $(+, -, \pm)$.
  - The curvature (convex or concave).
  - Monotonicity (nondecreasing or nonincreasing).
- The composition of functions is the base rule for the construction of expressions on the CVX family [6].
- One shall use the atoms functions in order to build expressions on CVX [5].

Consider the following examples:

- $f(\mathbf{x}) = \texttt{max}(\texttt{abs}(\mathbf{x}))$
  - $h = \texttt{max}(\cdot)$ is a convex and $\tilde{h}$ is nondecreasing in any argument. Therefore, if g is convex for any element in $\mathbf{x} \in \mathbb{R}^n$, so is $f = h \circ g$. Hence, the function $f = h \circ g = \texttt{max}(\texttt{abs}(\mathbf{x}))$ is convex for any $\mathbf{x} \in \mathbb{R}^n$.
- $f(\mathbf{x}) = \texttt{sqrt}(\langle \mathbf{k}, \mathbf{x} \rangle) + \texttt{min}(4, 1.3 - \texttt{norm}(\mathbf{A} * \mathbf{x} - \mathbf{b}))$, where $\mathbf{k}, \mathbf{A}, \mathbf{b}$ are constants.
  - $h_1 = \texttt{sqrt}(\cdot)$ is concave and nondecreasing.
  - $g_1 = \langle \cdot, \cdot \rangle$ is linear, consequently affine. Hence, it is both convex and concave.
  - Then $f_1 = h_1 \circ g_1 = \texttt{sqrt}(\langle \cdot, \cdot \rangle)$ is concave.
  - $h_2 = \texttt{min}(\cdot)$ is concave and nondecreasing.
  - $g_2 = 1.3 - \texttt{norm}(\cdot)$ is concave as it is a difference of a constant and a concave function, $\texttt{norm}(\cdot)$.
  - Then, $f_2 = h_2 \circ g_2$ is also concave.
  - Finally, $f = f_1 + f_2$ is concave since it is the sum of two concave functions (vide nonnegative weighted sum).
- $f(x) = (x\text{\^{}}2 + 1)\text{\^{}}2$
  - $g_1 = x\text{\^{}}2$ is a convex function (vide power function). $g = g_1 + 1$ is convex (vide addition/subtraction by a constant). Although $f = g^2$ is convex, the power function guarantees convexity only when the power base is solely x. For instance, the function $(x\text{\^{}}2 - 1)\text{\^{}}2$ is nonconvex. Therefore, the function $(x\text{\^{}}2 + 1)\text{\^{}}2$ would be rejected by CVX.
  - To circumvent it, one can rewrite as f as $x\text{\^{}}4 + 2 * x\text{\^{}}2 + 1$. Now, the power function guarantees that f is convex, thus this expression is DCP-compliant.
  - Another approach is to use the atom library $\texttt{square\_pos}(\cdot)$, which represents the function $(x_+)^2$, where $x_+ = \max\{0, x\}$. Now, since $h = \texttt{square\_pos}(\cdot)$ is convex and $\tilde{h}$ is nondecreasing, $f = h \circ g$ is guaranteed to be convex as long a g is convex as well. As $g = x\text{\^{}}2 + 1$ is convex, we conclude that f is convex and a valid DCP expression.
- $f(x) = 2 * x\text{\^{}}2 + 3$ [4]
  - $g(x) = x\text{\^{}}2$ is a convex function (vide power function).
  - $2 * g(x)$ is convex (vide nonnegative weighted sum).
  - Finally, $f(x) = 2 * g(x) = x\text{\^{}}2$ is a convex function (vide power function). is also convex (vide addition/subtraction by a constant).



- $f(x) = \texttt{sqrt}(1 + x\text{\^{}}2)$ [4]
  - $g(x) = x\text{\^{}}2$ is a convex function (vide power function).
  - $g(x) + 1$ is a convex function (vide addition/subtraction by a constant).
  - $h(\cdot) = \texttt{sqrt}(\cdot)$ is a concave function (vide power function) and nondecreasing. g should be convex to $f = h \circ g$ be concave. But, since g is concave, this expression is not DCP-compliant.



5 Constraints

- Type of constraints:
  - Equality constraint.
  - Inequality constraint $(\leq, \geq, \preceq_K, \succeq_K)$.
  - Strict inequality constraint $(<, >, \prec_K, \succ_K)$.
- Nonequalities is *nerver* a constraint.
- For CVX packages, strict inequalities $(<, >, \prec_K, \succ_K)$ are analyzed as inequalities $(\leq, \geq, \preceq_K, \succeq_K)$. Thus, *it is strongly recommended to only deal with nonstrict inequalities.*
- Convex and concave functions on CVX are interpreted as their *extended-valued extensions* [5]. This has the effect of automatically constraining the argument of a function to be in the function's domain.
  - For example, if we form `sqrt(x+1)` in a CVX specification, `x` will automatically be constrained to be larger than or equal to `-1`.
  - There is no need to add a separate constraint, `x>=-1`, to enforce this.

6 Methods of each optimization problem [3]

Linear Optimization	Simplex method
Convex Optimization	Branch-and-bound method
Unconstrained Optimization	subgradient, pattern search (also known as direct search, derivative-free search or black-box search)
Constrained Optimization	Interior-points method

References

[1] Stephen Boyd and Michael Grant. "Disciplined Convex Programming". In: *Convex optimization* (), p. 53.

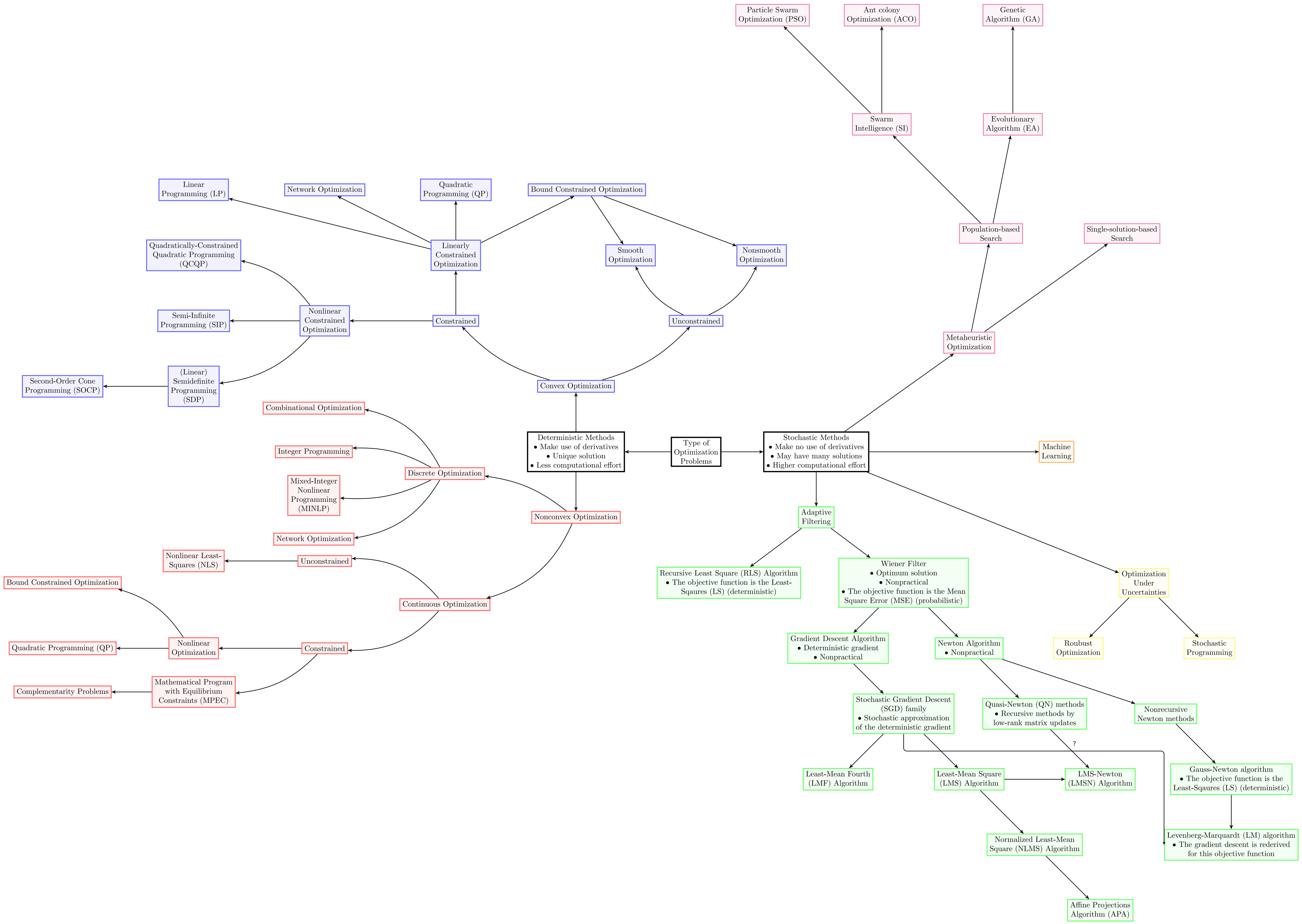
[2] *Home · Convex.Jl*. URL: `https://jump.dev/Convex.jl/stable/` (visited on 12/06/2022).

[3] Tarcisio F Maciel. "Slides - Otimização não-linear". In: (), p. 204.

[4] *Rules*. URL: `https://dcp.stanford.edu/rules` (visited on 12/08/2022).

[5] *The DCP Ruleset — CVX Users' Guide*. URL: `http://cvxr.com/cvx/doc/dcp.html` (visited on 11/27/2022).

[6] *The DCP Ruleset — CVX Users' Guide*. URL: `http://cvxr.com/cvx/doc/dcp.html#compositions` (visited on 12/04/2022).



Type of Optimization Problems
Deterministic Methods
• Make use of derivatives
• Unique solution
• Less computational effort
Stochastic Methods
• Make no use of derivatives
• May have many solutions
• Higher computational effort
Convex Optimization
Constrained
Unconstrained
Linearly Constrained Optimization
Linear Programming (LP)
Network Optimization
Quadratic Programming (QP)
Bound Constrained Optimization
Smooth Optimization
Nonsmooth Optimization
Nonlinear Constrained Optimization
Quadratically-Constrained Quadratic Programming (QCQP)
Semi-Infinite Programming (SIP)
(Linear) Semidefinite Programming (SDP)
Second-Order Cone Programming (SOCP)
Nonconvex Optimization
Discrete Optimization
Combinational Optimization
Integer Programming
Mixed-Integer Nonlinear Programming (MINLP)
Network Optimization
Continuous Optimization
Unconstrained
Nonlinear Least-Squares (NLS)
Constrained
Nonlinear Optimization
Bound Constrained Optimization
Quadratic Programming (QP)
Mathematical Program with Equilibrium Constraints (MPEC)
Complementarity Problems
Metaheuristic Optimization
Population-based Search
Single-solution-based Search
Swarm Intelligence (SI)
Evolutionary Algorithm (EA)
Particle Swarm Optimization (PSO)
Ant colony Optimization (ACO)
Genetic Algorithm (GA)
Machine Learning
Adaptive Filtering
Recursive Least Square (RLS) Algorithm
• The objective function is the Least-Sqaures (LS) (deterministic)
Wiener Filter
• Optimum solution
• Nonpractical
• The objective function is the Mean Square Error (MSE) (probabilistic)
Gradient Descent Algorithm
• Deterministic gradient
• Nonpractical
Newton Algorithm
• Nonpractical
Stochastic Gradient Descent (SGD) family
• Stochastic approximation of the deterministic gradient
Quasi-Newton (QN) methods
• Recursive methods by low-rank matrix updates
Nonrecursive Newton methods
?
Least-Mean Fourth (LMF) Algorithm
Least-Mean Square (LMS) Algorithm
LMS-Newton (LMSN) Algorithm
Gauss-Newton algorithm
• The objective function is the Least-Sqaures (LS) (deterministic)
Levenberg-Marquardt (LM) algorithm
• The gradient descent is rederived for this objective function
Normalized Least-Mean Square (NLMS) Algorithm
Affine Projections Algorithm (APA)
Optimization Under Uncertainties
Roubust Optimization
Stochastic Programming