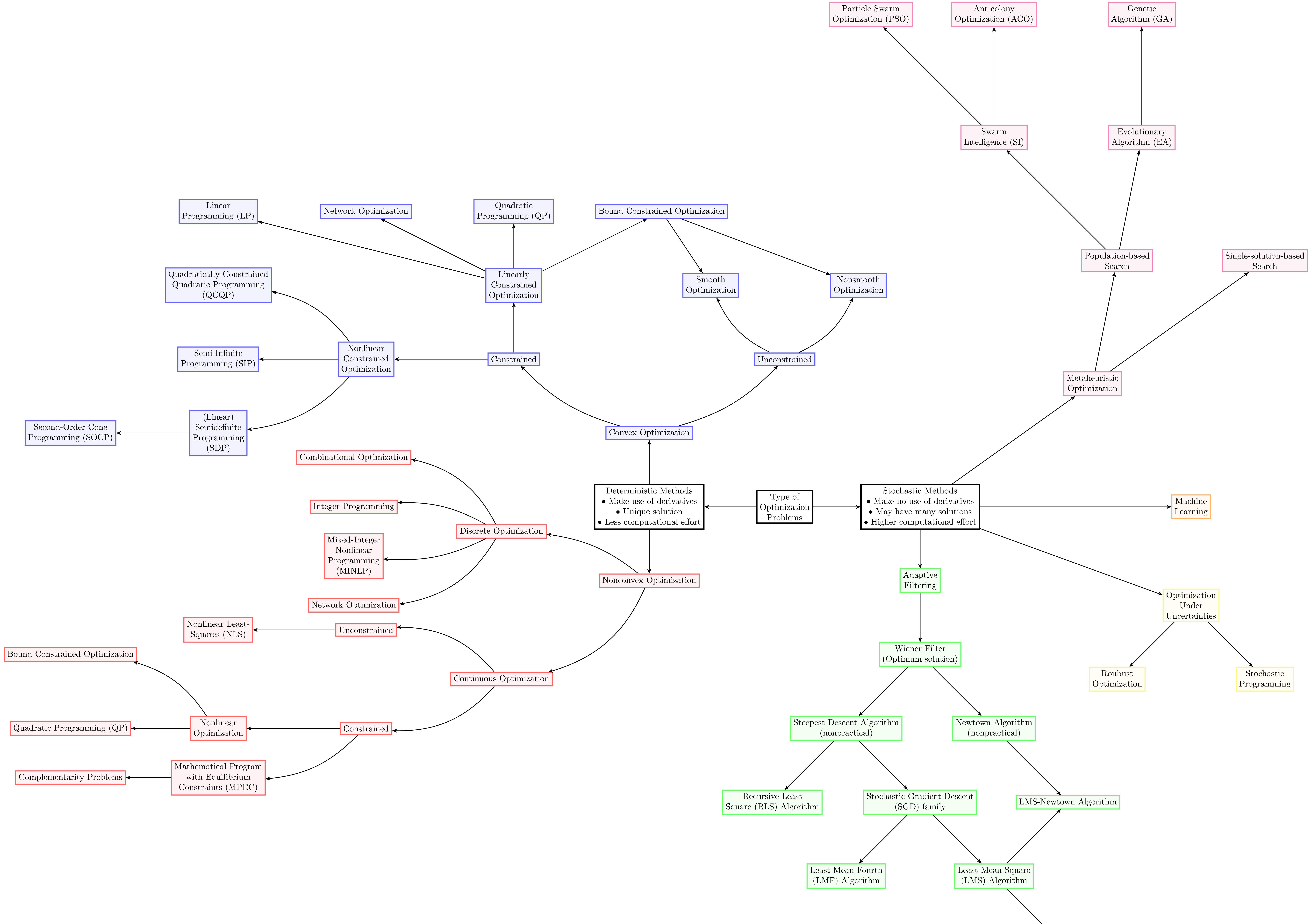
Convex sets

Set	Comments
Convex hull: • $\mathrm{conv}\ C = \left\{\sum_{i=1}^{k} \theta_i \mathbf{x}_i \mid \mathbf{x}_i \in C \text{ for } i = 1, \cdots, k, \mathbf{0} \le \boldsymbol{\theta} \le \mathbf{1}, \mathbf{1}^\top \boldsymbol{\theta} = 1\right\}$	• conv C is the smallest convex set that contains C. • conv C is a finite set as long as C is also finite.
Affine hull: • $\mathrm{aff}\ C = \left\{\sum_{i=1}^{k} \theta_i \mathbf{x}_i \mid \mathbf{x}_i \in C \text{ for } i = 1, \cdots, k, \mathbf{1}^\top \boldsymbol{\theta} = 1\right\}$	• aff C is the smallest affine set that contains C. • aff C is always an infinite set. If 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: • $B(\mathbf{x}_c, r) = \{\mathbf{x} \mid \lVert \mathbf{x} - \mathbf{x}_c \rVert \le r\}$ • $B(\mathbf{x}_c, r) = \left\{\mathbf{x} \mid (\mathbf{x} - \mathbf{x}_c)^\top (\mathbf{x} - \mathbf{x}_c) \le r^2\right\}$ • $B(\mathbf{x}_c, r) = \{\mathbf{x}_c + r \lVert \mathbf{u} \rVert \mid \lVert \mathbf{u} \rVert \le 1\}$	• $B(\mathbf{x}_c, r)$ is a finite set as long as $r < \infty$. • $\mathbf{x}_c$ is the center of the ball. • r is its radius.
Ellipsoid: • $\mathcal{E} = \left\{\mathbf{x} \mid (\mathbf{x} - \mathbf{x}_c)^\top \mathbf{P}^{-1} (\mathbf{x} - \mathbf{x}_c) \le 1\right\}$ • $\mathcal{E} = \left\{\mathbf{x}_c + \mathbf{P}^{1/2} \mathbf{u} \mid \lVert \mathbf{u} \rVert \le 1\right\}$	• $\mathcal{E}$ is a finite set as long as $\mathbf{P}$ is a finite matrix. • $\mathbf{P}$ is symmetric and positive definite, that is, $\mathbf{P} = \mathbf{P}^\top > \mathbf{0}$. 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 $\mathbf{P}^{1/2} \ge \mathbf{0}$ but singular, we say that $\mathcal{E}$ is a degenerated ellipsoid (degenerated ellipsoids are also convex).
Norm cone: • $C = \left\{(x_1, x_2, \cdots, x_n, t) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \mathbb{R}^n, \lVert \mathbf{x} \rVert_p \le t\right\} \subseteq \mathbb{R}^{n+1}$	• 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., $-K \cap K = \{\mathbf{0}\}$.	• The proper cone K is used to define the *generalized inequality* (or *partial ordering*) in some set S. For the generalized inequality, one must define both the proper cone K and the set S. • $\mathbf{x} \preceq \mathbf{y} \iff \mathbf{y} - \mathbf{x} \in K$ for $\mathbf{x}, \mathbf{y} \in S$ (generalized inequality) • $\mathbf{x} \prec \mathbf{y} \iff \mathbf{y} - \mathbf{x} \in \mathrm{int}\ K$ for $\mathbf{x}, \mathbf{y} \in S$ (strict generalized inequality). • There are two cases where K and S are understood from context and the subscript K is dropped out: ▹ When $S = \mathbb{R}^n$ and $K = \mathbb{R}^n_+$ (the nonnegative orthant). In this case, $\mathbf{x} \preceq \mathbf{y}$ means that $x_i \le y_i$. ▹ When $S = \mathcal{S}^n$ and $K = \mathcal{S}^n_+$ or $K = \mathcal{S}^n_{++}$, where $\mathcal{S}^n$ denotes the set of symmetric $n \times n$ matrices, $\mathcal{S}^n_+$ is the space of the positive semidefinite matrices, and $\mathcal{S}^n_{++}$ is the space of the positive definite matrices. $\mathcal{S}^n_+$ is a proper cone in $\mathcal{S}^n$ (??). 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 $\mathcal{S}^n_+$ in the subspace of symmetric matrices $\mathcal{S}^n$. It is usual to denote $\mathbf{X} > \mathbf{0}$ and $\mathbf{X} \ge \mathbf{0}$ 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 $S = \mathbb{R}^n$ and $K = \left\{\mathbf{c} \in \mathbb{R}^n \mid c_1 + c_2 t + \cdots + c_n t^{n-1} \ge 0, \text{ for } 0 \le t \le 1\right\}$. 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 $\mathbf{x}, \mathbf{y} \in S$. When it happens, we say that $\mathbf{x}$ and $\mathbf{y}$ are not comparable (this case does not happen in ordinary inequality, $<$ and $>$). • $\mathbf{x} \in S$ is the *minimum* element of 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 mean. The set S does not necessarily have a minimum (maximum), but the minimum (maximum) is unique if it does. • $\mathbf{x} \in S$ 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 $\mathbf{x}$. Note that 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 different minimal (maximal) elements. • 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: • $\mathcal{S}^n = \left\{\mathbf{X} \in \mathbb{R}^{n \times n} \mid \mathbf{X} = \mathbf{X}^\top\right\}$	• The positive semidefinite cone is given by $\mathcal{S}^n_+ = \{\mathbf{X} \in \mathbb{R}^{n \times n} \mid \mathbf{X} \ge \mathbf{0}\} \subset \mathcal{S}^n$. 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 $\mathcal{S}^n_{++} = \{\mathbf{X} \in \mathbb{R}^{n \times n} \mid \mathbf{X} > \mathbf{0}\} \subseteq \mathcal{S}^n_+$. 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: • $S = \mathrm{conv}\ \{\mathbf{v}_m\}_{m=0}^{k} = \left\{\sum_{i=0}^{k} \theta_i \mathbf{v}_i \mid \mathbf{0} \le \boldsymbol{\theta} \le \mathbf{1}, \mathbf{1}^\top \boldsymbol{\theta} = 1\right\}$ • $S = \{\mathbf{x} \mid \mathbf{x} = \mathbf{v}_0 + \mathbf{V}\boldsymbol{\theta}\}$, 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 $\mathbf{A}\mathbf{V} = \begin{bmatrix}\mathbf{I}_{k \times k} \\ \mathbf{0}_{n-k \times n-k}\end{bmatrix}$	• 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., $\mathrm{rank}(\mathbf{V}) = k$. All its column vectors are independent. The matrix $\mathbf{A}$ is its left pseudoinverse.
α-sublevel set: • $C_\alpha = \{\mathbf{x} \in \mathrm{dom}\,(f) \mid f(\mathbf{x}) \le \alpha\}$ (regarding convexity), where $f : \mathbb{R}^n \to \mathbb{R}$ • $C_\alpha = \{\mathbf{x} \in \mathrm{dom}\,(f) \mid f(\mathbf{x}) \ge \alpha\}$ (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. • $C_\alpha \subseteq \mathrm{dom}\,(f)$

Functions (or operators) and their implications regarding convexity

Function	Convexity	Comments
Union: $C = A \cup B$	Not in most of the cases.	
Intersection: $C = A \cap B$	Yes, if A and B are convex sets.	
Convex function: $f : \mathrm{dom}\,(f) \to \mathbb{R}$ • $f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \le \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y})$, where $0 \le \theta \le 1$. • $\mathrm{dom}\,(f)$ shall be a convex set to f be considered a convex function.	Yes.	• Graphically, the line segment between $(\mathbf{x}, f(\mathbf{x}))$ and $(\mathbf{y}, f(\mathbf{y}))$ lies always above the graph f. • In terms of sets, a function is convex iff a line segment within $\mathrm{dom}\,(f)$, which is a convex set, gives an image set that is also convex. • $\mathrm{dom} f$ is convex iff all points for any line segment within $\mathrm{dom}\,(f)$ belong to it. • *First-order condition*: f is convex iff $\mathrm{dom}\,(f)$ is convex and $f(\mathbf{y}) \ge f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}), \forall\ \mathbf{x}, \mathbf{y} \in \mathrm{dom}\,(f), \mathbf{x} \ne \mathbf{y}$, being $\nabla f(\mathbf{x})$ the 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. • If $\nabla f(\mathbf{x}) = \mathbf{0}$, then $f(\mathbf{y}) \ge f(\mathbf{x}), \forall\ \mathbf{y} \in \mathrm{dom}\,(f)$ and $\mathbf{x}$ is a global minimum. • *Second-order condition*: f is convex iff $\mathrm{dom}\,(f)$ is convex and $\mathbf{H} \ge \mathbf{0}$, that is, the Hessian matrix $\mathbf{H}$ is a positive semidefinite matrix. It means that the graphic of the curvature has a positive (upward) curvature at $\mathbf{x}$. It is important to note that, if $\mathbf{H} > \mathbf{0}, \forall\ \mathbf{x} \in \mathrm{dom}\,(f)$, then f is strictly convex. But is f is strictly convex, not necessarily that $\mathbf{H} > \mathbf{0}, \forall\ \mathbf{x} \in \mathrm{dom}\,(f)$. Therefore, strict convexity can only be partially characterized.
Convex function: $f : \mathrm{dom}\,(f) \to \mathbb{R}$ • $f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \ge \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y})$, where $0 \le \theta \le 1$. • $\mathrm{dom}\,(f)$ shall be a concave set to f be considered a concave function.	Concave	yes
Affine function $f : \mathbb{R}^n \to \mathbb{R}^m$ • $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m, \mathbf{x} \in \mathbb{R}^n$	Yes, if the domain $S \subseteq \mathbb{R}^n$ is a convex set, then its image $f(S) = \{f(\mathbf{x}) \mid \mathbf{x} \in S\} \subseteq \mathbb{R}^m$ is also convex.	• f is an affine function iff $f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) = \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y})$, where $\theta \in \mathbb{R}$. • The affine function, $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, is a broader category that encompasses the linear function, $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$. The linear function has its origin fixed at $\mathbf{0}$ after the transformation, whereas the affine function does not necessarily have it (when not, this makes the affine function nonlinear). Graphically, we can think of an affine function as a linear transformation plus a shift from the origin of $\mathbf{b}$. • A special case of the linear function is when $\mathbf{A} = \mathbf{c}^\top$. In this case, we have $f(\mathbf{x}) = \mathbf{c}^\top \mathbf{x}$, 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 affine function. 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.
Constant function $f : \mathbb{R} \to \mathbb{R}$ • $f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) = f(\mathbf{x})$, where $\theta \in \mathbb{R}$.	Convex and concave.	
Exponential function $f : \mathbb{R} \to \mathbb{R}$ • $f(x) = e^{ax} \in \mathbb{R}$, where $a \in \mathbb{R}$	Convex.	
Quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ • $f(\mathbf{x}) = a\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{p}^\top \mathbf{x} + r \in \mathbb{R}$, 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 $\mathbf{P} \ge \mathbf{0}$. • f is strictly convex iff $\mathbf{P} > \mathbf{0}$. • f is concave iff $\mathbf{P} \le \mathbf{0}$. • f is strictly concave iff $\mathbf{P} < \mathbf{0}$.	
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 $0 \le a \le 1$.	
Power of absolute value: $f : \mathbb{R} \to \mathbb{R}$ • $f(x) = \lvert x \rvert^p$, where $p \le 1$.	Convex.	
Logarithm function: $f : \mathbb{R}_{++} \to \mathbb{R}$ • $f(x) = \log x$	Concave.	
Negative entropy function: $f : \mathbb{R}_+ \to \mathbb{R}$ • $f(x) = x \log x$	Convex.	• When it is defined $f(x)\rvert_{x=0} = 0$, $\mathrm{dom}\,(f) = \mathbb{R}$.
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.	
Pointwise maximum (maximum function): $f : \mathbb{R}^n \to \mathbb{R}$ • $f(\mathbf{x}) = \max\{f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})\}$.	f is convex if $f_1, \ldots, f_n$ are convex functions.	• Its domain $\mathrm{dom}\,(f) = \bigcap_{i=1}^{n} \mathrm{dom}\,(f_i)$ is also convex.
Pointwise infimum: • $f(\mathbf{x}) = \inf_{\mathbf{y} \in \mathcal{A}} g(\mathbf{x}, \mathbf{y})$.	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. • $\mathrm{dom}\,(f) = \left\{x \mid (x, y) \in \mathrm{dom}\,(g)\ \forall\ y \in \mathcal{A}, \inf_{y \in \mathcal{A}} g(x, y) > -\infty\right\}$.
Pointwise supremum: • $f(\mathbf{x}) = \sup_{\mathbf{y} \in \mathcal{A}} g(\mathbf{x}, \mathbf{y})$.	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. • $\mathrm{dom}\,(f) = \left\{x \mid (x, y) \in \mathrm{dom}\,(g)\ \forall\ y \in \mathcal{A}, \sup_{y \in \mathcal{A}} g(x, y) < \infty\right\}$. • 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: $\mathrm{epi}\ f = \bigcap_{y \in \mathcal{A}} \mathrm{epi}\ g(\cdot, y)$
Minimum function: $f : \mathbb{R}^n \to \mathbb{R}$ • $f(\mathbf{x}) = \min\{f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})\}$.	Nonconvex and nonconcave in most of the cases.	
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 $f : \mathcal{S}^n_{++} \to \mathbb{R}$ • $f(\mathbf{X}) = \log \lvert \mathbf{X} \rvert$	Convex.	• $\mathbf{X}$ is positive semidefinite, i.e., $\mathbf{X} > \mathbf{0} \therefore \mathbf{X} \in \mathcal{S}^n_{++}$.
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: ▹ $g : \mathbb{R}^n \to \mathbb{R}^k$. ▹ $h : \mathbb{R}^k \to \mathbb{R}$. ▹ $\mathrm{dom}\,(f) = \{\mathbf{x} \in \mathrm{dom}\,(g) \mid g(\mathbf{x}) \in \mathrm{dom}\,(h)\}$.	• 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, $\mathrm{dom}\,(h)$ 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, $\mathrm{dom}\,(h)$ 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 $\mathrm{dom}\,(h)$ 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 $f(\mathbf{x}) = h(g(\mathbf{x})) = \exp g(\mathbf{x})$ is convex. ▹ If g is concave and $\mathrm{dom}\,(g) \subseteq \mathbb{R}_{++}$, then $f(\mathbf{x}) = h(g(\mathbf{x})) = \log g(\mathbf{x})$ is concave. ▹ If g is concave and $\mathrm{dom}\,(g) \subseteq \mathbb{R}_{++}$, then $f(\mathbf{x}) = h(g(\mathbf{x})) = 1/g(\mathbf{x})$ is convex. ▹ If g is convex and $\mathrm{dom}\,(g) \subseteq \mathbb{R}_+$, 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 $\mathrm{dom}\,(f) = \{\mathbf{x} \mid g(\mathbf{x}) < 0\}$. • For vector composition, we have the following examples: ▹ If $g(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ 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 $\mathrm{dom}\,(g_i) = \mathbb{R}^n$, 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 $\mathrm{dom}\,(h) = \mathbb{R}^n_+$, 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: $f : \mathrm{dom}\,(f) \to \mathbb{R}$ • $f(\mathbf{x}) = \sum_{i=1}^{m} w_i f_i(\mathbf{x})$, 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 is when $f = wf$ (a nonnegative scaling) and $f = f_1 + f_2$ (sum).
Integral function $f : \mathbb{R}^n \to \mathbb{R}$: • $f(\mathbf{x}) = \int_{\mathcal{A}} w(\mathbf{y}) g(\mathbf{x}, \mathbf{y})\, \mathrm{d}\mathbf{y}$, 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 $S \subseteq \mathrm{dom}\,(f)$ is a convex set, then its image, $f(S) = \{f(\mathbf{x}) \mid \mathbf{x} \in S\} \subseteq \mathbb{R}^n$, is also convex.	• The perspective function decreases the dimension of the function domain since $\dim(\mathrm{dom}\,(f)) = n + 1$. • 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 $g(\mathbf{x}) = \begin{bmatrix}\mathbf{A} \\ \mathbf{c}^\top\end{bmatrix} \mathbf{x} + \begin{bmatrix}\mathbf{b} \\ d\end{bmatrix}$, 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. • $f(\mathbf{x}) = \mathcal{P}^{-1}(\mathbf{Q}\mathcal{P}(\mathbf{x}))$ ▹ $\mathcal{P}(\mathbf{x}) = \{(t\mathbf{x}, t) \mid t \ge 0\} \subset \mathbb{R}^{n+1}$ ▹ $\mathbf{Q} = \begin{bmatrix}\mathbf{A} & \mathbf{b} \\ \mathbf{c}^\top & d\end{bmatrix} \in \mathbb{R}^{(m+1) \times (n+1)}$	Yes, if $S \subseteq \mathrm{dom}\,(f)$ is a convex set, then its image, $f(S) = \{f(\mathbf{x}) \mid \mathbf{x} \in S\} \subseteq \mathbb{R}^n$, is also convex.	• The linear and affine functions are special cases of the linear-fractional function. • $\mathrm{dom}\,(f) = \left\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{c}^\top \mathbf{x} + d > 0\right\}$ • $\mathcal{P}(\mathbf{x}) \subset \mathbb{R}^{n+1}$ is a ray set that begins at the origin and its last component takes only positive values. For each $\mathbf{x} \in \mathrm{dom}\,(f)$, it is associated a ray set in $\mathbb{R}^{n+1}$ in this form. This (projective) correspondence between all points in $\mathrm{dom}\,(f)$ 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: • $\mathrm{epi}\ f = \{(\mathbf{x}, t) \mid \mathbf{x} \in \mathrm{dom}\,(f), t \ge f(\mathbf{x})\}$	• The function f is convex iff its epigraph is convex.	• Visually, it is the graph above the $(\mathbf{x}, f(\mathbf{x}))$ curve.
Hypograph: • $\mathrm{hypo}\ f = \{(\mathbf{x}, t) \mid \mathbf{x} \in \mathrm{dom}\,(f), t \ge f(\mathbf{x})\}$	• The function f is concave iff its hypograph is convex.	• Visually, it is the graph below the $(\mathbf{x}, f(\mathbf{x}))$ curve.



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
Nonlinear Constrained Optimization
Linear Programming (LP)
Network Optimization
Quadratic Programming (QP)
Bound Constrained Optimization
Smooth Optimization
Nonsmooth 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
Optimization Under Uncertainties
Roubust Optimization
Stochastic Programming
Adaptive Filtering
Wiener Filter (Optimum solution)
Steepest Descent Algorithm (nonpractical)
Newtown Algorithm (nonpractical)
Recursive Least Square (RLS) Algorithm
Stochastic Gradient Descent (SGD) family
LMS-Newtown Algorithm
Least-Mean Fourth (LMF) Algorithm
Least-Mean Square (LMS) Algorithm