Optimization Cheatsheet

Hongyi Guo*

October 27, 2020

1 Linear Optimization

Definition 1 (Standard Form).

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & \sum_{j=1}^{n} a_j x_j = b \\ & x \geq 0. \end{aligned}$$

Definition 2 (Convex set). A set X is convex if $\forall x, y \in X$, $\lambda x + (1-\lambda)y \in X$, $0 \leq \lambda \leq 1$.

Definition 3 (Convex cone). A convex cone C is a convex set s.t. $\forall x \in C$, $\lambda x \in C$, $\forall \lambda \geq 0$.

Definition 4 (Hyperplane, half-space). The set $\{x \in \mathbb{R}^n : a^\top x = b\}$ is called a Hyperplane. The set $\{x \in \mathbb{R}^n : a^\top x \geq b\}$ is called a half-space.

Definition 5 (Polyhedron). A polyhedron is a set that can be described as the intersection of finitely many inequalities (half spaces) and equalities (hyperplanes).

Definition 6 (Bounded). A set $S \in \mathbb{R}^n$ is bounded if there exists a finite constant K s.t. $\forall x \in S, |x_i| \leq K, i = 1, 2, \ldots, n$.

Definition 7 (Dimension). Given a polyhedron $P = \{x \in \mathbb{R}^n : Ax \geq b\}$, the dimension $\dim(P)$ is the smallest affine subspace that contains it. P is full dimensional if $\dim(P) = n$.

Theorem 8. Two approaches to get the dimension of a polyhedron $P = \{x \in \mathbb{R}^n : Ax \geq b\}$.

1. Let $A^=$ be the submatrix of A consisting of the rows a^i, the i-th row of A s.t. $a^i x = b_i$, $\forall x \in P$ for $i = 1, \ldots, m$. Then $\dim(P) = n - \text{rank}(A^=)$.

*Northwestern University; hongyiguo2025@u.northwestern.edu

2. If the maximum number of affinely independent points in P is k, then $\dim(P) = k - 1$.

Remark 9. Normally, Theorem 8.1 gives us an upper bound of $\dim(P)$, while Theorem 8.2 gives us a lower bound of $\dim(P)$. By showing that they meet each other, we obtain the exact $\dim(P)$.

Definition 10. The definition of Linear, affine, conic, and convex combinations.

(a). Linear combination $\operatorname{span}(S)$: $\bar{x} = \sum_{i=1}^{k} \lambda_i x^i$.

(b). Affine combination $\operatorname{aff}(S)$: $\bar{x} = \sum_{i=1}^{k} \lambda_i x^i, \sum_{i=1}^{k} \lambda_i = 1$.

(c). Conic combination $\operatorname{cone}(S)$: $\bar{x} = \sum_{i=1}^{k} \lambda_i x^i, \lambda_i \geq 0, \forall i$.

(d). Convex combination $\operatorname{conv}(S)$: $\bar{x} = \sum_{i=1}^{k} \lambda_i x^i, \sum_{i=1}^{k} \lambda_i = 1, 0 \leq \lambda \leq 1, \forall i$.

Definition 11 (Linearly/affinely independent)**.** The vectors $x^1, x^2, \ldots, x^k$ are linearly independent if $\sum_{i=1}^{k} \lambda_i x^1 = 0$ has a unique solution $\lambda_i = 0, \forall i$. The vectors $x^1, x^2, \ldots, x^k$ are affinely independent if $\sum_{i=1}^{k} \lambda_i x^1 = 0, \sum_{i=0}^{k} \lambda_i = 0$ has a unique solution $\lambda_i = 0, \forall i$.

Proposition 12. $x^1, \ldots, x^k$ are affinely independent iff. $x^1 - x^k, \ldots, x^{k-1} - x^k$ are linearly independent.

Proposition 13. $x^1, \ldots, x^k$ are affinely independent iff. $(1, x^1), \ldots, (1, x^{k-1})$ are linearly independent.

Theorem 14. The intersection of convex sets is convex.

Theorem 15. Every polyhedron is a convex set.

Theorem 16. A convex combination of a finite number of vectors in a convex set also belongs to that set.

Theorem 17. The convex hull of a finite number of vectors is a convex set.

Definition 18 (Ray)**.** A ray is given by $x_0 + \lambda d, \lambda \geq 0$, for a given point x_0 and a direction vector $d \in \mathbb{R}^n$ s.t. $d \neq \mathbf{0}$.

Definition 19 (Recession direction)**.** A direction $d \in \mathbb{R}^n, d \neq \mathbf{0}$ is called a recession direction of a convex set S, if $\forall x \in S$, we have that $x + \lambda d \in S, \forall \lambda \geq 0$.

Definition 20 (Extreme point)**.** A point $x \in S$ is an extreme point of S, if it's not a <u>convex</u> combination of any $x^1, x^2 \in S, x^1 \neq x^2$.

Definition 21 (Extreme ray)**.** A direction d is an extreme ray of S if it is not a <u>conical</u> combination of any $d^1 \neq d^2 \neq d$ that are rays of S.

Definition 22 (Active/binding/tight)**.** Consider $P \in \mathbb{R}^n$ is defined by constraints that are inequalities and equalities. If a given point $\bar{x}$ satisfies a constraint at equality, then the constraint is active/binding/tight at $\bar{x}$.

Definition 23 (Basic solution (b.s.) & basic feasible solution (b.f.s.))**.** Let $P \in \mathbb{R}^n$ be a polyhedron. The vector $\bar{x}$ is a b.s. to P if all equality constraints are active at $\bar{x}$, and there are n linearly independent constraints active at $\bar{x}$. Furthermore, if $\bar{x} \in P$, then $\bar{x}$ is also a b.f.s. to P.

Definition 24 (Degenerate)**.** A b.s. $\bar{x}$ is degenerate if there are more than n constrains active at $\bar{x}$.

Definition 25 (Adjacent)**.** Two distinct b.s. $\bar{x}, \bar{y}$ are adjacent if the $n-1$ linearly independent constraints active at $\bar{x}$ and $\bar{y}$ are common, and one active constraint is not common.

Theorem 26. *Let $P = \{x \in \mathbb{R}^n : Ax \geq b\}, P \neq \emptyset$. Then $\bar{x} \in P$ is an extreme point of P iff. $\bar{x}$ is a b.f.s.*

Definition 27 (Line)**.** A polyhedron $P \subset \mathbb{R}^n$ contains a line, if there exists a vector $x \in P$ and direction $d \in \mathbb{R}^n, d \neq \mathbf{0}$ s.t. $x + \lambda d \in P, \forall \lambda \in \mathbb{R}$.

Theorem 28. *Let $P \subset \mathbb{R}^n$ be a polyhedron and $P \neq \emptyset$. Then P does not contain a line iff. P has at least one extreme point.*

Theorem 29. *If there exists an optimal solution to an LP that has an extreme point (i.e., the LP is feasible and it has a bounded optimal solution), then there exists an optimal solution that is an extreme point.*

Theorem 30. *Consider $P = \{x \in \mathbb{R}^n : Ax = b, x_j \geq 0, j = 1, \ldots, n\}$, where A is an $m \times n$ matrix with $m \leq n$ and $\mathrm{rank}(A) = m$. Suppose $P \neq \emptyset$. We say that $\bar{x}$ is a b.s. iff. $\underline{A\bar{x} = b}$ and there exists m linearly independent columns $A^{B(1)}, \ldots, A^{B(m)}$ of A, with $\underline{\bar{x}_j = 0, \forall j \notin \{B(1), \ldots, B(m)\}}$. The variables $x_j, j \in \{B(1), \ldots, B(m)\}$ are called basic variables, and the other variables are called nonbasic.*

Definition 31 (Convex)**.** A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if for every $x, y \in \mathbb{R}^n, \lambda \in [0, 1]$, we have $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$.

Definition 32 (Convex optimization problems)**.** Convex optimization problem is to minimize a convex function over a convex set.

Definition 33 (Local/global optima)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function, and $S \subset \mathbb{R}^n$ be a convex set. A given point $x^* \in S$ is a local optimum for minimizing $f(x)$ over $x \in S$ if $\exists \epsilon > 0$ s.t. $\forall x \in S$ with $\|x^* - x\| \leq \epsilon$, we have $f(x^*) \leq f(x)$. The point x^* is also a global optimum if $f(x^*) \leq f(x), \forall x \in S$.

Algorithm 1 Simplex Method

Input: A standard form LP: $\min z = c^\top x$ s.t. $Ax = b, x \geq 0$. Initial b.f.s. x.

while TRUE **do**
  Let $x = (x_B, x_N)$, $c = (c_B, c_N)$, $d = (d_B, d_N)$, $B = [A_{B(1)}, \dots, A_{B(m)}]$ be the basis matrix.
  Calculate the <u>reduced cost</u> of variable x_j: $\bar{c}_j = c_j - c_B^\top B^{-1} A_j$.
  <u>Steepest descent rule</u>: Choose the variable x_j with the smallest $\bar{c}_j (< 0)$.
  if $\bar{c} \geq \mathbf{0}$ **then**
    $z^* \leftarrow c^\top x$. // Optimal solution
    break.
  end if
  Calculate the <u>feasible direction</u>: $d_B \leftarrow -B^{-1} A_j, d_N \leftarrow e_j$.
  if $d_B \geq \mathbf{0}$ **then**
    $z^* = \leftarrow -\infty$. // Unbounded
    break.
  end if
  <u>Ratio test</u>: $\theta^* = \min_{i=1,\dots,m: d_{B(i)} < 0} -x_{B(i)} / d_{B(i)}$.
  Update: $x \leftarrow x + \theta^* d$
end while

Output: Optimal solution x if exists.

Definition 34 (Feasible direction)**.** Let $x \in P$. A vector $d \in \mathbb{R}^n, d \neq \mathbf{0}$ is called a feasible direction at x if $\exists \theta > 0$ for which $x + \theta d \in P$.

Theorem 35 (Revised simplex)**.** $[B_i^{-1} | u] \to [B_{i+1}^{-1} | \mathbf{e}_\ell]$, where $u = -d_B = B_i^{-1} A_\ell$, and ℓ is s.t. $x_{B(\ell)}$ leaves the basis. Perform row operations on B_i^{-1}, with the ℓ-th row as the pivot row, until we turn u into $\mathbf{e}_\ell$. Notice that the order of variables in c_B must be the same as that of B^{-1}.

Remark 36 (Full Tableau Implementation of simplex)**.** The tableau is in the following form.

$-z$	x_B	x_N
$-c_B^\top B^{-1} b$	$\mathbf{0}$	$c_N^\top - c_B^\top B^{-1} N$
$x_B = B^{-1} b$	$\boldsymbol{I}$	$B^{-1} N$

Theorem 37. Assume that the LP is feasible, and every b.f.s. is nondegenerate. Then the simplex method stops after a finite number of iterations, with an optimal b.f.s. and a finite objective function value, or with a vector d s.t. $Ad = 0, d \geq 0$ and $c^\top d < 0$, and $-\infty$ as the optimal cost.

Definition 38 (Lexicographical larger (smaller))**.** A vector $u \in \mathbb{R}^n$ is said to be lexicographical larger (or smaller) than another vector $v \in \mathbb{R}^n$, if $v \neq u$ and the first nonzero component of $u - v$ is positive (or negative, respectively). We denote this by $u >_L v$ (or $u <_L v$, respectively).