Fundamentals

- Outcomes: Infeasible, Unbounded, Optimal
- (Weierstrass) If f continuous and X compact, then $\min_x\{f(x) : x \in X\}$ where $X = \{x : g_i(x) \leq x, \forall i\}$ has an optimal solution.

Convexity

- Equivalent defintions for f convex (for all x, y):
  - $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$
  - $f(y) \geq f(x) + \nabla f(x)^T(y-x)$ for differentiable f
  - $\nabla^2 f(x)$ is PSD for differentiable $\nabla f(x)$
  - $epi(f)$ is a convex set
- Convexity for convex f is preserved under:
  - Conic combinations of convex functions
  - Maximization of convex functions
  - $g(x)$ convex and component-wise non-decreasing $\implies f(x) = g(f_1(x), ..., f_k(x))$ is convex for convex $f_i(x)$
  - $g(x) = f(Ax+b)$ is convex
  - $g(x) = \max_{y \in Y}\{f(x,y)\}$is convex if $\forall y$, $f(x,y)$ is convex in x
  - $g(x) = \min_{y \in Y}\{f(x,y)\}$ is convex if Y is a convex set
- If f is convex, then $C_\alpha = \{x : f(x) \leq \alpha\}$ is a convex set, but the converse is not true.
- (Separation) Let X be a nonempty closed convex set. If $\hat{x} \notin X \implies \exists(\pi, \pi_0)$ such that $\pi^T\hat{x} < \pi_0$ and $\pi^T x \geq \pi_0, \forall x \in X$.

Polyhedral Theory

- X has a polyhedral representation (p.r.) if $X = \{x : \exists y, Ax + By \leq b\}$
- If X has a p.r. then $\min_x\{c^T x : x \in X\}$ is equivalent to $\min_{x,y}\{c^T x : Ax + By \leq b\}$
  - A function g is polyhedral if $epi(g)$ has a p.r.
- Projection of a polyhedron is a polyhedron, but not true for lifting of a polyhedron
- If $X = \{x \in \mathbb{R}^n : Ax = b\} \neq \emptyset$ then the maximum number of affinely independent points in X is $n + 1 - \text{rank}(A)$.
- Let $X = \{x \in \mathbb{R}^n : Ax = b, Cx \leq d\}$ and $\exists \hat{x} \in X$ such that $C\hat{x} < d$. Then, $\dim(X) = n - \text{rank}(A)$.
- $\dim(X) = k$ if the **maximum** number of affinely independent points in X is $k+1$.
- (Caratheodory) If $\hat{x} \in \text{conv}(X)$ and $\dim(X) = k$ then $\exists$ d points $X' = \{x^1, ..., x^d\} \subseteq X$ with $d \leq k+1$ such that $\hat{x} \in \text{conv}(X')$.
- (Radon) For $\{x^1, ..., x^k\} \subseteq \mathbb{R}^n, k > n+1$, there is a partition $I \sqcup J = \{1, ..., k\}$ with non-empty I, J such that

$$\text{conv}(\{x^i : i \in I\}) \cap \text{conv}(\{x^j : j \in J\}) = \emptyset$$

- (Helley) For convex sets $X_1, ..., X_k$ each of dimension d, the every $d+1$ sets has a common point, then all k sets have a common point.
- **Farkas' Lemma:** Use duality to derive.
  - $\pi^T x \leq \pi_0$ is valid $\iff$ $\exists u \geq 0$ such that $u^T A = \pi^T, u^T b \leq \pi_0$
- **Extreme points** $x \in P$:
  - Cannot be expressed as a convex combination of two distinct points of P
  - There are n linearly independent constraints binding at x
  - Zero dimensional face and unique minimizer of some objective function over P
- **Rays or directions**
  - If $P = \{x : Ax \geq b\}$ then d is a direction $\iff$ $d \in \{d : Ad \geq 0\}$
  - Rays d are **extreme rays** $\iff$ d cannot be expressed as the sum of two different rays up to scaling $\iff$ there are $(n-1)$ active constraints at d in the related recession cone.
  - Minimization (LP) is unbounded $\iff$ $\exists$ extreme ray d such that $c^T d < 0$.
- (Representation) Polyhedrons are Minkowski sums of a polytope and a recessive cone.

Simplex Method (Basic Definitions)

- A **bfs** has the form $x = (x_B, x_N) = (A_B^{-1}b, 0)$
- Alternate form: $c^T x = c_N^T\left(A_B^{-1}b - A_B^{-1}A_N x_B\right) + c_N^T x_N$
- (Pricing) **Reduced costs** are $r^T = c^T - c_B^T A_B^{-1} A$; we want $r^T \geq 0$ to stop or choose basic index i such that $r_i^T < 0$
- **Direction** $d^j = (d_B^j, d_N^j)$ where $d_B^j = -A_B^{-1}A_j, d_N^j = e_j$ for
  - We want $c^T d^j < 0$ (else ≥ 0 implies unbounded) with index from reduced cost probing
- (Ratio Test) Choose **leaving variable** i by finding argmax of $\theta^* = \min\left\{-\frac{x_i}{d_i^j} : i \in B, d_i^j < 0\right\}$
- **Degenerate** solutions are those where one or more of the basic indices i have value 0 in x_i.
  - These are vertices that have more than n hyperplanes that represent the point

Duality

Primal		Dual
min	$\geq$	max
# of (real) constraints	$\leftrightarrow$	# of variables
# of variables	$\leftrightarrow$	# of (real) constraints
obj. vector	$\leftrightarrow$	RHS vector
RHS vector	$\leftrightarrow$	obj. vector
≥ 0, free , ≤ 0 variables	$\leftrightarrow$	$\leq, =, \geq$ constraints
$\geq, =, \leq$ constraints	$\leftrightarrow$	≥ 0, free , ≤ 0 variables

- Optimality conditions
  - **Primal** feasibility and **Dual** feasibility
  - CSC: $\lambda_i^T[b_i - (Ax)_i] = 0, \forall i$ or $\left[c_j - (y^T A)_j\right] x_j = 0, \forall j$
- Dual Simplex
  - Start with a dual feasible solution and iterate to get a primal feasible solution
  - find l s.t. $[A_B^{-1}b]_l < 0$; check $v^T = [A_B^{-1}]_l A \geq 0$ (infeasible); ratio test $j \in \text{argmin}\left\{\frac{r_k}{|v_k|} : v_k < 0\right\}$; l leaves and j enters
- Let $f : \mathbb{R}^m \mapsto \mathbb{R}$ be a convex function. A vector $S \in \mathbb{R}^m$ is a **subgradient** of f at x^0 if

$$f(x) \geq f(x^0) + s^T(x - x^0)$$

Sensitivity

- New variable $\implies$ Check dual feasibility through r^T (may need to run primal Simplex)
- New $\leq$ constraint $\implies$ Check primal feasibility (may need to run dual Simplex)
- New $=$ constraint $\implies$ Check primal feasibility (may need to run primal Simplex)
- Changes in RHS $\implies$ Check range of (primal) feasibility (may need to run dual Simplex)
- Changes in costs $\implies$ Check range of (dual) feasibility (may need to run primal Simplex)
- Changes in nonbasic column of $A \implies$ Check range of (dual) feasibility (may need to run primal Simplex)
- Changes in basic column of $A \implies$ Check range of (dual+primal) feasibility (may need to run from scratch)
- Parametric Programming
  - Turns out this is a concave optimization problem related to the Lagrangian:

$$F(\theta) = \min_x\{c^T x : Ax = b + \theta d, x \geq 0\}$$
$$G(\theta) = \min_x\{(c + \theta d)^T x : Ax = b, x \geq 0\}$$

Decompositions

- Benders:
  - For systems where one variable is in every constraint (one column in the block system is filled in)
  - *Original Problem*: $\min_{x\geq 0, Ax=b}\left\{c^T x + \sum_{k=1}^K p_k Q_k(x)\right\}$ where $Q_k(x) = \min_y\{q_k^T y : W_k y = h_k - T_k x, y \geq 0\}$.
  - *Reformulation*:

$$\begin{aligned} \min_{x,\theta}\ & c^T x + \sum_{k=1}^K p_k\theta_k \\ \text{s.t. } & Ax = b, x \geq 0 \\ & \theta_k \geq (\pi_k^i)^T(h_k - T_k x), \forall i \\ & (\psi_k^j)^T(h_k - T_k x) \leq 0, \forall j \end{aligned}$$

  where $\{\pi_k^i\}, \{\psi_k^i\}$ are respective extreme points and rays of $\Pi_k = \{\pi : \pi^T W_k \leq q_k\} \neq \emptyset$.
- Dantzig-Wolfe
  - Exploits the Representation Theorem
  - *Original Problem*: $\min_{x^i\in X_i}\left\{\sum_{i=1}^m (c^i)^T x : \sum_{i=1}^m D_i x^i = b\right\}$
  - *Reformulation (for all i):*

$$\begin{aligned} \min_{\lambda,\mu\geq 0}\ & \sum_{i=1}^m\left[\sum_{k\in K_i}\lambda_k (c^i)^T u^k + \sum_{l\in L_i}(c^i)^T v^l\right] \\ \textit{s.t.}\ & \sum_{i=1}^m\left[\sum_{k\in K_i}\lambda_k D_i u^k + \sum_{l\in L_i}\mu_l D_i v^l\right] = b \\ & \sum_{k\in K_i}\lambda_k = 1 \end{aligned}$$

  where we may want to only choose a small subset of the vertices and extreme rays (RMP) and price in additional ones via

$$z_i = \min_x\{[(c^i)^T - \alpha^T D_i]x : x \in X_i\}$$

    * $z_i = -\infty \implies$ found an extreme ray
    * $z_i \leq \beta_i$ the RMP optimal solution $\implies$ found an extreme point

Network Flows

- A matrix is totally unimodular (TU) if all square submatrices have $\det(A) \in \{-1, 0, 1\}$.
  - The node-incidence matrix is TU
- Min-cost is equivalent to max-flow if the cost vector is 0
- Min-cost is equivalent to shortest path (Dijkstra's algorithm) if there are no capacity constraints

Ellipsoid Method

- An ellipsoid can be defined via a center z and positive definite matrix D via

$$E(z, D) = \{x \in \mathbb{R}^n : (x - z)^T D^{-1}(x - z) \leq 1\}$$

- Ellipsoids in the ellipsoid method (feasibility) follow $\frac{Vol(E')}{Vol(E)} \leq e^{-\frac{1}{2(n+1)}}$ and runs in $O(n^6\log(nU))$ iterations.
- Feasibility is equivalent to optimization
- Separation is equivalent to optimization if our separation oracle is fast.