Vectors: Remember determinant of 3×3 matrix has $-$ for second term. **Linear independence**: the unique solution to the system of equations $\sum_{i=1}^{K} \alpha_i \mathbf{x}^i = \mathbf{0}$ is $\alpha_i = 0$. **Inverse Matrix**: A 2×2 matrix has inverse: $A^{-1} = \frac{1}{ad-bc} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, From $A = IA \to I = AA^{-1}$. **Taylor's approximation**: Second term is $\frac{1}{2}(\mathbf{x} - \mathbf{x}^0)^\top \nabla^2 f(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0)$. **Closed Set**: it's limit point belongs to X. **Bounded Set**: There exists M s.t. $\|\mathbf{x}\| \leq M$. A **Compact Set** is both closed and bounded. **Convexity: first Order Condition** $f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x})$. **Convexity: Second Order Condition** Hession $\nabla^2 f(\mathbf{x})$ is psd.. **Nonnegative weighted sum**, **Maximum** of convex functions also convex. Composition of convex functions f and g_i $h(\mathbf{x}) = f(g_i(\mathbf{x}))$ is convex if **either** f is nondecreasing or if each g_i is linear. **Intersection**, **Sum** and **Product** of two convex sets are convex. Note **Union** of two convex sets is not convex in general. The epigraph epi$f = \{(y, \mathbf{x}) : y \geq f(\mathbf{x})\}$, α-level set $X_\alpha = \{\mathbf{x} : f(\mathbf{x}) \leq \alpha\}$ are convex. **Convex Optimization** write equality constraint in $\leq$ form and check its convexity. ℓ_1-norm $= \|\mathbf{x} - \mathbf{y}\|_1 = \sum_{i=1}^{n} |x_i - y_i|$. ℓ_∞-norm $= \|\mathbf{x} - \mathbf{y}\|_\infty = \max_{1 \leq i \leq n} |x_i - y_i|$. **Weierstrass's Theorem** If f is a continuous function and X is a nonempty, compact set, then (P) has an optimal solution. Sufficient but not necessary condition. If (P) is a convex program, i.e. f is a convex function and X is a convex set, then any local optimal solution is global. **Improving search** is an algorithm based on moving from one solution to a better solution. **Be careful about the statement of the question whether optimal solution refers to maximum or minimum**. Minimization problem over a compact set may have optimal solution in interior of X but has local max on the boundary. **Relaxation** Problem (Q) is a relaxation of (P) if: 1. $X \in Y$; 2. $f(x) \geq g(x)$. **Lagrangian Relaxation** $\mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \min(f(x) + \sum_{i \in I} \lambda_i(g_i(x) - b_i) + \sum_{j \in J} \mu_i(h_j(x) - d_i))$. **First order optimalityy condition**: $\nabla f(x^*)\Delta x \geq 0 \to \nabla f(x^*) = 0$ due to Δx can be either positive or negative. **Second order optimality condition**: $\Delta x^\top \nabla f(x^*)\Delta x \geq 0 \to \nabla^2 f(x^*)$ is psd. If $\nabla f(x^*)\Delta x \geq 0$ and $\nabla^2 f(x^*)$ is **positive definite**, then x^* is local min. **Gradient Descent**: The gradient is $\nabla f(x^k)$ and gradient direction $d^k = -\nabla f(x^k)$. Choose α that minimize $f(x^k + \alpha d^k)$. Update $x^{k+1} \leftarrow x^k - \alpha \nabla f(x^k)$. Behavior of gradient descent can be "zig-zags" because k and $k+1$ gradient descent directions can be perpendicular. **Newton Methods**: $x^{k+1} = x^k - (\nabla^2 f(x^k))^{-1} \nabla f(x^k)$. $d^k = -(\nabla^2 f(x^k))^{-1} \nabla f(x^k)$, and step size $\alpha = 1$. Newton's method is not guaranteed to converge. If the Hessian is singular, at some iteration we cannot proceed. **Golden Section Search**, **Quadratic Fit**, **Nelder-Mead method** are not derivative based. N-M method only depends on function value. In **Transportation problem**, objective function is $\min \sum c_{ij} x_{ij}$, the constraints are $\sum x_{ij} \geq d_j$ and $\sum x_{ij} \leq s_i$. In **Max Flow problem**, the objective function is to max the supply $\max b_s$ and the constraints are 1. flow conservation $\sum_{k \in O(i)} x_{ik} - \sum_{j \in I(i)} x_{ji} = b_i$; 2. total supply = total demand $b_t = -b_s$ and 3. Arc capacity $0 \leq x_{ij} \leq u_{ij}$. In **Min Cut problem**, the capacity of a cut is the total capacity of arcs that cross from S to its complement nodes C. In **Shortest Path problem**, $d_j \leq \min(d_i + c_{ij})$, for all incoming node i. **Here-and-Now** made before knowing uncertain parameters. **Wait-and-See** made after knowing uncertain parameters. **Two-stage stochastic inventory control** production quantity x is here-and-now variable. sell quantity and discount quantity y_i, z_i are wait-and-see variables

$$\min \quad cx + E_d(Q(x,d)), \quad Q(x,d) \quad \min -ry - sz \tag{1}$$
$$\text{s.t.} \quad 0 \leq x \leq \bar{x} \quad \text{s.t.} y \leq d \tag{2}$$
$$y + z \leq x, x, z \geq 0 \tag{3}$$

Solution depends on $\frac{r-c}{r-s}$.
Piecewise linear function: $|x| = \max\{x, -x\}$, is convex piecewise linear function. **Ray, Lines**: Ray definition $\{\mathbf{x} | \mathbf{x} = \boldsymbol{\alpha} + \theta \mathbf{d}, \forall \theta \geq 0\}$. Line definition: $\{\mathbf{x} | \mathbf{x} = \boldsymbol{\alpha} + \theta \mathbf{d}, \forall \theta \in \mathbb{R}\}$. **Hyperplane** a plane in $\mathbb{R}$ is defined as $\sum_{i=1} A x_i = c$. **Half space** a linear equity $\mathbf{a}^\top \mathbf{x} = \mathbf{c}$ can be expressed as two halfspaces: $H_1 = \mathbf{x} : \mathbf{a}^\top \mathbf{x} \leq \mathbf{c} \cap H_2 = \mathbf{x} : \mathbf{a}^\top \mathbf{x} \geq \mathbf{c}$. **Polyhedron** the intersection of a finite number of halfspaces $P = \{\mathbf{x} : A^\top \mathbf{x} \leq \mathbf{b}\}$. **Extreme Point** $\mathbf{x}$ is a extrem point if and only if it cannot be a convex combination of two different points in P. **Convex hull**: $conv\{a_1, ... a_n\} = \{x | x = \lambda_1 a_1 + \lambda_2 a_2 + ... + \lambda_n a_n, \sum_i \lambda_i = 1\}$. A nonempty and bounded polyhedron is the convex hull of it's extreme points. Also called **polytope**. **Unbounded polyhedron, recession directions**: iff there are directions to move t o infinity. These directions are called **recession directions**. A ray d is a **Conic combination** of two rays e_1, e_2 if d is a nonnegative weighted sum of e_1, e_2. A Conic hull or **cone** is $conic\{r_1, ... r_n\} = \{x | x = \lambda_1 r_1 + \lambda_2 r_2 + ... + \lambda_n r_n, \lambda_i \geq 0\}$.

1. if a polyhedron is bounded, there's no extrem ray.

2. if a polyhedron is bounded, there must be an extreme point.

3. if a polyhedron is unbounded and doesn't contain a line, it must have an extreme ray.

4. if a polyhedron is unbounded, it may not have an extreme point. e.g. $0 \leq x \leq 1$.

Weyl-Caratheodory Theorem Any nonempty polyhedron P with at least one extreme points can be formed by its extreme point and its extreme rays. An extreme point is the unique solution of n **l.i. active constraints**.

1. Active constraints: linear equality constraints.

2. Basic solution: unique solution of n linearly independent active constraints.

3. Basic feasible solution: basic solution that's feasible = extreme points.

n active constraints can be written as $\begin{bmatrix} B & N \\ 0 & I \end{bmatrix} \begin{bmatrix} x_B \\ x_N \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}$, $x_B = B^{-1}b, x_N = 0$. A polyhedron P has an extreme point iff it doesn't contain a line. **A feasible standard form LP always has a BFS**. An optimal solution in LP is a BFS but not all BFS are optimal. **From practice exam** A nonempty polyhedron in standard form must have an extreme point. Two **Adjacent BFS** share the same $n-1$ l.i.active constraints. **Simplex method**

$$d_N = \begin{bmatrix} 0 \\ \dots \\ 1\dots0 \end{bmatrix} = e_j, d_B = -B^{-1}A_j, \bar{c}_j = c_j - c_B^\top B^{-1} A_j$$

A BFS is nondegenerate in standard form LP if all basic variables are positive. If there's some basic variable is 0, then the BFS is degenerated. Simplex method can get stucked.

1. We start the simplex method with a basic feasible solution x and the associated basis consisting of the basis matrix $B = [A_{B(1)}, \dots, A_{B(m)}]$, where $A_{B(i)}$'s are the basic columns of A.

2. Compute the reduced cost $\bar{c}_j = c_j - c_B^\top B^{-1} A_j$ for each nonbasic variable x_j. There are two possibilities:

   (2.1) If all the reduced costs are nonnegative, the current BFS is optimal, and the algorithm terminates;

   (2.2) Otherwise, select some j with $\bar{c}_j < 0$, so x_j *enters the basis*.

3. Compute $d_B = -B^{-1}A_j$. There are two possibilities:

   (3.1) If $d_B \geq 0$, we know the optimal cost is $-\infty$, and the algorithm terminates.

   (3.2) Otherwise, if some entry of d_B is negative, continue to the next step.

4. If some entry of d_B is negative, compute the stepsize θ^* by the following *min-ratio test*:

$$\theta^* = \min_{\{i=1,\dots,m \,|\, d_{B(i)}<0\}} \frac{x_{B(i)}}{-d_{B(i)}}.$$

   Suppose the index $B(l)$ achieves the minimum: $x_{B(l)}$ *exits the basis*.

5. Form the new basis matrix $\overline{B}$ by replacing $A_{B(l)}$ column with A_j. The new BFS y has basic variable part $y_{B(i)} = x_{B(i)} + \theta^* d_{B(i)}$ and $y_j = \theta^*$. Start a new iteration.

Degeneracy: a BFS is nondegenerate in standard form LP if all basic variables are positive. **Bland's rule**: Choose the entering nonbasic variable x_j s.t. j is the smallest index with $\bar{c}_j \leq 0$. Choose exiting basic variable x_i s.t. i is the smallest index that achieves the min ratio in the test. Bland's rule guarantees termination. **Phase-I**

$$(P1) \quad z_{p1} = \min \sum_{i=1} y_i$$
$$\text{s.t.} \quad Ax + y = b$$
$$x, y \geq 0$$

1. if $z_{p1} \geq 0$, the original LP is infeasible.

2. if $z_{p1} = 0$, a feasible solution to the original LP is found, from which we can obtain a BFS for the original LP.

Weak Duality If x is any feasible solution to the primal LP and y is any feasible solution to the dual LP, then $c^\top x \geq b^\top y$. **Strong Duality** If a primal LP has finite optimal solution x^*, and it's dual LP must also have a finite optimal solution y^*, and the respective optimal objective values are equal, $c^\top x^* = b^\top y^*$. **Table of possible and impossible Relaxation**: lower objective function **on feasible points** and find a feasible region larger than the feasible region of the original problem.
Lagarangia Relaxation: For a minimization problem

	Finite optimal	Unbounded optimal	Infeasible
Finite optimal	possible	impossible	impossible
Unbounded optimal	impossible	impossible	possible(must)
Infeasible	impossible	possible	possible

1. If primal constraint is an equality, then the dual variable is free

2. If primal constraint has $\geq$, then the dual variable is ≥ 0

3. If primal constraint has $\leq$, then the dual variable is ≤ 0

4. If primal variable is free, then the dual constraint has $=$

5. If primal variable ≥ 0, then the dual constraint has $\leq$

6. If primal variable ≤ 0, then the dual constraint has $\geq$

Complementary Slackness $\boldsymbol{x}$ and $\boldsymbol{y}$ are optimal solutions for the primal and dual LP iff. they satisify following conditions:

1. Primal complementary slackness: $y_i(\boldsymbol{a}_i^\top \boldsymbol{x} - b_i) = 0$ for all i. That is either the i-th primal constraint is **active** or dual variable $y_i = 0$.

2. Dual complementary slackness: $x_j(\boldsymbol{c}_j - \boldsymbol{y}^\top \boldsymbol{A}_j) = 0$ for all j. That is either the j-th dual constraint is **active** or primal variables $x_j = 0$.

General rules to form LP Dual

1. Step 1 Relaxing the objective: Since $y_i(b_i - \boldsymbol{a}_i^\top \boldsymbol{x}) \leq 0$: if $\boldsymbol{a}_i^\top \boldsymbol{x} \geq b_i$, then $y_i \geq 0$;if $\boldsymbol{a}_i^\top \boldsymbol{x} \leq b_i$, then $y_i \leq 0$;if $\boldsymbol{a}_i^\top \boldsymbol{x} = b_i$, then y_i is free;

2. Step 2 Solving lagrangian relaxation: Since $(c_j - y^\top \boldsymbol{A})x_i \geq 0$:

Robust Constraint Column Generation Property 1: The BFS of Restricted Master Problem is BFS of Master Problem by setting all other variables to 0. So as Basis matrix and basic variables. **Solving Pricing problem** $w = min_{j=1,\ldots,N}\{c_j - c_B^T B^{-1}A_j\}$. For cutting stock problem, $c_j = 1$. Treat A_j as a variable and solve the maximization problem. **Knapsack Problem** $\sum_{i=1}^m w_i a_i \leq W, a_i \geq 0$. **Primal-Dual Relationship of Column Generation and Constraint Generation Separation Problem** In Restricted Dual Master Problem $\hat{Z} = min_{j=1,\ldots,N} c_j - (y^*)^\top A_j$ (constraint violation) $\to$ Let A_j be the constraint achieving the minimum. If $\hat{Z} \leq 0$, then add A_j to I. Primal Pricing Problem = Dual Separation Problem.
Coordination and Distributed Problem

$$(LP) \quad \min \quad \boldsymbol{c}^\top \boldsymbol{x} \quad \leftarrow \text{Complicating constraint}$$
$$\text{s.t.} \quad \boldsymbol{Dx} = \boldsymbol{b_0} \quad \leftarrow \text{Easy constraint}$$
$$\boldsymbol{Fx} = \boldsymbol{b}$$
$$\boldsymbol{x} \geq 0$$

Dantzig-Wolfe Decomposition Rewrite above

$$\min \quad \boldsymbol{c}^\top \boldsymbol{x}$$
$$\text{s.t.} \quad \boldsymbol{Dx} = \boldsymbol{b}$$
$$\boldsymbol{x} \in P$$

Where $P = \boldsymbol{F}\boldsymbol{x} = b, x \geq 0$ and assume it's a bounded polyhedron. Use extreme points to represent P and rewrite the problem as

$$\min_{\lambda_1, \ldots, \lambda_N} \sum_{i=1}^{N} \lambda_i (\boldsymbol{c}^\top \boldsymbol{x_i})$$

$$\text{s.t.} \sum_{i=1}^{N} \lambda_i (\boldsymbol{D}x_i) = \boldsymbol{b_0}$$

$$\sum_{i=1}^{N} \lambda_i = 1$$

$$\lambda_i \geq 0, \forall i = 1, \ldots, N$$

Apply Column Generation to solve (MP) to get optimal solution $\hat{\lambda}$ and dual optimal solution $(\hat{y}, \hat{r})$. $\hat{Z} = \min_{i=1,\ldots,N}\{\boldsymbol{c}^\top x_i - \hat{\boldsymbol{y}}^\top(\boldsymbol{D}x_i) - \hat{r}\}$. **Least Squares Optimality Condition** $\min_x (Ax - b)^\top (Ax - b) = x^\top A^\top Ax - 2(A^\top b)^\top x + b^\top b$. Optimality condition: $\nabla(Ax - b)^\top (Ax - b) = 0 \rightarrow A^\top Ax = A^\top b$(Normal Equation) $\rightarrow x = (A^\top A)^{-1}A^\top b$. **Moore-Penrose Pseudoinverse**: For a matrix $A \in R^{m \times n}$ with $m \geq n$ and full rank, the Moore-Penrose pseudoinverse of A is defined as: $A^+ = (A^\top A)^{-1}A^\top$. If A is an invertible square matrix, then $A^+ = A^{-1}$. **SVD**: Let A be an m-by-n matrix with $m \geq n$. Then we can write A as $A = U\Sigma V^\top = \sum_{i=1}^{n} \sigma_i u_i v_i^\top$, where U is an m-by-n matrix satisfying $U^\top U = I$, V is an n-by-n matrix satisfying $V^\top V = I$, and $\Sigma = diag(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$. Columns u_i of U are called left singular vectors of A, Columns v_i of V are called right singular vectors of A and $\sigma_1, \ldots, \sigma_n$ are called singular values of A. **Relation to Eigendecomposition**: if A is a symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ and orthonormal eigenvectors $u_1, \ldots, u_n$, e.g. $A = U\Lambda U^\top$, where $U = [u_1, \ldots, u_n]$, then $\sigma_i = |\lambda_i|$ and $v_i = sign(\lambda_i)u_i$. Eigenvalues of $A^\top A$ are $\sigma_1^2, \ldots, \sigma_n^2$, eigenvectors are $v_1, \ldots, v_n$, then $A^\top A = V\Sigma U^\top U\Sigma V^\top = V\Sigma^2 V^\top$. **Low Rank Approximation by SVD Theorem**: the rank r matrix closest to A in l_2 norma is $\hat{A} = U\Sigma_r V^\top$, where $\Sigma_r = diag(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0)$ and $||\hat{A} - A||_2 = \sigma_{r+1}$. **Convex Cone**: A set K is a convex cone if K is convex and $\alpha x \in K, \forall \alpha \geq 0, x \in K$. $a \succeq_K b$ means $a - b \in K$ and K is a convex cone. **Second Order Cone** $L^m = (x_1, x2, \ldots, x_m) : \sqrt{x_1^2 + x_2^2 + \ldots + x_{m-1}^2} \leq x_m$. **SOC Constraint** $Ax \succeq_{L^m} b, A \in R^{m \times n}, x \in R^n, b \in R^m$. **Positive Semi-definite**: A is a symmetric matrix. If the eigenvalues of A are nonnegative, or $x^\top Ax \geq 0$ for all $x \in R^n$, or A can be factored as $A = D^\top D$ for some matrix D. **PSD Constraint** set of all positive semidefinite matrices in $R^{n \times n}$ is S_+^n, is a convex cone. **NP-complete**: Within the set of problems in NP there's a set of problems such that every problem in NP can be easily converted to those problem are called $NP - complete$. **Problem not in NP** Problems cannot be verified in polynomial time. **NP-hard**: If an algorithm for it can be converted to the one solving NP problem. $NP - hard$ may not be in NP. All NP-complete are NP-hard but not vice versa. *Most discrete optimization problems are NP-hard*. **Set packing** is a maximization problem. **Set covering** is a minization problem. **Set partitioning** could be max or min problem. **Min Cost Network Flow**: If the supply data b_i and the capacities u_{ij} are integers, then every extreme point (or basic feasible solution) will have integer entries. **Subtour Elimination**: $x_ij + x_ji \leq 1$. **LP Relaxation Properties**: 1. If LP relaxation is unbounded then the IP can be either infeasible or unbounded. 2. If the LP relaxation has an optimal solution, then the IP could be infeasible or have an optimal solution. 3. If an optimal solution to the LP relaxation is integral then it is an optimal solution to the IP. **Distrance from a point to a hyperplane**: $d(a^\top x - b, p) = \frac{|a^\top p - b|}{||a||_2}$