Lecture Note: Mathematical Optimization Models and Applications

Lecturer

Yinyu Ye

Scriber

Bin Yu

Institute of Operations Research and Analytics

National University of Singapore

Singapore, July 29 2022

Contents

Chapter 1

Introduction

[1] **Ref Chapters 1, 2.1-2, 6.1-2, 7.2, 11.3, 11.6**

1.1 Introduction

In this section, we will give the belief introduction about the linear program and basic idea of optimization. The detail you can check out in the textbook Luenberger et al. (1984)

$$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & Ax = b \\ & x \in K \end{aligned} \tag{1.1}$$

This is the conic linear program depending on the set K

- linear program: when K is the non-negative orthant cone
- second order cone programming: when K is the second order cone
- semidefinite cone programming: when K is the semidefinite matrix cone

$$\begin{aligned} \min \quad & 2x_1 + x_2 + x_3 \\ s.t. \quad & x_1 + x_2 + x_3 = 1 \\ & x_1, x_2, x_3 \geq 0 \end{aligned} \qquad \begin{aligned} \min \quad & 2x_1 + x_2 + x_3 \\ s.t. \quad & x_1 + x_2 + x_3 = 1 \\ & \sqrt{x_2^2 + x_3^2} \leq x_1 \end{aligned} \qquad \begin{aligned} \min \quad & 2x_1 + x_2 + x_3 \\ s.t. \quad & x_1 + x_2 + x_3 = 1 \\ & \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \succeq 0 \end{aligned} \tag{1.2}$$

They are LP, SOCP, and SDP respectively.

Definition 1.1. *A symmetric matrix M with real entries is positive-definite if the real number $z^T M z$ is positive for every nonzero real column vector z*

[1] First Revision: TBC

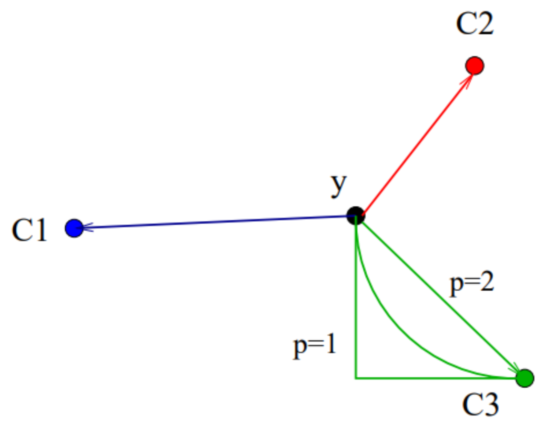

Figure 1.1: facility allocation

1.1.1 Facility Location Problem

Definition 1.2. *For a real number p, the p-norm or L^p-norm of x is defined by*

$$||x||_p = (\sum_i |x_i|^p)^{1/p}$$

Consider this unconstrained optimization, and let c_j be the location of client $j = 1, 2, ..., m$ and y be the location decision of a facility to be built. Then we solve

$$\min \quad \sum_j ||y - c_j||_p \tag{1.3}$$

In the sense, we will get the different optimal solution depending the value of p. The figure will illustrate Obviously, the green lines represents the different norm $p \in [1, 2]$

1.1.2 Sparse Linear Regression Problems

Our target is to minimize the number of non-zero entries in x such that $Ax = b$

$$\begin{aligned} \min \quad & ||x||_0 = |\{j : x_j \neq 0\}| \\ s.t. \quad & Ax = b \end{aligned} \tag{1.4}$$

Our target to minimize the number of zero in the vector or the rank in the matrix. And we use the linear regression to illustrate the idea of LASSO

Example 1.3.

$$\begin{aligned} \min_{\beta_0, \beta} \quad & \left\{ \sum_{i=1}^{N} \left(y_i - \beta_0 - x_i^T \beta \right)^2 \right\} \\ s.t. \quad & \sum_{j=1}^{p} |\beta_j| \leq t \end{aligned}$$

Here β_0 is the constant coefficient, $\beta := (\beta_1, \beta_2, \ldots, \beta_p)$ is the coefficient vector, and t is a prespecified free parameter that determines the degree of regularization.

Letting X be the covariate matrix, so that $X_{ij} = (x_i)_j$ and x_i^T is the i^{th} row of X, the expression can be written more compactly as

$$\begin{aligned} \min_{\beta_0,\beta} \quad & \|y - \beta_0 - X\beta\|_2^2 \\ s,t, \quad & \|\beta\|_1 \le t \end{aligned}$$

where $\|u\|_p = \left(\sum_{i=1}^N |u_i|^p\right)^{1/p}$ is the standard ℓ^p norm.

Denoting the scalar mean of the data points x_i by $\bar{x}$ and the mean of the response variables y_i by $\bar{y}$, the resulting estimate for β_0 is $\hat{\beta}_0 = \bar{y} - \bar{x}^T\beta$, so that

$$y_i - \hat{\beta}_0 - x_i^T\beta = y_i - \left(\bar{y} - \bar{x}^T\beta\right) - x_i^T\beta = (y_i - \bar{y}) - (x_i - \bar{x})^T \beta$$

It can be helpful to rewrite

$$\begin{aligned} \min_{\beta\in\mathbb{R}^p} \quad & \frac{1}{N}\|y - X\beta\|_2^2 \\ s.t. \quad & \|\beta\|_1 \le t \end{aligned}$$

in the so-called Lagrangian form

$$\min_{\beta\in\mathbb{R}^\mu} \frac{1}{N}\|y - X\beta\|_2^2 + \lambda\|\beta\|_1 - \lambda t$$

where the exact relationship between t and λ is data dependent.

Sometimes this objective can be accomplished by LASSO.

$$\begin{aligned} \min \quad & ||x||_1 = \sum_{i=1}^n |x_j| \\ s.t. \quad & Ax = b \end{aligned} \tag{1.5}$$

In this figure, we can illustrate why we could approximate the sparse problem by 1-norm. In the two dim, the polyhedron will touch the vertex of the 1-norm. In the sense, the vertex will $(1, 0)$ or $(0, 1)$. It will reduce the number of zero in the vector. Moreover, we also use the p-norm to solve this problem

$$\min \quad ||Ax - b||^2 + \beta(\sum_{j=1}^n |x_j|^p) \tag{1.6}$$

Usually, we let $p = \frac{1}{2}$. Then we use the cross validation method to estimate β

1.1.3 Support Vector Machine

Denote $a_i \in \mathbb{R}^d$ and $b_j \in \mathbb{R}^d$. We like to find a hyperplant, slope vector x and intersect scalar x_0

$$\begin{aligned} s.t. \quad & \mathbf{a}_i^T\mathbf{x} + x_0 + \ge 1, \forall i \\ & \mathbf{b}_j^T\mathbf{x} + x_0 \le -1, \forall j \end{aligned} \tag{1.7}$$

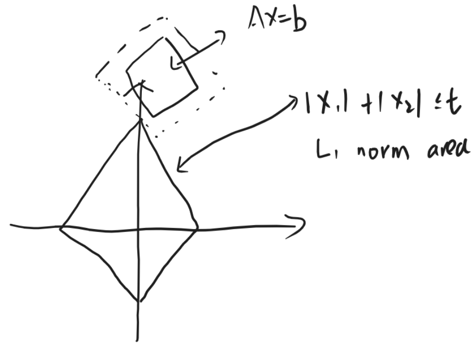

Figure 1.2: Approximation of Sparse Problem with 1-norm

This is a linear program with the null objective. Frequently we add the regularization term on the slope vector

$$
\begin{aligned}
\min \quad & \beta + \mu\|\mathbf{x}\|^2 \\
s.t. \quad & \mathbf{a}_i^T\mathbf{x} + x_0 + \beta \geq 1, \forall i \\
& \mathbf{b}_j^T\mathbf{x} + x_0 - \beta \leq -1, \forall j \\
& \beta \geq 0
\end{aligned}
\tag{1.8}
$$

The β is the error parameter. It is the relaxation method to compute the hyperplane. Moreover, we can imagine the separating boundary is the ellipsoid. Let recall the college knowledge

$$
\begin{aligned}
\frac{x^2}{a^2} + \frac{y^2}{b^2} &\leq 1 \\
\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{b^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &\leq 1
\end{aligned}
$$

Then we can have the ellipsoidal separation

$$
\begin{aligned}
& \min \operatorname{trace}(X) + \|\mathbf{x}\|^2 \\
\text{subject to } & \mathbf{a}_i^T X \mathbf{a}_i + \mathbf{a}_i^T\mathbf{x} + x_0 \geq 1, \forall i \\
& \mathbf{b}_j^T X \mathbf{b}_j + \mathbf{b}_j^T\mathbf{x} + x_0 \leq -1, \forall j \\
& X \succeq \mathbf{0}
\end{aligned}
\tag{1.9}
$$

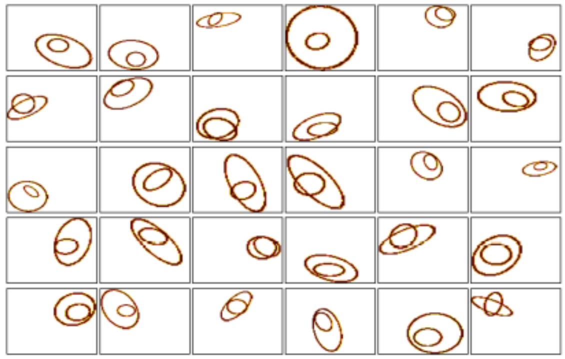

Figure 1.3: Sample

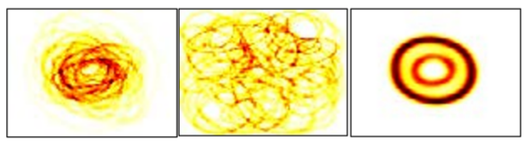

Figure 3: Mean picture constructed from the (a) Euclidean mean after re-centering images (b) Euclidean mean (c) Wasserstein Barycenter (self recenter, resize and rotate)

1.1.4 Transportation Problem

we consider the classic transportation problem. In the setting, there are seller and buyer, demand and supply respectively.

$$\begin{array}{lll} \min & \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} & \\ \text{s.t.} & \sum_{j=1}^{n} x_{ij} & = s_i, \forall i = 1,\dots,m \\ & \sum_{i=1}^{m} x_{ij} & = d_j, \forall j = 1,\dots,n \\ x_{ij} & \geq 0, \forall i,j. & \end{array} \tag{1.10}$$

The minimal transportation cost is called the Wasserstein Distance (WD) between supply distribution s and demand distribution d (can be interpreted as two probability distributions after normalization). This is a linear program.

The Wasserstein Barycenter Problem is to find a distribution such that the sum of its Wasserstein Distance to each of a set of distributions would be minimized. And you can check out this example 1.1.4 1.1.4 **Ref Appendices A, B, and C, Chapter 1**

1.2 Optimization Application

1.2.1 Graph Realization and Sensor Network Localization

Given a graph $G = (V, E)$ and sets of non–negative weights, say $\{d_{ij} : (i, j) \in E\}$, the goal is to compute a realization of G in the Euclidean space $\mathbb{R}^d$ for a given low dimension d, where the distance information is preserved

$$\begin{aligned} \|\mathbf{x}_i - \mathbf{x}_j\|^2 &= d_{ij}^2, \forall (i,j) \in N_x, i < j \\ \|\mathbf{a}_k - \mathbf{x}_j\|^2 &= \hat{d}_{kj}^2, \forall (k,j) \in N_a. \end{aligned} \tag{1.11}$$

This is the quadratic optimization

$$\min_{\mathbf{x}_i \forall i} \sum_{(i,j)\in N_x} \left(\|\mathbf{x}_i - \mathbf{x}_j\|^2 - d_{ij}^2\right)^2 + \sum_{(k,j)\in N_a} \left(\|\mathbf{a}_k - \mathbf{x}_j\|^2 - \hat{d}_{kj}^2\right)^2 \tag{1.12}$$

we have two directions to relax this problem: SOCP and SDP

1. change $=$ to $\leq$
2. SDP relaxation

   Let $X = [x_1 x_2 ... x_n]$ be the d $\times$ n matrix that needs to be determined and e_j be the vector of all zero except 1 at the jth position. Then

$$\|\mathbf{x}_i - \mathbf{x}_j\|^2 = (\mathbf{e}_i - \mathbf{e}_j)^T X^T X (\mathbf{e}_i - \mathbf{e}_j) \tag{1.13}$$

$$\begin{aligned} \|\mathbf{a}_k - \mathbf{x}_j\|^2 = (\mathbf{a}_k; -\mathbf{e}_j)^T \begin{bmatrix} I & X \end{bmatrix}^T \begin{bmatrix} I & X \end{bmatrix} (\mathbf{a}_k; -\mathbf{e}_j) = \\ (\mathbf{a}_k; -\mathbf{e}_j)^T \begin{pmatrix} I & X \\ X^T & X^T X \end{pmatrix} (\mathbf{a}_k; -\mathbf{e}_j). \end{aligned} \tag{1.14}$$

Convex relaxation first and steepest-descent-search second strategy

1.2.2 Markov Decision Process

- An MDP problem is defined by a given number of states, indexed by i, where each state has a number of actions, A_i , to take. Each action, say $j \in A_i$, is associated with an (immeidiate) cost $c_j$of taking, and a probability distribution p_j to transfer to all possible states at the next time period
- A stationary policy for the decision maker is a function $\pi = \{\pi_1, \pi_2, ..., \pi_m\}$ that specifies an action in each state,$\pi_i \in A_i$, that the decision maker will take at any time period; which also lead to a cost-to-go value for each state
- The MDP is to find a stationary policy to minimize/maximize the expected discounted sum over the infinite horizon with a discount factor $0 \leq \gamma \leq 1$

$$\sum_{t=0}^{\infty} \gamma^t E\left[c^{\pi_i t}\left(i^t, i^{t+1}\right)\right]. \tag{1.15}$$

- If the states are partitioned into two sets, one is to minimize and the other is to maximize the discounted sum, then the process becomes a two-person turn-based zero-sum stochastic game

Ref Appendix B, Chapters 3.1-2, 6.1-6.4

1.3 Global and Local Optimizer

A global minimizer for (P) is a vector $x^\star$ such that

$$\mathbf{x}^* \in \mathcal{X} \quad \text{and} \quad f\left(\mathbf{x}^*\right) \leq f(\mathbf{x}) \quad \forall \mathbf{x} \in \mathcal{X} \tag{1.16}$$

Sometimes one has to settle for a local minimizer, that is, a vector $\bar{x}$ such that

$$\overline{\mathbf{x}} \in \mathcal{X} \quad \text{and} \quad f(\overline{\mathbf{x}}) \leq f(\mathbf{x}) \quad \forall \mathbf{x} \in \mathcal{X} \cap N(\bar{x}) \tag{1.17}$$

where $N(\bar{x})$ is a neighborhood of $\bar{x}$. Typically, $N(\bar{x}) = B_\delta(\bar{x})$, an open ball centered at $\bar{x}$ having suitably small radius $\delta > 0$.

1.3.1 Size and Complexity of Problems

- number of decision variables
- number of constraints
- bit size/number required to store the problem input data (mean the memory consumption)
- problem difficulty or complexity number
- algorithm complexity or convergence speed

the definition of L_p norm is following

$$\|\mathbf{x}\|_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p} \tag{1.18}$$

For $x^k, x^\star \in R^n$ and $0 < \gamma < 1$ contraction sequence is

$$\left\|\mathbf{x}^{k+1} - \mathbf{x}^*\right\| \leq \gamma \left\|\mathbf{x}^k - \mathbf{x}^*\right\|, \forall k \geq 0 \tag{1.19}$$

inner product

$$A \bullet B = \operatorname{tr} A^T B = \sum_{i,j} a_{ij} b_{ij} \tag{1.20}$$

The operator norm of matrix A

$$\|A\|^2 := \max_{\mathbf{0} \neq \mathbf{x} \in \mathcal{R}^n} \frac{\|A\mathbf{x}\|^2}{\|\mathbf{x}\|^2} \tag{1.21}$$

The Frobenius norm of matrix A

$$\|A\|_f^2 := A \bullet A = \sum_{i,j} a_{ij}^2 \tag{1.22}$$

Theorem 1.4. *Perron-Frobenius Theorem: a real square matrix with positive entries has a unique largest real eigenvalue and that the corresponding eigenvector can be chosen to have strictly positive components.*

Stochastic Matrices: $A \leq 0$ with $e^T A = e^T$(Column-Stochastic), or $Ae = e$ (Row-Stochastic), or Doubly-Stochastic if both. It has a unique largest real eigenvalue 1 and corresponding non-negative right or left eigenvector.

1.3.2 Affine Set

$S \subset R^n$ is affine if

$$[\mathbf{x}, \mathbf{y} \in S \text{ and } \alpha \in R] \Longrightarrow \alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in S \tag{1.23}$$

When $\mathbf{x}$ and $\mathbf{y}$ are two distinct points in R^n and α runs over R,

$$\{\mathbf{z} : \mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}\} \tag{1.24}$$

is the affine combination of x and y. When $0 \leq \alpha \leq 1$, it is called the convex combination of $\mathbf{x}$ and $\mathbf{y}$. More points? For multipliers $\alpha \geq 0$ and for $\beta \geq 0$

$$\{\mathbf{z} : \mathbf{z} = \alpha\mathbf{x} + \beta\mathbf{y}\}, \tag{1.25}$$

is called the conic combination of x and y. It is called linear combination if both α and β are "free".

Convex Set

Ω is said to be a convex set if for every $\mathbf{x}^1, \mathbf{x}^2 \in \Omega$ and every real number $\alpha \in [0,1]$, the point $\alpha\mathbf{x}^1 + (1-\alpha)\mathbf{x}^2 \in \Omega$

Ball and Ellipsoid: for given $\mathbf{y} \in \mathcal{R}^n$ and positive definite matrix Q : $E(\mathbf{y}, Q) = \left\{\mathbf{x} : (\mathbf{x}-\mathbf{y})^T Q(\mathbf{x}-\mathbf{y}) \leq 1\right\}$

The intersection of convex sets is convex, the sum-set of convex sets is convex, the scaled-set of a convext set is convex

The convex hull of a set Ω is the intersection of all convex sets containing Ω. Given column-points of A, the convex hull is $\left\{\mathbf{z} = A\mathbf{x} : \mathbf{e}^T\mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}\right\}$

SVM Claim: two point sets are separable by a plane if any only if their convex hulls are separable.

An extreme point in a convex set is a point that cannot be expressed as a convex combination of other two distinct points of the set

A set is polyhedral if it has finitely many extreme points; $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ and $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ are convex polyhedral.

The dual norm

$$C^* := \{\mathbf{y} : \mathbf{x} \bullet \mathbf{y} \geq 0 \text{ for all } \mathbf{x} \in C\} \tag{1.26}$$

Theorem 1.5. *The dual is always a closed convex cone, and the dual of the dual is the closure of convex hall of C*

Cone Example

Example 1.6. *The n-dimensional non-negative orthant, $\mathcal{R}^n_+ = \{\mathbf{x} \in \mathcal{R}^n : \mathbf{x} \geq \mathbf{0}\}$, is a convex cone. Its dual is itself.*

Example 1.7. *The set of all PSD matrices in $\mathcal{S}^n, \mathcal{S}^n_+$, is a convex cone, called the PSD matrix cone. Its dual is itself.*

Example 1.8. *The set $\{(t;\mathbf{x}) \in \mathcal{R}^{n+1} : t \geq \|\mathbf{x}\|_p\}$ for $p \geq 1$ is a convex cone in $\mathcal{R}^{n+1}$, called the p-order cone. Its dual is the q-order cone with $\frac{1}{p} + \frac{1}{q} = 1$.*

The dual of the second-order cone ($p = 2$) is itself.

1.3.3 Convex Functions

f is a (strongly) convex function iff for $0 < \alpha < 1$,

$$f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y})(<) \leq \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y}) \tag{1.27}$$

The sum of convex functions is a convex function; the max of convex functions is a convex function

The Composed function $f(\phi(\mathbf{x}))$ is convex if $\phi(\mathbf{x})$ is a convex and $f(\cdot)$ is convex&non-decreasing. The (lower) level set of f is convex:

$$L(z) = \{\mathbf{x} : \quad f(\mathbf{x}) \leq z\} \tag{1.28}$$

Convex set $\{(z;\mathbf{x}) : f(\mathbf{x}) \leq z\}$ is called the epigraph of f. $tf(\mathbf{x}/t)$ is a convex function of $(t;\mathbf{x})$ for $t > 0$ if $f(\cdot)$ is a convex function; it's homogeneous with degree 1 Note that the difference between supreme and maximization, the maximal solution is achievable and supreme is not.

We need to clarify the function $tf(\mathbf{x}/t)$. It is jointly convex function of $(t;\mathbf{x})$. The jointly convex will be

$$(\alpha t_1 + (1-\alpha)t_2)\, f(\frac{\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2}{\alpha t_1 + (1-\alpha)t_2}) \leq \alpha t_1 f(\frac{\mathbf{x}_1}{t_1}) + (1-\alpha)t_2 f(\frac{\mathbf{x}_2}{t_2}) \tag{1.29}$$

The detail you can read chapter 2 of Boyd et al. (2004)

1.3.4 Convex Function Examples

$\|\mathbf{x}\|_p$ for $p \geq 1$ is convex because of the triangle inequality(norm property)

$$\|\alpha\mathbf{x} + (1-\alpha)\mathbf{y}\|_p \leq \|\alpha\mathbf{x}\|_p + \|(1-\alpha)\mathbf{y}\|_p \leq \alpha\|\mathbf{x}\|_p + (1-\alpha)\|\mathbf{y}\|_p, \tag{1.30}$$

Logistic function $\log\left(1 + e^{\mathbf{a}^T}\mathbf{x} + b\right)$ is convex. Consider the minimal-objective function of $\mathbf{b}$ for fixed A and $\mathbf{c}$

$$\begin{aligned} z(\mathbf{b}) := \min \quad & \mathbf{c}^T\mathbf{x} \\ s.t. \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{1.31}$$

Show that $z(\mathbf{b})$ is a convex function in b.

Proof. There are two separated cases

$$\begin{aligned} z(\mathbf{b}_1) := \min \quad & \mathbf{c}^T\mathbf{x}_1 & \qquad z(\mathbf{b}_2) := \min \quad & \mathbf{c}^T\mathbf{x}_2 \\ s.t. \quad & A\mathbf{x}_1 = \mathbf{b}_1 & s.t. \quad & A\mathbf{x}_2 = \mathbf{b}_2 \\ & \mathbf{x}_1 \geq \mathbf{0} & & \mathbf{x}_2 \geq \mathbf{0} \end{aligned} \tag{1.32}$$

we need to show

$$\mathbf{x} = \alpha\mathbf{x}_1^\star + (1-\alpha) * \mathbf{x}_2^\star$$

is also the feasible solution and $\mathbf{x}_1^\star, \mathbf{x}_2^\star$ are optimal solution of linear program respectively

$$\begin{aligned} z(\alpha\mathbf{b}_1 + (1-\alpha) * \mathbf{b}_2) &= \min_x \mathbf{c}^T\left(\alpha\mathbf{x}_1^\star + (1-\alpha) * \mathbf{x}_2^\star\right) \\ &\leq \mathbf{c}^T\left(\alpha\mathbf{x}_1^\star + (1-\alpha) * \mathbf{x}_2^\star\right) \\ &= \alpha \min_{\mathbf{x}_1} \mathbf{c}^T\mathbf{x}_1 + (1-\alpha)\min_{\mathbf{x}_2} \mathbf{c}^T\mathbf{x}_2 \end{aligned}$$

□

We are also curious the minimal-objective function of $\mathbf{c}$ for fixed A and $\mathbf{b}$

$$\begin{aligned} z(\mathbf{c}) := \min \quad & \mathbf{c}^T\mathbf{x} \\ s.t. \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{1.33}$$

Proof. There are two separated cases

$$\begin{aligned} z(\mathbf{c}_1) := \min \quad & \mathbf{c}_1^T\mathbf{x}_1 & \qquad z(\mathbf{c}_2) := \min \quad & \mathbf{c}_2^T\mathbf{x}_2 \\ s.t. \quad & A\mathbf{x}_1 = \mathbf{b} & s.t. \quad & A\mathbf{x}_2 = \mathbf{b} \\ & \mathbf{x}_1 \geq \mathbf{0} & & \mathbf{x}_2 \geq \mathbf{0} \end{aligned} \tag{1.34}$$

we have x is also feasible solution for two programming

$$\begin{aligned} \min_{\mathbf{x}_1} \alpha\mathbf{c}_1^T\mathbf{x}_1 + \min_{\mathbf{x}_2}(1-\alpha)\mathbf{c}_2^T\mathbf{x}_2 &\leq \alpha\mathbf{c}_1^T\mathbf{x} + (1-\alpha)\mathbf{c}_2^T\mathbf{x} \\ &= \min_{\mathbf{x}}\left(\alpha\mathbf{c}_1^T + (1-\alpha)\mathbf{c}_2^T\right)\mathbf{x} \end{aligned}$$

Since we can always find optimal solution of $\left(\alpha\mathbf{c}_1^T + (1-\alpha)\mathbf{c}_2^T\right)\mathbf{x}$ □

Lemma 1.9. *For an standard linear program, if we fix A and $\mathbf{c}$, function $z(\mathbf{b})$ is the convex function. If we fix A and $\mathbf{b}$, function $z(\mathbf{c})$ is the concave function*

1.4 Theorems on Functions

Theorem 1.10. *Let $f \in C^1$ be in a region containing the line segment $[\mathbf{x}, \mathbf{y}]$. Then there is a $\alpha, 0 \le \alpha \le 1$, such that*

$$f(\mathbf{y}) = f(\mathbf{x}) + \nabla f(\alpha \mathbf{x} + (1-\alpha)\mathbf{y})(\mathbf{y} - \mathbf{x}) \tag{1.35}$$

Furthermore, if $f \in C^2$ then there is a $\alpha, 0 \le \alpha \le 1$, such that

$$f(\mathbf{y}) = f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{y} - \mathbf{x}) + (1/2)(\mathbf{y} - \mathbf{x})^T \nabla^2 f(\alpha \mathbf{x} + (1-\alpha)\mathbf{y})(\mathbf{y} - \mathbf{x}) \tag{1.36}$$

Theorem 1.11. *Let $f \in C^1$. Then f is convex over a convex set Ω if and only if*

$$f(\mathbf{y}) \ge f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{y} - \mathbf{x}) \tag{1.37}$$

for all $\mathbf{x}, \mathrm{y} \in \Omega$

Proof. (Boyd et al. (2004))Proof of first-order convexity condition To prove, we first consider the case $n = 1:$ We show that a differentiable function $f : \mathbf{R} \to \mathbf{R}$ is convex if and only if

$$f(y) \ge f(x) + f'(x)(y - x)$$

for all x and y in $\mathrm{dom}\, f$. Assume first that f is convex and $x, y \in \mathrm{dom}\, f$. Since $\mathrm{dom}\, f$ is convex (i.e., an interval), we conclude that for all $0 < t \le 1, x + t(y - x) \in \mathrm{dom}\, f$, and by convexity of f

$$f(x + t(y - x)) \le (1 - t)f(x) + tf(y).$$

If we divide both sides by t, we obtain

$$f(y) \ge f(x) + \frac{f(x + t(y - x)) - f(x)}{t},$$

and taking the limit as $t \to 0$ yields 4.25. To show sufficiency, assume the function satisfies 4.25 for all x and y in dom f (which is an interval). Choose any $x \ne y$, and $0 \le \theta \le 1$, and let $z = \theta x + (1 - \theta)y$. Applying 4.25 twice yields

$$f(x) \ge f(z) + f'(z)(x - z), \quad f(y) \ge f(z) + f'(z)(y - z).$$

Multiplying the first inequality by θ, the second by $1 - \theta$, and adding them yields

$$\theta f(x) + (1 - \theta)f(y) \ge f(z),$$

which proves that f is convex. Now we can prove the general case, with $f : \mathbf{R}^n \to \mathbf{R}$. Let $x, y \in \mathbf{R}^n$ and consider f restricted to the line passing through them, i.e., the function defined by $g(t) = f(ty + (1 - t)x)$, so $g'(t) = \nabla f(ty + (1 - t)x)^T (y - x)$.

First assume f is convex, which implies g is convex, so by the argument above we have $g(1) \ge g(0) + g'(0)$, which means

$$f(y) \ge f(x) + \nabla f(x)^T (y - x).$$

Now assume that this inequality holds for any x and y, so if $ty+(1-t)x \in \operatorname{dom} f$ and $\tilde{t}y+(1-t)x \in \operatorname{dom} f$, we have

$$f(ty+(1-t)x) \geq f(\tilde{t}y+(1-\tilde{t})x)+\nabla f(\tilde{t}y+(1-\tilde{t})x)^T(y-x)(t-\tilde{t}),$$

i.e., $g(t) \geq g(\tilde{t})+g'(\tilde{t})(t-\tilde{t})$. We have seen that this implies that g is convex. □

Theorem 1.12. *Let $f \in C^2$. Then f is convex over a convex set Ω if and only if the Hessian matrix of f is positive semi-definite throughout Ω.*

$$\nabla^2 \boldsymbol{f}(\boldsymbol{x}) \succeq 0 \tag{1.38}$$

1.4.1 Lipschitz Functions

The first-order β-Lipschitz function: there is a positive number β such that for any two points $\mathbf{x}$ and y :

$$\|\nabla f(\mathbf{x})-\nabla f(\mathbf{y})\| \leq \beta\|\mathbf{x}-\mathbf{y}\| \tag{1.39}$$

This condition imples

$$\left|f(\mathbf{x})-f(\mathbf{y})-\nabla f(\mathbf{y})^T(\mathbf{x}-\mathbf{y})\right| \leq \frac{\beta}{2}\|\mathbf{x}-\mathbf{y}\|^2 \tag{1.40}$$

Proof. The key tool is Taylor's formula with integral remainder

Let $\Delta := y-x$ and $\phi(t)=f(x+t\Delta)$ where t is a scalar variable. Then we have $\phi(0)=f(x)$ and $\phi(1)=f(x+\Delta)=f(y)$. Moreover,

$$f(x+\Delta)-f(x)=\phi(1)-\phi(0)=\int_0^1 \mathrm{d}\phi(t)=\int_0^1 \Delta^T\nabla f(x+t\Delta)\mathrm{d}t$$

For the first implication inequality, noting $\Delta^T\nabla f(x)=\int_0^1 \Delta^T\nabla f(x)\mathrm{d}t$, we have

$$\begin{aligned}
&\left|f(x+\Delta)-f(x)-\nabla f(x)^T\Delta\right| = \left|\int_0^1 \Delta^T(\nabla f(x+t\Delta)-\nabla f(x))\mathrm{d}t\right| \\
&\leq \int_0^1 \left|\Delta^T(\nabla f(x+t\Delta)-\nabla f(x))\right|\mathrm{d}t \\
&\leq \int_0^1 \|\Delta\|\|\nabla f(x+t\Delta)-\nabla f(x)\|\mathrm{d}t \quad \text{(Cauchy-Schwartz inequality)} \\
&= \|\Delta\|\int_0^1 \|\nabla f(x+t\Delta)-\nabla f(x)\|\mathrm{d}t \\
&\leq \|\Delta\|\int_0^1 \beta\|t\Delta\|\mathrm{d}t \quad \text{(the first-order Lipschitz condition)} \\
&= \|\Delta\|\beta\|\Delta\|\int_0^1 t\ \mathrm{d}t = \frac{\beta}{2}\|\Delta\|^2.
\end{aligned}$$

□

The second-order β-Lipschitz function: there is a positive number β such that for any two points $\mathbf{x}$ and $\mathbf{y}$

$$\left\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y}) - \nabla^2 f(\mathbf{y})(\mathbf{x}-\mathbf{y})\right\| \leq \beta\|\mathbf{x}-\mathbf{y}\|^2 \tag{1.41}$$

This condition implies

$$\left|f(\mathbf{x}) - f(\mathbf{y}) - \nabla f(\mathbf{y})^T(\mathbf{x}-\mathbf{y}) - \frac{1}{2}(\mathbf{x}-\mathbf{y})^T\nabla^2 f(\mathbf{y})(\mathbf{x}-\mathbf{y})\right| \leq \frac{\beta}{3}\|\mathbf{x}-\mathbf{y}\|^3 \tag{1.42}$$

Proof. Similar idea we use to prove the first order Lipschitz function

Let $\Delta = x - y$ and $\phi(t) = f(y + t\Delta)$ where t is a scalar variable. Then we have $\phi(0) = f(y)$ and $\phi(1) = (y + \Delta) = f(x)$. Moreover,

$$\nabla f(y+\Delta) - \nabla f(y) = \phi(1) - \phi(0) = \int_0^1 \,\mathrm{d}\phi(t) = \int_0^1 \Delta^T\nabla^2 f(x+t\Delta)\mathrm{d}t$$

For the first implication inequality, noting $\Delta^T\nabla^2 f(x) = \int_0^1 \Delta^T\nabla^2 f(x)\mathrm{d}t$ and $\frac{1}{2}(\mathbf{x}-\mathbf{y})^T\nabla^2 f(\mathbf{y})(\mathbf{x}-\mathbf{y}) = \Delta^T\nabla^2 f(\mathbf{y})\Delta\int_0^1 t\mathrm{d}t$, we have

$$\begin{aligned}
\left|f(\mathbf{y}+\boldsymbol{\Delta}) - f(\mathbf{y}) - \nabla f(\mathbf{y})^T\Delta - \frac{\Delta^T\nabla^2 f(\mathbf{y})\Delta}{2}\right| &= \left|\int_0^1 \Delta^T(\nabla f(x+t\Delta) - \nabla f(x))\mathrm{d}t\right| \\
&\leq \int_0^1 \left|\Delta^T(\nabla f(y+t\Delta) - \nabla f(y) - t\nabla^2 f(\mathbf{y})\Delta)\right|\mathrm{d}t \\
&\leq \int_0^1 \|\Delta\|\|\nabla f(x+t\Delta) - \nabla f(x) - t\nabla^2 f(\mathbf{y})\Delta\|\mathrm{d}t \\
&= \|\Delta\|\int_0^1 \|\nabla f(x+t\Delta) - \nabla f(x) - t\nabla^2 f(\mathbf{y}\|\mathrm{d}t \\
&\leq \|\Delta\|\int_0^1 \beta\|t\Delta\|^2\mathrm{d}t \\
&= \|\Delta\|\beta\|\Delta\|^2\int_0^1 t\ \mathrm{d}t^2 \\
&= \frac{\beta}{3}\|\Delta\|^3
\end{aligned}$$

□

1.4.2 Known Inequalities

- Cauchy-Schwarz: given $\mathbf{x}, \mathbf{y} \in \mathcal{R}^n$, $\left|\mathbf{x}^T\mathbf{y}\right| \leq \|\mathbf{x}\|_p\|\mathbf{y}\|_q$, where $\frac{1}{p} + \frac{1}{q} = 1$ and $p \geq 1$
- Triangle: given $\mathbf{x}, \mathbf{y} \in \mathcal{R}^n$, $\|\mathbf{x}+\mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$ for $p \geq 1$
- Arithmetic-geometric mean: given $\mathbf{x} \in \mathcal{R}^n_+$,

$$\frac{\sum x_j}{n} \geq \left(\prod x_j\right)^{1/n}$$

1.4.3 Direct Solution

Given $A \in \mathcal{R}^{m \times n}$ and $\mathbf{c} \in \mathcal{R}^n$

$$\begin{aligned} \min \quad & \left\| A^T\mathbf{y} - \mathbf{c} \right\|^2 \\ s.t. \quad & \mathbf{y} \in \mathcal{R}^m \end{aligned} \tag{1.43}$$

Choleski Decomposition

$$AA^T = L\Lambda L^T, \text{ and then solve } L\Lambda L^T\mathbf{y} = A\mathbf{c}.$$

Projections Matrices: $A^T \left(AA^T\right)^{-1} A$ and $I - A^T \left(AA^T\right)^{-1} A$

1.5 Exercise

1. Given a symmetric matrix $A \in R^{n\times n}$ s.t. A has eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, show that for every $k = 1, 2, \cdots, n$, we have:

$$\begin{aligned} \lambda_k &= \max_U \left\{ \min_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \mid \mathbf{x} \in U, \mathbf{x} \neq \mathbf{0} \right\} \mid U \text{ is a linear subspace of} R^n \text{of dimension } k \right\} \\ &= \min_U \left\{ \max_{\mathbf{x}} \left\{ \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \mid \mathbf{x} \in U, \mathbf{x} \neq \mathbf{0} \right\} \mid U \text{ is a linear subspace of } R^n \text{ of dimension } n-k+1 \right\} \end{aligned} \tag{1.44}$$

Solution: This result is known as the Courant-Fischer Minimax Theorem. See Theorem 8.1.2 of Golub and Van Loan (2013) for a sample proof

Here we sketch the proof for 4.25. Let $\{v_k\}_{k=1}^n$ denote a set of orthonormal eigenbasis of A, with $Av_k = \lambda_k v_k$. Moreover, $A = \sum_{k=1}^n \lambda_k v_k v_k^T$.

When $k = 1$, the expression reduces to $\lambda_1 = \max_{x\neq 0} \frac{x^T Ax}{x^T x}$, which is true for symmetric matrices, with one maximizer U^1 being spanned by v_1.

Now suppose for the sake of induction that we have shown (1) for some k and that the maximizer U^k can be taken to be the span of the first k eigenvectors, and we need to show it holds for $k+1$ We show that a maximizer for λ_{k+1} is $U^{k+1} := U^k \cup \operatorname{span}(v_{k+1})$. To see this, note that

$$\lambda_{k+1} = \min_{x\in U^{k+1}} \frac{x^T Ax}{x^T x}$$

so that $\lambda_{k+1} \leq RHS$ of (1). On the other hand, for any subspace U of dimension $k+1$ that is not spanned by the first $k+1$ eigenvectors of A, minimization in RHS will choose an eigenvector corresponding to an eigenvalue smaller than λ_{k+1}. □

2. Given symmetric matrices $A, B, C \in R^{n\times n}$ s.t. A has eigenvalues $a_1 \geq a_2 \geq \cdots \geq a_n$, B has eigenvalues $b_1 \geq b_2 \geq \cdots \geq b_n$ and C has eigenvalues $c_1 \geq c_2 \geq \cdots \geq c_n$, if $A = B + C$, show that for every $k = 1, 2, \cdots, n$, we have:

$$b_k + c_n \leq a_k \leq b_k + c_1$$

Solution: We show that $a_k \leq b_k + c_1$. The other inequality is similar. According to (1), define U_k to be the dim- k linear subspace such that

$$a_k = \min_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in U_k, \boldsymbol{x} \neq \mathbf{0} \right\}$$

and let $\boldsymbol{x}^*$ be the minimizer of $\min_{\boldsymbol{x}} \left\{ \frac{x^T B x}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in U_k, \boldsymbol{x} \neq \mathbf{0} \right\}$. It follows that

$$\begin{aligned}
a_k &= \min_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T (B+C) \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in U_k, \boldsymbol{x} \neq \mathbf{0} \right\} \leq \frac{\boldsymbol{x}^{*T} (B+C) \boldsymbol{x}^*}{\boldsymbol{x}^{*T} \boldsymbol{x}^*} \\
&\leq \min_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T B \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in U_k, \boldsymbol{x} \neq \mathbf{0} \right\} + \max_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T C \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in R^n, \boldsymbol{x} \neq \mathbf{0} \right\} \\
&\leq \max_{U} \left\{ \min_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T B \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in U, \boldsymbol{x} \neq \mathbf{0} \right\} \mid \dim(U) = k \right\} + \max_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T C \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in R^n, \boldsymbol{x} \neq \mathbf{0} \right\} \\
&= b_k + c_1,
\end{aligned}$$

completing the proof. Similarly, According to 4.25, define U_k to be the $(n-k+1)$-dimensional linear subspace such that

$$a_k = \max_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in U_k, \boldsymbol{x} \neq \mathbf{0} \right\}$$

and let x^* be the maximizer of $\max_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T B \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in U_k, \boldsymbol{x} \neq \mathbf{0} \right\}$. It follows that

$$\begin{aligned}
a_k &= \max_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T (B+C) \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in U_k, \boldsymbol{x} \neq \mathbf{0} \right\} \geq \frac{\boldsymbol{x}^{*T} (B+C) \boldsymbol{x}^*}{\boldsymbol{x}^{*T} \boldsymbol{x}^*} \\
&\geq \max_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T B \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in U_k, \boldsymbol{x} \neq \mathbf{0} \right\} + \min_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T C \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in R^n, \boldsymbol{x} \neq \mathbf{0} \right\} \\
&\geq \min_{U} \left\{ \max_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T B \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in U, \boldsymbol{x} \neq \mathbf{0} \right\} \mid \dim(U) = n-k+1 \right\} + \min_{\boldsymbol{x}} \left\{ \frac{\boldsymbol{x}^T C \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \mid \boldsymbol{x} \in R^n, \boldsymbol{x} \neq \mathbf{0} \right\} \\
&= b_k + c_n
\end{aligned}$$

$\square$

Chapter 2

Duality

Ref Chapters 3.1-3.2 and 6.1-6.4

2.1 Basic Feasible Solution and Farkas Lemma

2.1.1 Caratheodory's theorem

Theorem 2.1. *Given matrix $A \in R^{m\times n}$, let convex polyhedral cone $C = \{A\mathbf{x} : \mathbf{x} \geq 0\}$. For any $\mathbf{b} \in C$*

$$\mathbf{b} = \sum_{i=1}^{d} \mathbf{a}_{j_i} x_{j_i}, x_{j_i} \geq 0, \forall i \tag{2.1}$$

for some linearly independent vectors $\mathbf{a}_{j_1}, \ldots, \mathbf{a}_{j_d}$ chosen from $\mathbf{a}_1, \ldots, \mathbf{a}_n$. There is a construct proof of the theorem (page 26 of the Luenberger et al. (1984)).

2.1.2 Basic Feasible Solution

Now consider the feasible set $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ for given data $A \in R^{m\times n}$ and $\mathbf{b} \in R^m$. Select m linearly independent columns, denoted by the variable index set B, from A. Solve $A_B\mathbf{x}_B = \mathbf{b}$ for the m-dimension vector $\mathbf{x}_B$, and set the remaining variables, $\mathbf{x}_N$, to zero. Then, we obtain a solution $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$, that is called a basic solution to with respect to the basis A_B. If a basic solution $\mathbf{x}_B \geq \mathbf{0}$, then x is called a basic feasible solution, or BFS.

Note that the the optimal solution is the extreme point. We show the proof of this argument below

Proof.

$$\begin{aligned} \min \quad & c^T x \\ s.t. \quad & x \in X \end{aligned} \tag{2.2}$$

Let $\dot{x} \in \text{int X}$. Thus we have $\{x : ||x - \dot{x}||_2 \leq r\} \in X$. Imply

$$\begin{aligned} x &= \dot{x} - r\frac{c}{||c||_2} \\ c^T x &= c^T(\dot{x} - r\frac{c}{||c||_2}) \\ &= c^T\dot{x} - r||c|| \\ &< c^T\dot{x} \end{aligned}$$

□

Theorem 2.2. *(Separating hyperplane theorem) Let C be a closed convex set in $\mathcal{R}^m$ and let b be a point exterior to C. Then there is a vector* $\mathrm{y} \in \mathcal{R}^m$ *such that*

$$\mathbf{b} \cdot \mathbf{y} > \sup_{\mathbf{x} \in C} \mathbf{x} \cdot \mathbf{y}. \tag{2.3}$$

Theorem 2.3. *(Supporting hyperplane theorem) Let C be a closed convex set and let* b *be a point on the boundary of C. Then there is a vector $y \in \mathcal{R}^m$ such that*

$$\mathbf{b} \cdot \mathbf{y} = \sup_{\mathbf{x} \in C} \mathbf{x} \cdot \mathbf{y}. \tag{2.4}$$

2.1.3 Farkas Lemma

Theorem 2.4. *Let $A \in \mathcal{R}^{m \times n}$ and* $\mathrm{b} \in \mathcal{R}^m$*. Then, the system* $\{\mathrm{x} : A\mathrm{x} = \mathrm{b}, \mathrm{x} \geq 0\}$ *has a feasible solution* x *if and only if that its alternative system $-A^T\mathbf{y} \geq 0$ and $\mathbf{b}^T\mathbf{y} > 0$ has no feasible solution* $\mathbf{y}$.

Geometrically, Farkas' lemma means that if a vector $b \in \mathcal{R}^m$ does not belong to the convex cone generated by $\mathbf{a}_{.1}, \ldots, \mathbf{a}_{.n}$, then there is a hyperplane separating $\mathbf{b}$ from cone $(\mathbf{a}_{.1}, \ldots, \mathbf{a}_{.n})$.

Proof. Let $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ have a feasible solution, say $\overline{\mathbf{x}}$. Then, $\left\{\mathbf{y} : A^T\mathbf{y} \leq \mathbf{0}, \mathbf{b}^T\mathbf{y} > 0\right\}$ is infeasible, since otherwise,

$$0 < \mathbf{b}^T\mathbf{y} = (A\mathbf{x})^T\mathbf{y} = \mathbf{x}^T\left(A^T\mathbf{y}\right) \leq 0$$

from $\mathbf{x} \geq \mathbf{0}$ and $A^T\mathbf{y} \leq \mathbf{0}$ Now let $\{\mathrm{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ have no feasible solution, or $\mathbf{b} \notin C := \{A\mathbf{x} : \mathbf{x} \geq 0\}$. We now prove that its alternative system has a solution. We first prove

Lemma 2.5. *$C = \{A\mathbf{x} : \mathbf{x} \geq \mathbf{0}\}$ is a closed convex set.*

That is, any convergent sequence $\mathbf{b}^k \in C, k = 1.2\ldots$ has its limit point $\bar{\mathrm{b}}$ also in C. Let $\mathbf{b}^k = A\mathbf{x}^k, \mathbf{x}^k \geq \mathbf{0}$. Then by Carathéodory's theorem, we must have $\mathbf{b}^k = A_{B^k}\mathbf{x}_{B^k}, \mathbf{x}_{B^k} \geq \mathbf{0}$ where A_{B^k} is a basis of A. Therefore, $\mathbf{x}_{B^k}$, together with zero values for the nonbasic variables, is bounded for all k, so that it has sub-sequence, say indexed by $l = 1, \ldots$, where $\mathbf{x}^l = \mathbf{x}_{B^l}$ has a limit point $\bar{\mathbf{x}}$ and $\bar{\mathbf{x}} \geq 0$. Consider this very sub-sequence $\mathbf{b}^l = A\mathbf{x}^l$ we must also have $\mathbf{b}^l \to \bar{\mathbf{b}}$. Then from

$$\|\bar{\mathbf{b}} - A\bar{\mathbf{x}}\| = \left\|\bar{\mathbf{b}} - \mathbf{b}^l + A\mathbf{x}^l - A\bar{\mathbf{x}}\right\| \leq \left\|\bar{\mathbf{b}} - \mathbf{b}^l\right\| + \left\|A\mathbf{x}^l - A\bar{\mathbf{x}}\right\| \leq \left\|\bar{\mathbf{b}} - \mathbf{b}^l\right\| + \|A\| \left\|\mathbf{x}^l - \bar{\mathbf{x}}\right\|$$

we must have $\bar{\mathrm{b}} = A\bar{\mathrm{x}}$, that is, $\bar{\mathrm{b}} \in C$; since otherwise the right-hand side of the above inequality is strictly greater than zero which is a contradiction. Now since C is a closed convex set, by the separating hyperplane theorem, there is y such that

$$\mathbf{y} \cdot \mathbf{b} > \sup_{\mathbf{c} \in C} \mathbf{y} \bullet \mathbf{c}$$

or

$$\mathbf{y} \bullet \mathbf{b} > \sup_{\mathbf{x} \geq \mathbf{0}} \mathbf{y} \bullet (A\mathbf{x}) = \sup_{\mathbf{x} \geq \mathbf{0}} A^T\mathbf{y} \bullet \mathbf{x} \tag{2.5}$$

From $0 \in C$ we have $\mathrm{y} \bullet > 0$. Furthermore, $A^T\mathbf{y} \leq \mathbf{0}$. Since otherwise, say $\left(A^T\mathbf{y}\right)_1 > 0$, one can have a vector $\bar{\mathbf{x}} \geq 0$ such that $\bar{x}_1 = \alpha > 0, \bar{x}_2 = \ldots = \bar{x}_n = 0$, from which

$$\sup_{\mathbf{x} \geq 0} A^T\mathbf{y} \bullet \mathbf{x} \geq A^T\mathbf{y} \bullet \bar{\mathbf{x}} = \left(A^T\mathbf{y}\right)_1 \cdot \alpha$$

and it tends to ∞ as $\alpha \to \infty$. This is a contradiction because $\sup_{\mathbf{x} \geq 0} A^T\mathbf{y} \bullet \mathbf{x}$ is bounded from above by (5). $\square$

Farkas Lemma Variant

Theorem 2.6. *Let $A \in \mathcal{R}^{m \times n}$ and $\mathrm{c} \in \mathcal{R}^n$. Then, the system $\left\{\mathbf{y} : \mathbf{c} - A^T\mathbf{y} \geq 0\right\}$ has a solution $\mathbf{y}$ if and only if that $A\mathrm{x} = 0, \mathrm{x} \geq 0$, and $\mathrm{c}^T\mathrm{x} < 0$ has no feasible solution x.*

Consider the pair:

$$\{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in K\}$$

and

$$\left\{\mathbf{y} : -A^T\mathbf{y} \in K^*, \quad \mathbf{b}^T\mathbf{y} > 0\right\}.$$

Or in operator form: given data vector or matrix $\mathbf{a}_i, i = 1, \ldots, m$, and $\mathbf{b} \in \mathcal{R}^m$, an "alternative" system pair would be

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in K$$

and

$$-\mathcal{A}^T\mathbf{y} \in K^*, \quad \mathbf{b}^T\mathbf{y} = 1(> 0)$$

where

$$\mathcal{A}\mathbf{x} = (\mathbf{a}_1 \bullet \mathbf{x}; \ldots; \mathbf{a}_m \bullet \mathbf{x}) \in \mathcal{R}^m \text{ and } \mathcal{A}^T\mathbf{y} = \sum_i^m y_i\mathbf{a}_i$$

Let K be a closed and convex cone in the rest of the course. If there is y such that $-\mathcal{A}^T\mathbf{y} \in \text{int}\, K^*$, then $C := \{\mathcal{A}\mathbf{x} : \mathbf{x} \in K\}$ is a closed convex cone. Consequently,

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in K$$

and

$$-\mathcal{A}^T\mathbf{y} \in K^*, \quad \mathbf{b}^T\mathbf{y} = 1(>0)$$

are an alternative system pair. And if there is $\mathbf{x}$ such that $\mathcal{A}^T\mathbf{x} = \mathbf{0}, \mathbf{x} \in \text{int}\, K$, then

$$\mathcal{A}\mathbf{x} = \mathbf{0}, \quad \mathbf{x} \in K, \quad \mathbf{c} \bullet \mathbf{x} = -1(<0)$$

and

$$\mathbf{c} - \mathcal{A}^T\mathbf{y} \in K^*$$

are an alternative system pair.

2.2 Conic Linear Program

$$\begin{aligned} (\text{C L P}) \min \quad & \mathbf{c} \bullet \mathbf{x} \\ s.t. \quad & \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, \ldots, m, \mathbf{x} \in K, \\ & \left(\mathcal{A}^T\mathbf{x} = \mathbf{b}\right) \end{aligned} \tag{2.6}$$

where K is a closed and pointed convex cone. Linear Programming (LP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $K = \mathcal{R}^n_+$ Second-Order Cone Programming (SOCP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $K = SOC = \{\mathbf{x} : x_1 \geq \|\mathbf{x}_{-1}\|_2\}$. Semidefinite Programming (SDP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{S}^n$ and $K = \mathcal{S}^n_+$ p-Order Cone Programming (POCP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $K = POC = \left\{\mathbf{x} : x_1 \geq \|\mathbf{x}_{-1}\|_p\right\}$. Here, $\mathbf{x}_{-1}$ is the vector $(x_2; \ldots; x_n) \in R^{n-1}$. Cone K can be also a product of different cones, that is, $\mathbf{x} = (\mathbf{x}_1; \mathbf{x}_2; \ldots)$ where $\mathbf{x}_1 \in K_1, \mathbf{x}_2 \in K_2, \ldots$ and so on with linear constraints:

$$\mathcal{A}_1\mathbf{x}_1 + \mathcal{A}_2\mathbf{x}_2 + \ldots = \mathbf{b}$$

2.2.1 Dual of Conic Linear Program

The dual problem to

$$\begin{aligned} (CLP) \min \quad & \mathbf{c} \bullet \mathbf{x} \\ s.t. \quad & \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, \ldots, m, \mathbf{x} \in K \end{aligned} \tag{2.7}$$

is

$$\begin{aligned} (CLD) \quad \min \quad & \mathbf{b}^T\mathbf{y} \\ s.t. \quad & \textstyle\sum_i^m y_i\mathbf{a}_i + \mathbf{s} = \mathbf{c}, \mathbf{s} \in K^*, \end{aligned} \tag{2.8}$$

where $y \in \mathcal{R}^m$, s is called the dual slack vector/matrix, and K^* is the dual cone of K. The former is called the primal problem, and the latter is called dual problem.

Ref Chapters 3.1-2, 6.4-5

2.3 Duality of Conic Linear Program

Recall the primal and dual program

$$\begin{array}{rlrl} \min & C^T x & \max & b^T y \\ s.t. & Ax = b & s.t. & A^y + s = c \\ & x \in K & & s \in K' \end{array} \tag{2.9}$$

Theorem 2.7. *(Weak duality theorem)* $\mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} = \mathbf{x}^T\mathrm{s} \geq 0$ *for any feasible* $\mathbf{x}$ *of* (CLP) *and* $(\mathbf{y}, \mathrm{s})$ *of* (CLD)

Corollary 2.8. *Let* $\mathrm{x}^* \in \mathcal{F}_p$ *and* $(\mathbf{y}^*,\ \mathrm{s}^*) \in \mathcal{F}_d$. *Then,* $\mathbf{c} \bullet \mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$ *implies that* x^* *is optimal for* (CLP) *and* $(\mathbf{y}^*,\ \mathrm{s}^*)$ *is optimal for* (CLD)

It is called the strong duality theorem, but it does not work in general. Here, operator $\mathcal{A}\mathbf{x}$ and Adjoint-Operator $\mathcal{A}^T\mathbf{y}$ minimic matrix-vector production $A\mathbf{x}$ and its transpose operation $A^T\mathbf{y}$, where

$$\mathcal{A} = (\mathbf{a}_1; \mathbf{a}_2; \ldots; \mathbf{a}_m), \quad \mathcal{A}\mathbf{x} = (\mathbf{a}_1 \bullet \mathbf{x}; \ldots; \mathbf{a}_m \bullet \mathbf{x}), \quad \text{and} \quad A^T\mathbf{y} = \sum_i y_i \mathbf{a}_i^T \tag{2.10}$$

Theorem 2.9. *The following statements hold for every pair of* (LP) *and (LD) :*

- *If* (LP) *and (LD) are both feasible, then both problems have optimal solutions and the optimal objective values of the objective functions are equal, that is, optimal solutions for both (LP) and (LD) exist and there is no duality gap*
- *If* (LP) *or* (LD) *is feasible and bounded, then the other is feasible and bounded*
- *If* (LP) *or* (LD) *is feasible and unbounded, then the other has no feasible solution*
- *If* (LP) *or* (LD) *is infeasible, then the other is either unbounded or has no feasible solution*

2.3.1 Farkas Lemma and Duality

The Farkas lemma concerns the system the system $\{\mathrm{x} : A\mathrm{x} = \mathbf{b}, \mathrm{x} \geq 0\}$ and its alternative $\left\{\mathbf{y} : -A^T\mathbf{y} \geq \mathbf{0}, \mathbf{b}^T\mathbf{y} > 0\right\}$ for given data $(A, \mathbf{b})$. This pair can be represented as a primal-dual LP pair

$$\begin{array}{rlrl} \min & \mathbf{0}^T\mathbf{x} & \max & \mathbf{b}^T\mathbf{y} \\ \text{s. t.} & A\mathbf{x} = \mathbf{b}, & \text{s.t.} & A^T\mathbf{y} \leq \mathbf{0} \\ & \mathbf{x} \geq \mathbf{0}; & & \end{array} \tag{2.11}$$

If the primal is infeasible, then the dual must be feasible and unbounded since it is always feasible.

2.3.2 Optimality Conditions for LP

$$\begin{cases} & C^T x - b^T y = 0 \\ (\mathbf{x}, \mathrm{y}, \mathrm{s}) \in (\mathcal{R}^n_+, \mathcal{R}^m, \mathcal{R}^n_+) & Ax = b \\ & A^T y + s = c \end{cases} \tag{2.12}$$

which is a system of linear inequalities and equations. Now it is easy to verify whether or not a pair $(\mathbf{x}, \mathrm{y}, \mathrm{s})$ is optimal.

2.3.3 Complementarity Condition

For feasible $\mathbf{x}$ and $(\mathbf{y}, \mathbf{s}), \mathbf{x}^T\mathbf{s} = \mathbf{x}^T\left(\mathbf{c} - A^T\mathbf{y}\right) = \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y}$ is called the complementarity gap. If $\mathbf{x}^T\mathbf{S} = 0$, then we say x and s are complementary to each other. Since both $\mathbf{x}$ and $\mathbf{s}$ are nonnegative, $\mathbf{x}^T\mathbf{s} = 0$ implies that $\mathbf{x}. * \mathbf{s} = 0$ or $x_j s_j = 0$ for all $j = 1, \ldots, n$.

$$\begin{aligned} \mathrm{x} \cdot * &= 0 \\ A\mathrm{x} &= \mathbf{b} \\ -A^T\mathbf{y} - \mathrm{s} &= -\mathbf{c}. \end{aligned} \tag{2.13}$$

This system has total $2n + m$ unknowns and $2n + m$ equations including n nonlinear equations. Interpretation of $s_j = 0:$ the j th inequality constraint of the dual is "binding" or "active".

2.3.4 Duality of Conic Program

Example 2.10. *The strong duality theorem may not hold for general convex cones:*

$$\mathbf{c} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{a}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

and

$$\mathbf{b} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

The problem is

$$\begin{array}{llll} \min & x_1 + x_2 & \max & y_2 \\ s.t. & \begin{bmatrix} x_2 \\ x_1 + x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \end{bmatrix} & s.t. & \begin{bmatrix} 0 & -y_2 & 0 \\ -y_2 & y_1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + s = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ & \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \succeq 0 & & s \succeq 0 \end{array}$$

Theorem 2.11. *The following statements hold for every pair of* (CLP) *and (CLD):*

- *If (CLP) and (CLD) both are feasible, and furthermore one of them have an interior, then there is no duality gap between (CLP) and (CLD). However, one of the optimal solution may not be attainable.*

- *If (CLP) and (CLD) both are feasible and have interior, then, then both have attainable optimal solutions with no duality gap.*

- *If (CLP) or (CLD) is feasible and unbounded, then the other has no feasible solution.*

- *If (CLP) or (CLD) is infeasible, and furthermore the other is feasible and has an interior, then the other is unbounded.*

Construct the Dual Cone

Consider set

$$\{(\tau, x) : \tau > 0, \tau c_i(x/\tau) \leq 0\}$$

The dual cone is the set of all points $(\kappa; s)$ such that

$$\kappa\tau + s^T x \geq 0, \quad \forall(\tau; x) \quad s.t.\tau > 0, \tau c_i(x/\tau) \leq 0, i = 1, ..., m$$

Without loss of generality, we can set $\tau = 1$ and the condition becomes

$$\kappa + s^T x \geq 0, \quad \forall x \ s.t. c_i(x) \leq 0, i = 1, ..., m$$

Then, consider the optimization problem

$$\begin{aligned} &\psi(s) := inf \quad s^T x \\ &s.t. \quad c_i(x) \leq 0, i = 1, 2, ..., m \end{aligned}$$

Then, the dual cone coan be represented as

$$K^\star = \{(\kappa, s) : \kappa + \psi(x) \geq 0\}$$

2.4 Combinationrial Auction Pricing

Given the m different states that are mutually exclusive and exactly one of them will be true at the maturity. A contract on a state is a paper agreement so that on maturity it is worth a notional \$1 if it is on the winning state and worth \$0 if is not on the winning state. There are n orders betting on one or a combination of states, with a price limit and a quantity limit

Order:	#1	#2	#3	#4	#5
Argentina	1	0	1	1	0
Brazil	1	0	0	1	1
Italy	1	0	1	1	0
Germany	0	1	0	1	1
France	0	0	1	0	0
Bidding Prize: π	0.75	0.35	0.4	0.95	0.75
Quantity limit:q	10	5	10	10	5
Order fill:x	x_1	x_2	x_3	x_4	x_5

Lethal x_j be the number of contacts awarded to the jth order. Then jth better will pay the amount

$$\pi_j \times x_j$$

and the total collected amount is $\sum_{i=1}^{n} \pi_j \times x_j = \pi^T x$. If the ith state is the winning state, then the auction organizer need to pay.

$$\sum_{j=1}^{n} a_{ij} x_j$$

We can formulate the primal and dual problem

$$\begin{array}{llll} \max & \pi^T \mathbf{x} - z & \min & \mathbf{q}^T \mathbf{y} \\ \text{s.t.} & A\mathbf{x} - \mathbf{e} \cdot z \leq \mathbf{0} & \text{s.t.} & A^T \mathbf{p} + \mathbf{y} \geq \pi, \\ & \mathbf{x} \leq \mathbf{q} & & \mathbf{e}^T \mathbf{p} = 1, \\ & \mathbf{x} \geq 0 & & (\mathbf{p}, \mathbf{y}) \geq 0. \end{array} \tag{2.14}$$

2.4.1 Online Linear Programming

The main idea of linear program is we don't know the coefficient matrix. We need to make the decision and reveal the information sequentially.

$$\begin{aligned} \max \quad & \sum_{t=1}^{n} \pi_t x_t \\ s.t. \quad & \sum_{t=1}^{n} a_{it} x_t \leq b_i, \quad \forall i = 1, \ldots, m \\ & 0 \leq x_t \leq 1, \quad \forall t = 1, \ldots, n \end{aligned} \tag{2.15}$$

Each bid/activity t requests a bundle of m resources, and the payment is π_t. Online Decision Making: we only know $(n, \mathbf{b})$ at the start, but - the (bounded) order-data of each variable x_t is revealed sequentially. - an irrevocable decision must be made as soon as an order arrives without observing or knowing the future data.

The algorithm/mechanism quality is evaluated on the expected performance over all the permutations comparing to the offline optimal solution, i.e., an algorithm $\mathcal{A}$ is c-competitive if and only if

$$E_\sigma \left[\sum_{t=1}^{n} \pi_t x_t(\sigma, \mathcal{A}) \right] \geq c \cdot OPT(A, \pi), \forall (A, \pi). \tag{2.16}$$

Then we will introduce the algorithm how to solve the online linear program.

1. Set $x_t = 0$ for all $1 \leq t \leq \epsilon n$

2. Solve the ϵ portion of the problem

$$\begin{array}{lll} \text{maximize}_{\mathbf{x}} & \sum_{t=1}^{\epsilon n} \pi_t x_t & \\ \text{subject to} & \sum_{t=1}^{\epsilon n} a_{it} x_t \leq \epsilon b_i \quad i = 1, \ldots, m & \\ & 0 \leq x_t \leq 1 & t = 1, \ldots, \epsilon n \end{array}$$

   and get the optimal dual solution $\hat{\mathbf{p}}$ of the sample LP;

3. Determine the future allocation x_t as:

$$x_t = \begin{cases} 0 & \text{if } \pi_t \leq \hat{\mathbf{p}}^T \mathbf{a}_t \\ 1 & \text{if } \pi_t > \hat{\mathbf{p}}^T \mathbf{a}_t \end{cases}$$

   as long as $a_{it} x_t \leq b_i - \sum_{j=1}^{t-1} a_{ij} x_j$ for all i; otherwise, set $x_t = 0$. Online Learning: Periodically resolve the sample LP with all arrived orders and update the "ideal" prices...

Chapter 3

Optimality Conditions and KKT Condition

Ref Chapters 7.1-7.3, 11.1-11.8

3.1 Optimality Conditions

Recall the primal and dual program

$$\left\{ (\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \left(\mathcal{R}_+^n, \mathcal{R}^m, \mathcal{R}_+^n\right) : \quad \begin{aligned} \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} &= \mathbf{0} \\ A\mathbf{x} &= \mathbf{b} \\ -A^T\mathbf{y} - \mathbf{s} &= -\mathbf{c} \end{aligned} \right\} \tag{3.1}$$

Let $x^\star$ and $s^\star$ be optimal solutions with zero duality gap

$$|support(x^\star)| + |support(s^\star)| \leq n \tag{3.2}$$

Note that support size means the number of non-zero components. Then we will have

Theorem 3.1. *If (LP) and (LD) are both feasible, then there exists a pair of strictly complementary solutions $\mathbf{x}^* \in \mathcal{F}_p$ and $(\mathbf{y}^*, \mathbf{s}^*) \in \mathcal{F}_d$ such that*

$$\mathbf{x}^* \cdot \mathbf{s}^* = \mathbf{0} \text{ and } |\text{supp}\,(\mathbf{x}^*)| + |\text{supp}\,(\mathbf{s}^*)| = n. \tag{3.3}$$

Moreover, the supports

$$P^* = \left\{j : x_j^* > 0\right\} \quad \text{and} \quad Z^* = \left\{j : s_j^* > 0\right\}$$

are invariant for all strictly complementary solution pairs.

3.1.1 Uniqueness Theorem for Linear Program

If we have the optimal solution $x^\star$ and how to clarify the uniqueness of $x^\star$

Theorem 3.2. *An LP optimal solution* $\mathbf{x}^*$ *is unique if and only if the size of supp* $(\mathbf{x}^*)$ *is maximal among all optimal solutions and the columns of* $A_{Supp}(x^*)$ *are linear independent.*

Proof. It is easy to see both conditions are necessary, since otherwise, one can find an optimal solution with a different support size. To see sufficiency, suppose there there is another optimal solution $\mathbf{y}^*$ such that $\mathbf{x}^* - \mathbf{y}^* \neq \mathbf{0}$. We must have $\operatorname{supp}(\mathbf{y}^*) \subset \operatorname{supp}(\mathbf{x}^*)$, since, otherwise, $(0.5\mathbf{x}^* + 0.5\mathbf{y}^*)$ remains optimal and its support size is greater than that of x^* which is a contradiction. Then we see

$$\mathbf{0} = A\mathbf{x}^* - A\mathbf{y}^* = A(\mathbf{x}^* - \mathbf{y}^*) = A_{\operatorname{supp}(\mathbf{x}^*)}(\mathbf{x}^* - \mathbf{y}^*)\operatorname{supp}(\mathbf{x}^*)$$

which implies that columns of $A_{\operatorname{Supp}(\mathbf{x}^*)}$ are linearly dependent. Think the $x^\star = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $y^\star = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ ☐

3.1.2 Uniqueness Theorem for Semidefinite Program

Theorem 3.3. *An SDP optimal and complementary solution* X^* *is unique if and only if the rank of* X^* *is maximal among all optimal solutions and* $V^* A_i (V^*)^T, i = 1, \ldots, m$*, are linearly independent, where* $X^* = (V^*)^T V^*, V^* \in \mathcal{R}^{r \times n}$*, and* r *is the rank of* X^**.*

3.2 Relaxation Example

3.2.1 Sensor Localization Problem

Given $\mathbf{a}_k \in \mathbf{R}^d, d_{ij} \in N_x$, and $\hat{d}_{kj} \in N_a$, find $\mathbf{x}_i \in \mathbf{R}^d$ such that

$$\begin{aligned} \|\mathbf{x}_i - \mathbf{x}_j\|^2 &= d_{ij}^2, \forall (i,j) \in N_x, i < j, \\ \|\mathbf{a}_k - \mathbf{x}_j\|^2 &= \hat{d}_{kj}^2, \forall (k,j) \in N_a, \end{aligned}$$

We can transform to the matrix form and SDP relaxation

$$\begin{aligned}
&(\mathbf{e}_i - \mathbf{e}_j)^T Y (\mathbf{e}_i - \mathbf{e}_j) = d_{ij}^2, \forall i, j \in N_x, i < j \\
&(\mathbf{a}_k; -\mathbf{e}_j)^T \begin{pmatrix} I & X \\ X^T & Y \end{pmatrix} (\mathbf{a}_k; -\mathbf{e}_j) = \hat{d}_{kj}^2, \forall k, j \in N_a \\
&\begin{pmatrix} I & X \\ X^T & Y \end{pmatrix} \succeq \mathbf{0}
\end{aligned} \tag{3.4}$$

Theorem 3.4. *Let $\bar{Z}$ be a feasible solution for SDP and $\bar{U}$ be an optimal slack matrix of the dual. Then,*

1. *complementarity condition holds: $\bar{Z} \bullet \bar{U} = 0$ or $\bar{Z}\bar{U} = 0$*
2. $\mathrm{Rank}(\bar{Z}) + \mathrm{Rank}(\bar{U}) \leq 2 + n$
3. $\mathrm{Rank}(\bar{Z}) \geq 2$ *and* $\mathrm{Rank}(\bar{U}) \leq n$

3.3 Rank-Reduction for SDP

In the most SDP case, it is difficult to find the rank-minimal SDP solution

$$\begin{array}{lll} (SDP) & \min & C \bullet X \\ & \text{subject to} & A_i \bullet X = b_i, i = 1, 2, \ldots, m, X \succeq 0 \end{array} \tag{3.5}$$

where $C, A_i \in S^n$

Theorem 3.5. *(Carathéodory's theorem)*

- *If there is a minimizer for (LP), then there is a minimizer of (LP) whose support size r satisfying $r \leq m$*
- *If there is a minimizer for (SDP), then there is a minimizer of (SDP) whose rank r satisfying $\frac{r(r+1)}{2} \leq m$. Moreover, such a solution can be find in polynomial time.*

Then if we simply the SDP feasiblility problem

$$A_i \bullet X = b_i \quad i = 1, .., m, \quad X \succeq 0$$

we try to find an approximate $\hat{X} \succeq 0$ of rank at most d

$$\beta(m, n, d) \cdot b_i \leq A_i \bullet \hat{X} \leq \alpha(m, n, d) \cdot b_i \quad \forall i = 1, \ldots, m \tag{3.6}$$

Here, $\alpha \geq 1$ and $\beta \in (0, 1]$ are called the distortion factors. Clearly, the closer are both to 1 , the better.

Theorem 3.6. *Let* $r = \max\{\text{rank}(A_i)\}$ *and* $\bar{X}$ *be a feasible solution. Then, for any* $d \geq 1$, *the randomly generated*

$$\hat{X} = \sum_i^d \xi_i \xi_i^T, \quad \xi_i \in N\left(\mathbf{0}, \frac{1}{d}\bar{X}\right)$$

$$\alpha(m,n,d) = \begin{cases} 1 + \frac{12\ln(4mr)}{d} & \textit{for } 1 \leq d \leq 12\ln(4mr) \\ 1 + \sqrt{\frac{12\ln(4mr)}{d}} & \textit{for } d > 12\ln(4mr) \end{cases}$$

and

$$\beta(m,n,d) = \begin{cases} \frac{1}{e(2m)^{2/d}} & \textit{for } 1 \leq d \leq 4\ln(2m) \\ \max\left\{\frac{1}{e(2m)^{2/d}}, 1 - \sqrt{\frac{4\ln(2m)}{d}}\right\} & \textit{for } d > 4\ln(2m) \end{cases}$$

Here is the some remarks from theorem 9

- There is always a low-rank, or sparse, approximate SDP solution with respect to a bounded relative residual distortion. As the allowable rank increases, the distortion bounds become smaller and smaller.
- The lower distortion factor is independent of n and the rank of A_i s.
- The factors can be improved if we only consider one-sided inequalities.
- This result contains as special cases several well-known results in the literature.
- Can the distortion upper bound be improved such that it's independent of rank of A_i ?
- Is there deterministic rank-reduction procedure? Choose the largest d eigenvalue component of X ?
- General symmetric A_i ?
- In practical applications, we see much smaller distortion, why?

3.4 Max-Cut Problem

This is the Max-Cut problem on an undirected graph $G = (V, E)$ with non-negative weights w_{ij} for each edge in E (and $w_{ij} = 0$ if $(i,j) \notin E$), which is the problem of partitioning the nodes of V into two sets S and $V \backslash S$ so that

$$w(S) := \sum_{i \in S, j \in V \backslash S} w_{ij}$$

is maximized. A problem of this type arises from many network planning, circuit design, and scheduling applications.

$$\begin{aligned} w^* := \quad \text{Maximize} \quad & w(\mathbf{x}) := \frac{1}{4}\sum_{i,j} w_{ij}(1 - x_i x_j) \\ s.t. \quad & (x_j)^2 = 1, j = 1, \ldots, n \end{aligned} \tag{3.7}$$

Then we do the semidefinite relaxation reformulation. Let $Z_{ij} = x_i x_j$

$$\begin{aligned} z^{SDP} := \ & \text{Maximize } \frac{1}{4}\sum_{i,j} w_{ij}\,(1 - Z_{ij}) \\ & \textit{Subject to} \quad Z_{ii} = 1, \quad i = 1, \ldots, n \\ & \qquad\qquad\quad Z \succeq \mathbf{0}, \operatorname{rank}(Z) = 1 \end{aligned} \tag{3.8}$$

If we remove the rank-one constraint, it will be SDP relaxation problem

Theorem 3.7. *(Goemans and Williamson)*

$$\mathrm{E}[w(\hat{\mathbf{x}})] \geq .878 z^{SDP} \geq .878 w^* \tag{3.9}$$

Ref Chapters 11.7-8, 14.1-2

3.5 Optimality Conditions for Nonlinear Optimization

3.5.1 KKT Optimality Condition

A differentiable function f of one variable defined on an interval $F = [ae]$. If an interior-point $\bar{x}$ is a local/global minimizer, then $f'(\bar{x}) = 0$; if the left-end-point $\bar{x} = a$ is a local minimizer, then $f'(a) \geq 0$; if the right-end-point $\bar{x} = e$ is a local minimizer, then $f'(e) \leq 0$. first-order necessary condition (FONC) summarizes the three cases by a unified set of optimality/complementarity slackness conditions

If $f'(\bar{x}) = 0$, then it is also necessary that $f(x)$ is locally convex at $\bar{x}$ for it being a local minimizer. How to tell the function is locally convex at the solution? It is necessary $f''(\bar{x}) \geq 0$, which is called the second-order necessary condition (SONC), which we would explored further.

These conditions are still not, in general, sufficient. It does not distinguish between local minimizers, local maximizers, or saddle points.

If the second-order sufficient condition (SOSC): $f''(\bar{x}) > 0$, is satisfied or the function is strictly locally convex, then $\bar{x}$ is a local minimizer. Thus, if the function is convex everywhere, the first-order necessary condition is already sufficient.

Then we want to explore the second order optimality under unconstrained optimization.

Theorem 3.8. *(First-Order Necessary Condition) Let $f(\mathbf{x})$ be a C^1 function where $\mathbf{x} \in R^n$. Then, if $\overline{\mathbf{x}}$ is a minimizer, it is necessarily $\nabla f(\overline{\mathbf{x}}) = 0$.*

Theorem 3.9. *(Second-Order Necessary Condition) Let $f(\mathbf{x})$ be a C^2 function where $\mathbf{x} \in R^n$. Then, if $\overline{\mathbf{x}}$ is a minimizer, it is necessarily*

$$\nabla f(\overline{\mathbf{x}}) = \mathbf{0} \textit{ and } \nabla^2 f(\overline{\mathbf{x}}) \succeq \mathbf{0}.$$

Note that the Hessian matrix is the semidefinite

Furthermore, if $\nabla^2 f(\overline{\mathbf{x}}) \succ 0$, then the condition becomes sufficient. The proofs would be based on 2nd-order Taylor's expansion at $\overline{\mathrm{x}}$

Proof. we have the second order Taylor's expansion

$$f(x) = f(\bar{x}) + (x - \bar{x})^T \bigtriangledown f(\bar{x}) + (x - \bar{x})^T \bigtriangledown^2 f(\bar{x})(x - \bar{x}) + o(||x - \bar{x}||^2)$$

If $\bigtriangledown f(\bar{x}) = 0$ and $\bigtriangledown^2 f(\bar{x}) = 0$, the $\bar{x}$ is not the local optimal. □

Such that if these conditions are not satisfied, then one would be find a descent-direction d and a small constant $\bar{\alpha} > 0$ such that $f(\overline{\mathbf{x}} + \alpha \mathbf{d}) < f(\overline{\mathbf{x}}), \forall 0 < \alpha \leq \bar{\alpha}$

For example, if $\nabla f(\overline{\mathbf{x}}) = \mathbf{0}$ and $\nabla^2 f(\overline{\mathbf{x}}) \not\succeq 0$, the eigenvector of a negative eigenvalue of the Hessian would be a descent direction from $\overline{\mathrm{x}}$. Again, they may still not be sufficient, e.g., $f(x) = x^3$.

3.5.2 Descent Direction

Let f be a differentiable function on R^n. If point $\overline{\mathrm{x}} \in R^n$ and there exists a vector d such that

$$\nabla f(\overline{\mathbf{x}})\mathbf{d} < 0$$

then there exists a scalar $\bar{\tau} > 0$ such that

$$f(\overline{\mathbf{x}} + \tau \mathbf{d}) < f(\overline{\mathbf{x}}) \text{ for all } \tau \in (0, \bar{\tau}).$$

Note that $\exists \tau, \frac{f(\overline{\mathbf{x}}+\tau\mathbf{d})-f(\overline{\mathbf{x}})}{\tau\mathbf{d}}\mathbf{d} < 0$

The vector $\mathbf{d}$ (above) is called a descent direction at $\overline{\mathbf{x}}$. If $\nabla f(\overline{\mathbf{x}}) \neq 0$, then $\nabla f(\overline{\mathbf{x}})$ is the direction of steepest ascent and $-\nabla f(\overline{\mathbf{x}})$ is the direction of steepest descent at $\overline{\mathbf{x}}$. Denote by $\mathcal{D}^d_{\overline{\mathrm{x}}}$ the set of descent directions at $\overline{\mathrm{x}}$, that is,

$$\mathcal{D}^d_{\overline{\mathrm{x}}} = \{\mathbf{d} \in R^n : \nabla f(\overline{\mathbf{x}})\mathbf{d} < 0\}$$

At feasible point $\overline{\mathbf{X}}$, a feasible direction is

$$\mathcal{D}^f_{\overline{\mathbf{x}}} := \{\mathbf{d} \in R^n : \mathbf{d} \neq \mathbf{0}, \overline{\mathbf{x}} + \lambda \mathbf{d} \in \mathcal{F} \text{ for all small } \lambda > 0\}.$$

3.5.3 Optimality Condition of Problem

Roughly the optimality condition of unconstrained problem is

Theorem 3.10. *Let $\overline{\mathrm{x}}$ be a (local) minimizer of (UP). If the functions f is continuously differentiable at $\overline{\mathrm{x}}$, then*

$$\nabla f(\overline{\mathbf{x}}) = \mathbf{0}.$$

This condition is also sufficient for global optimality if f is a convex function.

Consider the linear equality-constrained problem, where f is differentiable on R^n,

$$\begin{array}{lrl} \text{(LEP)} & & \min f(\mathbf{x}) \\ & \text{s.t.} & A\mathbf{x} = \mathbf{b}. \end{array}$$

Theorem 3.11. *(the Lagrange Theorem) Let $\bar{\mathrm{x}}$ be a (local) minimizer of (LEP). If the functions f is continuously differentiable at $\bar{\mathrm{x}}$, then*

$$\nabla f(\bar{\mathbf{x}}) = \bar{\mathbf{y}}^T A$$

for some $\bar{\mathbf{y}} = (\bar{y}_1; \ldots; \bar{y}_m) \in R^m$, which are called Lagrange or dual multipliers. This condition is also sufficient for global optimality if f is a convex function.

The Lagrange formula is

$$L(x, y) = f(x) + y^T b - y^T Ax$$

Then take the derivative

$$\triangledown L(\bar{x}, y) = \triangledown f(\bar{x}) - \bar{y}^T A = 0$$

The geometric interpretation: the objective gradient vector is perpendicular to or the objective level set tangents the constraint hyperplanes. Let $v(\mathbf{b})$ be the minimal value function of $\mathbf{b}$ of (LEP). Simliarly

$$\triangledown L(b) = \bar{y} = 0$$

Proof. Consider feasible direction space

$$\mathcal{F} = \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\} \Rightarrow \mathcal{D}^f_{\bar{\mathrm{x}}} = \{\mathrm{d} : A\ \mathrm{d} = 0\}$$

If $\bar{\mathbf{x}}$ is a local optimizer, then the intersection of the descent and feasible direction sets at $\bar{x}$ must be empty or

$$A\mathbf{d} = \mathbf{0}, \nabla f(\bar{\mathbf{x}})\mathbf{d} \neq 0$$

has no feasible solution for d. By the Alternative System Theorem it must be true that its alternative system has a solution, that is, there is $\bar{y} \in R^n$ such that

$$\nabla f(\bar{\mathbf{x}}) = \bar{\mathbf{y}}^T A = \sum_{i=1}^{m} \bar{y}_i A_i$$

□

3.5.4 Barrier Optimization

Consider the problem

$$\begin{array}{ll} \min & -\sum_{j=1}^n \log x_j \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{array} \tag{3.10}$$

The non-negativity constraint can be removed if the feasible region has an "interior", that is, there is a feasible solution such that $\mathrm{x} > 0$. Thus, if a minimizer $\bar{\mathrm{x}}$ exists, then $\bar{\mathrm{x}} > 0$ and

$$-\mathbf{e}^T \bar{X}^{-1} = \bar{\mathbf{y}}^T A = \sum_{i=1}^{m} \bar{y}_i A_i.$$

3.6 KKT Condition

Let us now consider the inequality-constrained problem

$$\begin{array}{ll} & \min \quad f(\mathbf{x}) \\ \text{s.t.} & A\mathbf{x} \geq \mathbf{b}. \end{array}$$

Theorem 3.12. *(the KKT Theorem) Let $\bar{\mathrm{x}}$ be a (local) minimizer of (LIP). If the functions f is continuously differentiable at $\bar{\mathrm{x}}$, then*

$$\nabla f(\bar{\mathbf{x}}) = \bar{\mathbf{y}}^T A, \bar{\mathbf{y}} \geq \mathbf{0}$$

for some $\bar{\mathbf{y}} = (\bar{y}_1; \ldots; \bar{y}_m) \in R^m$, which are called Lagrange or dual multipliers, and $\bar{y}_i = 0$, if $i \notin \mathcal{A}(\bar{\mathbf{x}})$. These conditions are also sufficient for global optimality if f is a convex function.

Then we can have the KKT constraint. We now consider optimality conditions for problems having both inequality and equality constraints. These can be denoted

$$\begin{array}{lll} \min & f(\mathbf{x}) & \\ \text{s.t.} & \mathbf{a}_i\mathbf{x} & (\leq, =, \geq) \quad b_i, i = 1, \ldots, m, \end{array}$$

For any feasible point $\bar{x}$ of (P) we have the sets

$$\begin{aligned} \mathcal{A}(\bar{\mathbf{x}}) &= \{i : \mathbf{a}_i\mathbf{x} = b_i\} \\ \mathcal{D}^d_{\bar{\mathbf{x}}} &= \{\mathbf{d} : \nabla f(\bar{\mathbf{x}})\mathbf{d} < 0\} \end{aligned}$$

Theorem 3.13. *Let $\bar{\mathrm{x}}$ be a local minimizer for (P). Then there exist multipliers $\bar{\mathrm{y}}$ such that*

$$\begin{array}{rccl} \mathbf{a}_i\bar{\mathbf{x}} & (\leq, =, \geq) & b_i, i = 1, \ldots, m, & \\ \textit{(Original Problem Constraints } (OPC)) & & & \\ \nabla f(\bar{\mathbf{x}}) & = & \bar{\mathbf{y}}^T A & \\ \textit{(Lagrangian Multiplier Conditions (LMC))} & & & \\ \bar{y}_i & (\leq,' \textit{ free }', \geq) & 0, i = 1, \ldots, m, & \\ \textit{Multiplier Sign Constraints } (MSC) & & & \\ \bar{y}_i & = & 0 & \textit{if } i \notin \mathcal{A}(\bar{\mathbf{x}}) \\ \textit{(Complementarity Slackness Conditions (CSC)).} & & & \end{array} \tag{3.11}$$

These conditions are also sufficient for global optimality if f is a convex function.

And we will emphasize again to sufficient and necessary the convexity. Like $f(x) = x^3$

- Hessian matrix is PSD in the feasible region
- Epigraph is a convex set

3.6.1 LCOP Examples: Linear Optimization

$$
\begin{aligned}
(LP) \quad \min \quad & \mathbf{c}^T\mathbf{x} \\
& s.t. A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0 \\
& \text{s}
\end{aligned}
$$

For any feasible x of (LP), it's optimal if for some y,

$$
\begin{aligned}
x_j s_j &= 0, \forall j = 1, \ldots, n \\
A\mathbf{x} &= \mathbf{b} \\
\nabla \left(\mathbf{c}^T\mathbf{x}\right) = \mathbf{c}^T &= \mathbf{y}^T A + \mathbf{s}^T \\
\mathbf{x}, \mathbf{s} &\geq \mathbf{0}.
\end{aligned}
$$

Here, $\mathbf{y}$ (shadow prices in LP) are Lagrange or dual multipliers of equality constraints, and $\mathbf{s}$ (reduced gradient/costs in LP) are Lagrange or dual multipliers for x $\geq \mathbf{0}$.

3.6.2 LCOP Examples : Quadratic Optimization

$$
\begin{aligned}
(QP) \quad \min \quad & \mathbf{x}^T Q\mathbf{x} - 2\mathbf{c}^T\mathbf{x} \\
\text{s.t.} \quad & A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}
\end{aligned}
$$

Optimality Conditions:

$$
\begin{aligned}
x_j s_j &= 0, \forall j = 1, \ldots, n \\
A\mathbf{x} &= \mathbf{b} \\
2Q\mathbf{x} - 2\mathbf{c} - A^T\mathbf{y} - \mathbf{s} &= \mathbf{0} \\
\mathbf{x}, \mathbf{s} &\geq \mathbf{0}
\end{aligned}
$$

3.6.3 LCOP Examples: Linear Barrier Optimization

$$
\min f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} - \mu \sum_{j=1}^{n} \log\left(x_j\right), \text{ s.t. } A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}
$$

for some fixed $\mu > 0$. Assume that interior of the feasible region is not empty:

$$
\begin{aligned}
A\mathbf{x} &= \mathbf{b} \\
c_j - \frac{\mu}{x_j} - \left(\mathbf{y}^T A\right)_j &= 0, \forall j = 1, \ldots, n \\
\mathbf{x} &> \mathbf{0}.
\end{aligned}
$$

Let $s_j = \frac{\mu}{x_j}$ for all j (note that this s is not the s in the KKT condition of $f(\mathbf{x})$). Then

$$\begin{aligned} x_j s_j &= \mu, \forall j = 1, \ldots, n, \\ A\mathbf{x} &= \mathbf{b}, \\ A^T\mathbf{y} + \mathbf{s} &= \mathbf{c}, \\ (\mathbf{x}, \mathbf{s}) &> \mathbf{0}. \end{aligned}$$

3.7 Inverse Optimization

We know that the KKT theorem could be applied into the equilibrium, and the dual variable reflects the econ perproties. We can explore which are we can use the KKT theorem, such as the inverse optimization. This area is introduced by Ahuja and Orlin (2001). In this paper, we study inverse optimization problems defined as follows. Let **S** denote the set of feasible solutions of an optimization problem P, let c be a specified cost vector, and x^0 be a given feasible solution. The solution x^0 may or may not be an optimal solution of **P** with respect to the cost vector c. The inverse optimization problem is to perturb the cost vector c to d so that x^0 is an optimal solution of **P** with respect to d and $\|d-c\|_p$ is minimum, where $\|d-c\|_p$ is some selected L_p norm. In this paper, we consider the inverse linear programming problem under L_1 norm (where $\|d-c\|_p = \sum_{i\in J} w_j |d_j - c_j|$, with J denoting the index set of variables x_j and w_j denoting the weight of the variable j) and under L_∞ norm (where $\|d-c\|_p = \max_{j\in J} \{w_j |d_j - c_j|\}$). We prove the following results:

1. If the problem **P** is a linear programming problem, then its inverse problem under the L_1 as well as L_∞ norm is also a linear programming problem.

2. If the problem **P** is a shortest path, assignment or minimum cut problem, then its inverse problem under the L_1 norm and unit weights can be solved by solving a problem of the same kind. For the nonunit weight case, the inverse problem reduces to solving a minimum cost flow problem.

3. If the problem **P** is a minimum cost flow problem, then its inverse problem under the L_1 norm and unit weights reduces to solving a unit-capacity minimum cost flow problem. For the nonunit weight case, the inverse problem reduces to solving a minimum cost flow problem.

4. If the problem **P** is a minimum cost flow problem, then its inverse problem under the L_∞ norm and unit weights reduces to solving a minimum mean cycle problem. For the nonunit weight case, the inverse problem reduces to solving a minimum cost-to-time ratio cycle problem.

5. If the problem **P** is polynomially solvable for linear cost functions, then inverse versions of **P** under the L_1 and L_∞ norms are also polynomially solvable.

Ref Chapters 7.6-7, and 8.1-5

3.8 General Constrained Optimization

We show the example of general constrained optimization

$$\begin{array}{lll} (GCO) & \min & f(\mathbf{x}) \\ \text{s.t.} & & \mathbf{h}(\mathbf{x}) = \mathbf{0} \in R^m \\ & & \mathbf{c}(\mathbf{x}) \geq \mathbf{0} \in R^p \end{array} \tag{3.12}$$

We establish to optimality condition to verify the local minimizer or an KKT solution. Consider the intersection of hypersurfaces

$$\{x \in R^n : h(x) = 0 \in R^m, m \leq n\}$$

when function $h_i(x)$ are differentiable, we say the surface is smooth.

Definition 3.14. *A point $\mathbf{x}^*$ satisfying the constraint $\mathbf{h}(\mathbf{x}^*) = \mathbf{0}$ is said to be a regular point of the constraint if the gradient vectors $\nabla h_1(\mathbf{x}^*), \nabla h_2(\mathbf{x}^*), \ldots, \nabla h_m(\mathbf{x}^*)$ are linearly independent.*

Based on the Implicit Function Theorem (Appendix A of the Text), if $\overline{\mathbf{x}}$ is a regular point and $m < n$, then for every $\mathrm{d} \in \mathcal{T}_{\overline{\mathrm{x}}} = \{\mathbf{z} : \nabla \mathbf{h}(\overline{\mathbf{x}})\mathbf{z} = 0\}$ there exists a curve $\mathbf{x}(t)$ on the hypersurface, parametrized by a scalar t in a sufficiently small interval $\begin{bmatrix} -a & a \end{bmatrix}$, such that

$$\mathbf{h}(\mathbf{x}(t)) = \mathbf{0}, \quad \mathbf{x}(0) = \overline{\mathbf{x}}, \quad \dot{\mathbf{x}}(0) = \mathbf{d}.$$

$\mathcal{T}_{\overline{\mathrm{x}}}$ is called the tangent-space or tangent-plane of the constraints at $\overline{\mathrm{x}}$.

Lemma 3.15. *Let $\overline{\mathrm{x}}$ be a feasible solution and a regular point of the hypersurface of*

$$\{\mathbf{x} : \mathbf{h}(\mathrm{x}) = \mathbf{0}, c_i(\mathrm{x}) = 0, i \in \mathcal{A}_{\overline{\mathrm{x}}}\}$$

where active-constraint set $\mathcal{A}_{\overline{\mathrm{x}}} = \{i : c_i(\overline{\mathrm{x}}) = 0\}$. If $\overline{\mathrm{x}}$ is a (local) minimizer of (GCO), then there must be no d to satisfy linear constraints:

$$\begin{aligned} &\nabla f(\overline{\mathbf{x}})\mathbf{d} < 0 \\ &\nabla \mathbf{h}(\overline{\mathbf{x}})\mathbf{d} = \mathbf{0} \in R^m \\ &\nabla c_i(\overline{\mathbf{x}})\mathbf{d} \geq 0, \forall i \in \mathcal{A}_{\overline{\mathbf{x}}}. \end{aligned} \tag{3.13}$$

3.8.1 First-Order Necessary Conditions

(First-Order or KKT Optimality Condition) Let $\overline{\mathrm{x}}$ be a (local) minimizer of (GCO) and it is a regular point of $\{\mathbf{x} : \mathbf{h}(\mathbf{x}) = \mathbf{0}, c_i(\mathbf{x}) = 0, i \in \mathcal{A}_{\overline{\mathbf{x}}}\}$. Then, for some multipliers $(\overline{\mathbf{y}}, \overline{\mathbf{s}} \geq \mathbf{0})$

$$\nabla f(\overline{\mathbf{x}}) = \overline{\mathbf{y}}^T \nabla \mathbf{h}(\overline{\mathbf{x}}) + \overline{\mathbf{s}}^T \nabla \mathbf{c}(\overline{\mathbf{x}})$$

and (complementarity slackness)

$$\bar{s}_i c_i(\overline{\mathbf{x}}) = 0, \forall i$$

The proof is again based on the Alternative System Theory or Farkas Lemma. The complementarity slackness condition is from that $c_i(\bar{\mathbf{x}}) = 0$ for all $i \in \mathcal{A}_{\bar{\mathbf{x}}}$, and for $i \notin \mathcal{A}_{\bar{\mathbf{x}}}$, we simply set $\bar{s}_i = 0$. A solution who satisfies these conditions is called an KKT point or solution of (GCO) - any local minimizer $\bar{\mathrm{x}}$, if it is also a regular point, must be an KKT solution; but the reverse may not be true.

3.8.2 KKT via the Lagrangian Function

It is more convenient to introduce the Lagrangian Function associated with generally constrained optimization:

$$L(\mathbf{x}, \mathbf{y}, \mathbf{s}) = f(\mathbf{x}) - \mathbf{y}^T \mathbf{h}(\mathbf{x}) - \mathbf{s}^T \mathbf{c}(\mathbf{x}) \tag{3.14}$$

where multipliers y of the equality constraints are "free" and s ≥ 0 for the "greater or equal to" inequality constraints, so that the KKT condition (2) can be written as

$$\nabla_{\mathbf{x}} L(\bar{\mathbf{x}}, \bar{\mathbf{y}}, \bar{\mathbf{s}}) = 0 \tag{3.15}$$

We now consider optimality conditions for problems having three types of inequalities:

$$\begin{array}{llll} \text{(GCO)} & \min & f(\mathbf{x}) & \\ & \text{s.t.} & c_i(\mathbf{x}) \quad (\leq, =, \geq) \quad 0, i = 1, \dots, m, & \text{(Original Problem Constraints (OPC))} \end{array}$$

For any feasible point $\mathbf{x}$ of (GCO) define the active constraint set by $\mathcal{A}_{\mathbf{x}} = \{i : c_i(\mathbf{x}) = 0\}$. Let $\bar{\mathrm{x}}$ be a local minimizer for (GCO) and $\bar{\mathrm{x}}$ is a regular point on the hypersurface of the active constraints Then there exist multipliers $\bar{\mathbf{y}}$ such that The complete First-Order KKT Conditions consist of these four parts!

Example 3.16. *Given $\mathbf{a}_k \in \mathbf{R}^2$ and Euclidean distances $d_k, k = 1, 2, 3$, find $\mathbf{x} \in \mathbf{R}^2$ such that*

$$\begin{array}{ll} \min_{\mathbf{x}} & \mathbf{0}^T \mathbf{x}, \\ s.t. & \|\mathbf{x} - \mathbf{a}_k\|^2 - d_k^2 \leq 0, k = 1, 2, 3, \end{array}$$

$$L(\mathbf{x}, \mathbf{y}) = \mathbf{0}^T \mathbf{x} - \sum_{k=1}^3 y_k \left(\|\mathbf{x} - \mathbf{a}_k\|^2 - d_k^2 \right),$$

$$\begin{array}{rll} \mathbf{0} & = \sum_{k=1}^3 y_k (\mathbf{x} - \mathbf{a}_k) & \textit{(LDC)} \\ y_k & \leq 0, k = 1, 2, 3, & \textit{(MSC)} \\ y_k \left(\|\mathbf{x} - \mathbf{a}_k\|^2 - d_k^2 \right) & = 0. & \textit{(CSC).} \end{array}$$

Then we talk about the Arrow-Debreu's exchange market.

Example 3.17. *(Arrow-Debreu's exchange market)*

$$\begin{array}{ll} \max & \mathbf{u}_i^T \mathbf{x}_i \\ s.t. & \mathbf{p}^T \mathbf{x}_i \leq \mathbf{p}^T \mathbf{w}_i, \\ & x_{ij} \geq 0, \quad \forall j, \end{array} \tag{3.16}$$

take the dual

$$\begin{array}{ll} \min & \lambda_i \mathbf{p}^T \mathbf{w}_i \\ s.t. & \lambda_i \mathbf{p} \geq \mathbf{u}_i \\ & \lambda_i \geq 0 \end{array} \tag{3.17}$$

Since we hold the weak duality theorem $\lambda_i \mathbf{p}^T\mathbf{w}_i \geq \mathbf{u}_i^T\mathbf{x}_i$. The setting indicate that the term are non-negative.

$$\begin{aligned} \lambda_i \mathbf{p}^T\mathbf{w}_i &\geq \mathbf{u}_i^T\mathbf{x}_i \\ \frac{1}{\lambda_i} &\leq \frac{\mathbf{p}^T\mathbf{w}_i}{\mathbf{u}_i^T\mathbf{x}_i} \\ p_j &\geq u_{ij}\frac{\mathbf{p}^T\mathbf{w}_i}{\mathbf{u}_i^T\mathbf{x}_i} \end{aligned}$$

We want to find the equilibrium. The necessary and sufficient equilibrium conditions of the Arrow-Debreu market are

$$\begin{array}{ll} p_j \geq u_{ij}\cdot\frac{\mathbf{p}^T\mathbf{w}_i}{\mathbf{u}_i^T\mathbf{x}_i}\cdot & \forall i,j \\ \sum_i x_{ij} = \sum_i w_{ij} & \forall j, \\ p_j > 0, \mathbf{x}_i \geq \mathbf{0}, & \forall i,j \end{array} \tag{3.18}$$

The first inequality is nonlinear

$$\begin{aligned} p_j * \mathbf{p}^T\mathbf{w}_i &\geq u_{ij}\mathbf{u}_i^T\mathbf{x}_i \geq 0 \\ \log(p_j * \mathbf{p}^T\mathbf{w}_i) &\geq \log(u_{ij}\mathbf{u}_i^T\mathbf{x}_i) \\ \log(p_j) + \log(\mathbf{p}^T\mathbf{w}_i) - \log(\mathbf{u}_i^T\mathbf{x}_i) &\geq \log(u_{ij}) \end{aligned}$$

Let $y_j = \log(p_j)$ or $p_j = \exp^{y_j}$. Note that left term is concave in x_i and y_j

$$y_j + \log(\sum_j \exp^{y_j} w_{ij}) - \log(\mathbf{u}_i^T\mathbf{x}_i) \geq \log(u_{ij}) \tag{3.19}$$

They are different utility function in the exchange markets, such as

Cobb-Douglas utility:

$$u_i(x_i) = \prod_j x_{ij}^{u_{ij}}$$

Leontief utility:

$$u_i(x_i) = \min_j \left\{\frac{x_{ij}}{u_{ij}}, x_{ij} \geq 0\right\}$$

If we use the above utility and do the linear algebra steps, we can also show KKT solution suffices.

3.9 Second Order Sufficient Condition

Next we discuss the second order sufficient condition. We assume the function are twice continuously differentiable. Recall the tangent linear subspace at $\bar{x}$

$$T_{\bar{x}} = \{\mathbf{z} : \mathbf{h}(\overline{\mathbf{x}})\mathbf{z} = \mathbf{0}, \nabla c_i(\overline{\mathbf{x}})\mathbf{z} = 0, \forall i \in \mathcal{A}_{\overline{\mathbf{x}}}\}$$

Theorem 3.18. *Let* $\overline{\mathrm{x}}$ *be a (local) minimizer of* (GCO) *and a regular point of hypersurface* $\{\mathbf{x} : \mathbf{h}(\mathbf{x}) = 0, c_i(\mathbf{x}) = 0, i \in \mathcal{A}_{\overline{\mathbf{x}}}\}$, *and let* $\overline{\mathbf{y}}, \overline{\mathbf{s}}$ *denote Lagrange multipliers such that* $(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \overline{\mathbf{s}})$ *satisfies the (first-order) KKT conditions of (GCO). Then, it is necessary to have*

$$\mathbf{d}^T\nabla_{\mathrm{x}}^2 L(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \overline{\mathbf{s}})\mathbf{d} \geq 0 \quad \forall \mathbf{d} \in T_{\overline{\mathbf{x}}}. \tag{3.20}$$

The Hessian of the Lagrangian function need to be positive semidefinite on the tangent-space. The details of this theorem could be checked out in the textbook Luenberger et al. (1984)

Proof. The proof reduces to one-dimensional case by considering the objective function $\phi(t) = f(x(t))$ for the feasible curve $\mathbf{x}(t)$ on the surface of ALL active constraints. Since 0 is a (local) minimizer of $\phi(t)$ in an interval $[-aa]$ for a sufficiently small $a > 0$, we must have $\phi'(0) = 0$ so that

$$0 \leq \phi''(t)|_{t=0} = \dot{\mathbf{x}}(0)^T \nabla^2 f(\overline{\mathbf{x}}) \dot{\mathbf{x}}(0) + \nabla f(\overline{\mathbf{x}}) \ddot{\mathbf{x}}(0) = \mathbf{d}^T \nabla^2 f(\overline{\mathbf{x}}) \mathbf{d} + \nabla f(\overline{\mathbf{x}}) \ddot{\mathbf{x}}(0).$$

Let all active constraints (including the equality ones) be $\mathrm{h}(\mathrm{x}) = 0$ and differentiating equations $\bar{y}^T \mathbf{h}(\mathbf{x}(t)) = \sum_i \bar{y}_i h_i(\mathbf{x}(t)) = 0$ twice, we obtain

$$0 = \dot{\mathbf{x}}(0)^T \left[\sum_i \bar{y}_i \nabla^2 h_i(\overline{\mathbf{x}}) \right] \dot{\mathbf{x}}(0) + \bar{y}^T \nabla \mathbf{h}(\overline{\mathbf{x}}) \ddot{\mathbf{x}}(0) = \mathbf{d}^T \left[\sum_i \bar{y}_i \nabla^2 h_i(\overline{\mathbf{x}}) \right] \mathbf{d} + \bar{y}^T \nabla \mathbf{h}(\overline{\mathbf{x}}) \ddot{\mathbf{x}}(0).$$

Let the second expression subtracted from the first one on both sides and use the FONC:

$$\begin{aligned} 0 &\leq \mathbf{d}^T \nabla^2 f(\overline{\mathbf{x}}) \mathbf{d} - \mathbf{d}^T \left[\sum_i \bar{y}_i \nabla^2 h_i(\overline{\mathbf{x}}) \right] \mathbf{d} + \nabla f(\overline{\mathbf{x}}) \ddot{\mathbf{x}}(0) - \bar{y}^T \nabla \mathbf{h}(\overline{\mathbf{x}}) \ddot{\mathbf{x}}(0) \\ &= \mathbf{d}^T \nabla^2 f(\overline{\mathbf{x}}) \mathbf{d} - \mathbf{d}^T \left[\sum_i \bar{y}_i \nabla^2 h_i(\overline{\mathbf{x}}) \right] \mathbf{d} \\ &= \mathbf{d}^T \nabla^2_{\mathbf{x}} L(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \overline{\mathbf{s}}) \mathbf{d} \end{aligned}$$

Note that this inequality holds for every $\mathrm{d} \in T_{\overline{\mathrm{x}}}$. □

Furthermore we could show similar theorem

Theorem 3.19. *Let $\overline{\mathrm{x}}$ be a regular point of (GCO) with equality constraints only and let $\overline{\mathrm{y}}$ be the Lagrange multipliers such that $(\overline{\mathbf{x}}, \overline{\mathbf{y}})$ satisfies the (first-order) KKT conditions of* (GCO)*. Then, if in addition*

$$\mathbf{d}^T \nabla^2_{\mathrm{x}} L(\overline{\mathbf{x}}, \overline{\mathbf{y}}) \mathbf{d} > 0 \quad \forall \mathbf{0} \neq \mathbf{d} \in T_{\overline{\mathbf{x}}}, \tag{3.21}$$

then $\overline{\mathrm{x}}$ is a local minimizer of (GCO).

The simple example illustrates the regular points and KKT points.

Example 3.20.

$$\min (x_1)^2 + (x_2)^2 \ \ s.t. \ \ (x_1)^2/4 + (x_2)^2 - 1 = 0$$

$$L(x_1, x_2, y) = (x_1)^2 + (x_2)^2 - y\left(-(x_1)^2/4 - (x_2)^2 + 1\right)$$

$$\nabla_x L(x_1, x_2, y) = (2x_1(1 + y/4), 2x_2(1 + y))$$

$$\nabla^2_x L(x_1, x_2, y) = \begin{pmatrix} 2(1 + y/4) & 0 \\ 0 & 2(1 + y) \end{pmatrix}$$

$$T_{\mathrm{x}} := \{(z_1, z_2) : (x_1/4) z_1 + x_2 z_2 = 0\}$$

We see that there are two possible values for y : either -4 or -1, which lead to total four KKT points:

$$\begin{pmatrix} x_1 \\ x_2 \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -4 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ -4 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \text{ and } \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}$$

Consider the first KKT point:

$$\nabla_x^2 L(2,0,-4) = \begin{pmatrix} 0 & 0 \\ 0 & -6 \end{pmatrix}, T_{\overline{\mathbf{x}}} = \{(z_1, z_2) : z_1 = 0\}$$

Then the Hessian is not positive semidefinite on $T_{\overline{\mathbf{x}}}$ since

$$\mathbf{d}^T \nabla_x^2 L(2,0,-4)\mathbf{d} = -6d_2^2 \leq 0.$$

Consider the third KKT point:

$$\nabla_x^2 L(0,1,-1) = \begin{pmatrix} 3/2 & 0 \\ 0 & 0 \end{pmatrix}, T_{\overline{\mathbf{x}}} = \{(z_1, z_2) : z_2 = 0\}$$

Then the Hessian is positive definite on $T_{\overline{\mathbf{x}}}$ since

$$\mathbf{d}^T \nabla_x^2 L(0,0,-1)\mathbf{d} = (3/2)d_1^2 > 0, \forall \mathbf{0} \neq \mathbf{d} \in T_{\overline{\mathbf{x}}}.$$

This would be sufficient for the third KKT solution to be a local minimizer.

We also try to apply the KKT condition on the noncovex problem, like spherical constrained nonconvex quadratic optimization

$$\begin{array}{lll} (SCQP) & \min & \mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} \\ & \text{s.t.} & \|\mathbf{x}\|^2 (\leq, =) 1 \end{array} \tag{3.22}$$

Here we can also relax the norm constraint to the SDP. And have rank 1 solution

Theorem 3.21. *The FONC and SONC, that is, the following conditions on* x*, together with the multiplier y,*

$$\begin{array}{rcl} \|\mathbf{x}\|^2 & (\leq, =) & 1, (OPC) \\ 2Q\mathbf{x} + \mathbf{c} - 2y\mathbf{x} & = & 0, (LDC) \\ y \quad \left(\leq, \textit{ free }'\right) & 0, (MSC) & \\ y\left(1 - \|\mathbf{x}\|^2\right) & = & 1, (CSC) \\ (Q - yI) & \succeq & \mathbf{0}, (SOC) \end{array}$$

are necessary and sufficient for finding the global minimizer of (SCQP).

3.10 Lagrangian Duality Theory

A bullet point is the conic duality form is the subset of Lagrangian duality. Recall the dual of conic linear program

$$\begin{array}{lll} (CLP) & \min & \mathbf{c} \bullet \mathbf{x} \\ & s.t. & \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, \ldots, m, \mathbf{x} \in K; \end{array}$$

and it dual problem

$$\begin{aligned}(CLD) \quad \max \quad & \mathbf{b}^T\mathbf{y} \\ s.t. \quad & \sum_i^m y_i\mathbf{a}_i + \mathbf{s} = \mathbf{c}, \mathbf{s} \in K^*,\end{aligned}$$

where $\mathbf{y} \in \mathcal{R}^m$, s is called the dual slack vector/matrix, and K^* is the dual cone of K. In general, K can be decomposed to $K = K_1 \oplus K_2 \oplus \ldots \oplus K_p$, that is,

$$\mathbf{x} = (\mathbf{x}_1; \mathbf{x}_2; \ldots; \mathbf{x}_p), \mathbf{x}_i \in K_i, \forall i$$

Note that $K^* = K_1^* \oplus K_2^* \oplus \ldots \oplus K_p^*$, or

$$\mathbf{s} = (\mathbf{s}_1; \mathbf{s}_2; \ldots; \mathbf{s}_p), \mathbf{s}_i \in K_i^*, \forall i.$$

This is a powerful but very structured duality form. We now develop the Lagrangian Duality theory as an alternative to Conic Duality theory. For general nonlinear constraints, the Lagrangian Duality theory is more applicable.

We explain the Lagrangian duality idea by toy example

$$\begin{aligned}\min \quad & (x_1-1)^2 + (x_2-1)^2 \\ s.t. \quad & x_1 + 2x_2 - 1 \le 0 \\ & 2x_1 + x_2 - 1 \le 0\end{aligned}$$

$$L(\mathbf{x},\mathbf{y}) = f(\mathbf{x}) - \mathbf{y}^T\mathbf{c}(\mathbf{x}) = f(\mathbf{x}) - \sum_{i=1}^2 y_i c_i(\mathbf{x}) =$$
$$= (x_1-1)^2 + (x_2-1)^2 - y_1(x_1+2x_2-1) - y_2(2x_1+x_2-1), (y_1;y_2) \le \mathbf{0}$$

where

$$\nabla L_x(\mathbf{x},\mathbf{y}) = \begin{pmatrix} 2(x_1-1) - y_1 - 2y_2 \\ 2(x_2-1) - 2y_1 - y_2 \end{pmatrix}$$

For given multipliers $\mathbf{y} \in Y$, consider problem

$$\begin{aligned}(LRP) \quad \inf \quad & L(\mathbf{x},\mathbf{y}) = f(\mathbf{x}) - \mathbf{y}^T\mathbf{c}(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in R^n\end{aligned}$$

Again, $\mathbf{y}_i$ can be viewed as a **penalty parameter** to penalize constraint violation $c_i(\mathbf{x}), i = 1, \ldots, m$. In the toy example, for given $(y_1; y_2) \le \mathbf{0}$, the LRP is:

$$\inf (x_1-1)^2 + (x_2-1)^2 - y_1(x_1+2x_2-1) - y_2(2x_1+x_2-1)$$

s.t. $(x_1; x_2) \in R^2$, let above infimum formula equal 0, and it has a close form solution x for any given y :

$$x_1 = \frac{y_1+2y_2}{2} + 1 \quad \text{and} \quad x_2 = \frac{2y_1+y_2}{2} + 1$$

with the minimal or infimum value function $= -1.25y_1^2 - 1.25y_2^2 - 2y_1y_2 - 2y_1 - 2y_2$. Note that the difference between minimum and infimum, the minimal solution is achievable but infimum is not.

For any y $\in Y$, the minimal value function (including unbounded from below or infeasible cases) and the Lagrangian Dual Problem (LDP) are given by:

$$\begin{array}{lll} \phi(\mathbf{y}) := & \inf_{\mathbf{x}} L(\mathbf{x},\mathbf{y}), \quad \text{s.t.} & \mathbf{x} \in R^n \\ (LDP) & \sup_{\mathbf{y}} \phi(\mathbf{y}), \quad \text{s.t.} & \mathbf{y} \in Y \end{array}$$

Theorem 3.22. *The Lagrangian dual objective $\phi(y)$ is a concave function*

Proof. For any given two multiply vectors $\mathbf{y}^1 \in Y$ and $\mathbf{y}^2 \in Y$,

$$\begin{aligned} \phi\left(\alpha \mathbf{y}^1 + (1-\alpha)\mathbf{y}^2\right) &= \inf_{\mathbf{x}} L\left(\mathbf{x}, \alpha \mathbf{y}^1 + (1-\alpha)\mathbf{y}^2\right) \\ &= \inf_{\mathbf{x}} \left[f(\mathbf{x}) - \left(\alpha \mathbf{y}^1 + (1-\alpha)\mathbf{y}^2\right)^T \mathbf{c}(\mathbf{x})\right] \\ &= \inf_{\mathbf{x}} \left[\alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{x}) - \alpha \left(\mathbf{y}^1\right)^T \mathbf{c}(\mathbf{x}) - (1-\alpha)\left(\mathbf{y}^2\right)^T \mathbf{c}(\mathbf{x})\right] \\ &= \inf_{\mathbf{x}} \left[\alpha L\left(\mathbf{x}, \mathbf{y}^1\right) + (1-\alpha) L\left(\mathbf{x}, \mathbf{y}^2\right)\right] \\ &\geq \alpha \left[\inf_{\mathbf{x}} L\left(\mathbf{x}, \mathbf{y}^1\right)\right] + (1-\alpha)\left[\inf_{\mathbf{x}} L\left(\mathbf{x}, \mathbf{y}^2\right)\right] \\ &= \alpha \phi\left(\mathbf{y}^1\right) + (1-\alpha)\phi\left(\mathbf{y}^2\right) \end{aligned}$$

□

Intuition, we have the weakly duality

Theorem 3.23. *(Weak duality theorem) For every $y \in Y$, the Lagrangian dual function $\phi(\mathbf{y})$ is less or equal to the infimum value of the original GCO problem.*

Proof.

$$\begin{aligned} \phi(\mathbf{y}) &= \inf_{\mathbf{x}} \left\{f(\mathbf{x}) - \mathbf{y}^T \mathbf{c}(\mathbf{x})\right\} \\ &\leq \inf_{\mathbf{x}} \left\{f(\mathbf{x}) - \mathbf{y}^T \mathbf{c}(\mathbf{x}) \text{ s.t. } \mathbf{c}(\mathbf{x})(\leq, =, \geq)\mathbf{0}\right\} \\ &\leq \inf_{\mathbf{x}} \{f(\mathbf{x}): \text{ s.t. } \mathbf{c}(\mathbf{x})(\leq, =, \geq)\mathbf{0}\} \end{aligned}$$

The first inequality is from the fact that the unconstrained inf-value is no greater than the constrained one. The second inequality is from $\mathbf{c}(\mathbf{x})(\leq, =, \geq)\mathbf{0}$ and $\mathbf{y}\,(\leq{}' \text{ free}, \geq)\,\mathbf{0}$ imply $-\mathbf{y}^T\mathbf{c}(\mathbf{x}) \leq 0$ □

In the toy example. we will have

$$\begin{aligned} \min \quad & (x_1 - 1)^2 + (x_2 - 1)^2 \\ s.t. \quad & x_1 + 2x_2 - 1 \leq 0 \\ & 2x_1 + x_2 - 1 \leq 0 \end{aligned}$$

where $\mathbf{x}^* = \left(\frac{1}{3}; \frac{1}{3}\right)$. The Lagrangian formula is

$$\phi(\mathbf{y}) = -1.25y_1^2 - 1.25y_2^2 - 2y_1y_2 - 2y_1 - 2y_2, \mathbf{y} \leq \mathbf{0}$$

The dual program is

$$\begin{aligned} \max \quad & -1.25y_1^2 - 1.25y_2^2 - 2y_1y_2 - 2y_1 - 2y_2 \\ s.t. \quad & (y_1; y_2) \le \mathbf{0} \end{aligned}$$

where y$^* = \left(\frac{-4}{9}; \frac{-4}{9}\right)$.

3.10.1 Example of Program

Example of Dual Linear Program

Consider LP problem

$$\begin{aligned} (LP) \quad \min \quad & \mathbf{c}^T\mathbf{x} \\ s.t. \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \ge \mathbf{0} \end{aligned}$$

and its conic dual problem is given by

$$\begin{aligned} (LD) \quad \max \quad & \mathbf{b}^T\mathbf{y} \\ s.t. \quad & A^T\mathbf{y} + \mathbf{s} = \mathbf{c} \\ & \mathbf{s} \ge \mathbf{0} \end{aligned}$$

We now derive the Lagrangian Dual of (LP). Let the Lagrangian multipliers be y('free') for equalities and s ≥ 0 for constraints x ≥ 0. Then the Lagrangian function would be

$$L(\mathbf{x}, \mathbf{y}, \mathbf{s}) = \mathbf{c}^T\mathbf{x} - \mathbf{y}^T(A\mathbf{x} - \mathbf{b}) - \mathbf{s}^T\mathbf{x} = \left(\mathbf{c} - A^T\mathbf{y} - \mathbf{s}\right)^T \mathbf{x} + \mathbf{b}^T\mathbf{y}$$

where x is "free".

Now consider the Lagrangian dual objective

$$\phi(\mathbf{y}, \mathbf{s}) = \inf_{\mathbf{x} \in R^n} L(\mathbf{x}, \mathbf{y}, \mathbf{s}) = \inf_{\mathbf{x} \in R^n} \left[\left(\mathbf{c} - A^T\mathbf{y} - \mathbf{s}\right)^T \mathbf{x} + \mathbf{b}^T\mathbf{y}\right].$$

If $\left(\mathbf{c} - A^T\mathbf{y} - \mathbf{s}\right) \ne \mathbf{0}$, then $\phi(\mathbf{y}, \mathbf{s}) = -\infty$. Thus, in order to maximize $\phi(\mathbf{y}, \mathbf{s})$, the dual must choose its variables $(\mathbf{y}, \mathbf{s} \ge \mathbf{0})$ such that $\left(\mathbf{c} - A^T\mathbf{y} - \mathbf{s}\right) = \mathbf{0}$. This constraint, together with the sign constraint s ≥ 0, establish the Lagrangian dual problem:

$$\begin{aligned} (LDP) \quad \max \quad & \mathbf{b}^T\mathbf{y} \\ s.t. \quad & A^T\mathbf{y} + \mathbf{s} = \mathbf{c}, \mathbf{s} \ge \mathbf{0}. \end{aligned}$$

which is consistent with the conic dual of LP.

Example of Dual Linear Program with Log-Barrier

For a fixed $\mu > 0$, consider the problem

$$\begin{array}{ll} \min & \mathbf{c}^T\mathbf{x} - \mu\sum_{j=1}^n \log(x_j) \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{array}$$

Again, the non-negativity constraints can be "ignored" if the feasible region has an "interior", that is, any minimizer must have $x(\mu) > 0$. Thus, the Lagrangian function would be simply given by

$$L(\mathbf{x},\mathbf{y}) = \mathbf{c}^T\mathbf{x} - \mu\sum_{j=1}^n \log(x_j) - \mathbf{y}^T(A\mathbf{x}-\mathbf{b}) = \left(\mathbf{c} - A^T\mathbf{y}\right)^T\mathbf{x} - \mu\sum_{j=1}^n \log(x_j) + \mathbf{b}^T\mathbf{y}$$

Then, the Lagrangian dual objective (we implicitly need x > 0 for the function to be defined)

$$\phi(\mathbf{y}) := \inf_{\mathbf{x}} L(\mathbf{x},\mathbf{y}) = \inf_{\mathbf{x}} \left[\left(\mathbf{c} - A^T\mathbf{y}\right)^T\mathbf{x} - \mu\sum_{j=1}^n \log(x_j) + \mathbf{b}^T\mathbf{y}\right].$$

First, from the view point of the dual, the dual needs to choose y such that c $- A^T\mathbf{y} > \mathbf{0}$, since otherwise the primal can choose $\mathbf{x} > 0$ to make $\phi(\mathbf{y})$ go to $-\infty$. Now for any given y such that c $- A^T\mathbf{y} > \mathbf{0}$, the inf problem has a unique finite close-form minimizer $\mathbf{x}$ by first order derivative equaling zero

$$x_j = \frac{\mu}{\left(\mathbf{c} - A^T\mathbf{y}\right)_j}, \forall j = 1,\ldots,n.$$

Thus,

$$\phi(\mathbf{y}) = \mathbf{b}^T\mathbf{y} + \mu\sum_{j=1}^n \log\left(\mathbf{c} - A^T\mathbf{y}\right)_j + n\mu(1-\log(\mu))$$

Therefore, the dual problem, for any fixed μ, can be written as

$$\max_{\mathbf{y}} \phi(\mathbf{y}) = n\mu(1-\log(\mu)) + \max_{\mathbf{y}}\left[\mathbf{b}^T\mathbf{y} + \mu\sum_{j=1}^n \log\left(\mathbf{c} - A^T\mathbf{y}\right)_j\right]$$

This is actually the LP dual with the Log-Barrier on dual inequality constraints c $- A^T\mathbf{y} \geq \mathbf{0}$.

Example of Dual Linear Program with Fisher Market

$$\begin{array}{ll} \max & \sum_{i\in B} w_i \log\left(\mathbf{u}_i^T\mathbf{x}_i\right) \\ \text{s.t.} & \sum_{i\in B}\mathbf{x}_i = \mathbf{b}, \quad \forall j \in G \\ & x_{ij} \geq 0, \quad \forall i,j \end{array}$$

The Lagrangian function would be simply given by

$$L\left(\mathbf{x}_i \geq \mathbf{0}, i \in B, \mathbf{y}\right) = \sum_{i\in B} w_i \log\left(\mathbf{u}_i^T\mathbf{x}_i\right) - \mathbf{y}^T\left(\sum_{i\in B}\mathbf{x}_i - \mathbf{b}\right) = \sum_{i\in B}\left(w_i\log\left(\mathbf{u}_i^T\mathbf{x}_i\right) - \mathbf{y}^T\mathbf{x}_i\right) + \mathbf{b}^T\mathbf{y}$$

Then, the Lagrangian dual objective, for any given $y > 0$, would be

$$\phi(\mathbf{y}) := \inf_{\mathbf{x}_i \geq \mathbf{0}, i \in B} L(\mathbf{x}_i, i \in B, \mathbf{y}) = \inf_{\mathbf{x}_i \geq \mathbf{0}, i \in B} \sum_{i \in B} \left(w_i \log\left(\mathbf{u}_i^T \mathbf{x}_i\right) - \mathbf{y}^T \mathbf{x}_i \right) + \mathbf{b}^T \mathbf{y}$$

Similar idea we use in the linear program with barrier function. For each $i \in B$, the sup-solution is

$$x_{ij^*} = \frac{w_i}{y_{j^*}} > 0, j^* = \arg\min_j \frac{y_j}{u_{ij}}, x_{ij} = 0 \forall j \neq j^*.$$

Thus,

$$\phi(\mathbf{y}) = \mathbf{b}^T \mathbf{y} - \sum_{i \in B} w_i \log\left(\min_j \left[\frac{y_j}{u_{ij}}\right]\right) + \sum_{i \in B} w_i \left(\log(w_i) - 1\right).$$

3.10.2 Lagrangian Strong Duality Theorem

Theorem 3.24. *Theorem 3 Let* (GCO) *be a convex minimization problem and the infimum* f^* *of* (GCO) *be finite, and the suprermum of* (LDP) *be* ϕ^*. *In addition, let* (GCO) *have an interior-point feasible solution with respect to inequality constraints, that is, there is* $\hat{\mathrm{x}}$ *such that all inequality constraints are strictly held. Then,* $f^* = \phi^*$, *and* (LDP) *admits a maximizer* y^* *such that*

$$\phi(\mathbf{y}^*) = f^*.$$

Furthermore, if (GCO) *admits a minimizer* x^*, *then*

$$y_i^* c_i(\mathbf{x}^*) = 0, \forall i = 1, \ldots, m.$$

The assumption of "interior-point feasible solution" is called Constraint Qualification condition, which was also needed as a condition to prove the strong duality theorem for general Conic Linear Optimization.

Proof. Consider the convex set

$$C := \{(\kappa; \mathbf{s}) : \exists \mathbf{x} \text{ s.t. } f(\mathbf{x}) \leq \kappa, -\mathbf{c}(\mathbf{x}) \leq \mathbf{s}\}.$$

Then, $(f^*; 0)$ is on the closure of C. From the supporting hyperplane theorem, there exists $(y_0^*; \mathbf{y}^*) \neq 0$ such that

$$y_0^* f^* \leq \inf_{(\kappa; \mathbf{s}) \in C} \left(y_0^* \kappa + (\mathbf{y}^*)^T \mathbf{s}\right)$$

First, we show $\mathrm{y}^* \geq \mathbf{0}$, since otherwise one can choose some $(0; \mathrm{s} \geq 0)$ such that the inequality is violated. Secondly, we show $y_0^* > 0$, since otherwise one can choose $(\kappa \to \infty; 0)$ if $y^* < 0$, or $(0; \mathbf{s} = -\mathbf{c}(\hat{\mathbf{x}}) < 0)$ if $y^* = 0$ (then $\mathbf{y}^* \neq 0$), such that the above inequality is violated. Now let us divide both sides by y_0^* and let $\mathbf{y}^* := \mathbf{y}^*/y_0^*$, we have

$$f^* \leq \inf_{(\kappa; \mathbf{s}) \in C} \left(\kappa + (\mathbf{y}^*)^T \mathbf{s}\right) = \inf_{\mathbf{x}} \left(f(\mathbf{x}) - (\mathbf{y}^*)^T \mathbf{c}(\mathbf{x})\right) = \phi(\mathbf{y}^*) \leq \phi^*.$$

Then, from the weak duality theorem, we must have $f^* = \phi^*$

If (GCO) admits a minimizer $\mathbf{x}^*$, then $f(\mathbf{x}^*) = f^*$ so that

$$f(\mathbf{x}^*) \leq \inf_{\mathbf{x}} \left[f(\mathbf{x}) - (\mathbf{y}^*)^T \mathbf{c}(\mathbf{x}) \right] \leq f(\mathbf{x}^*) - (\mathbf{y}^*)^T \mathbf{c}(\mathbf{x}^*) = f(\mathbf{x}^*) - \sum_i^m y_i^* c_i(\mathbf{x}^*),$$

which implies that

$$\sum_i^m y_i^* c_i(\mathbf{x}^*) \leq 0.$$

Since $y_i^* \geq 0$ and $c_i(\mathbf{x}^*) \geq 0$ for all i, it must be true $y_i^* c_i(\mathbf{x}^*) = 0$ for all i. □

Theorem 3.25. *The Lagrangian dual function $\phi(y)$ is a concave function*

The proof of theorem is 3.10

3.10.3 Support Vector Machine

The primal formulation is

$$\begin{aligned} \min_{\mathbf{x}, x_0, \beta} \quad & \beta + \mu \|\mathbf{x}\|^2 \\ s.t. \quad & \mathbf{a}_i^T \mathbf{x} + x_0 + \beta \geq 1, \forall i, (\mathbf{y}_a \geq \mathbf{0}) \\ & -\mathbf{b}_j^T \mathbf{x} - x_0 + \beta \geq 1, \forall j, (\mathbf{y}_b \geq \mathbf{0}) \\ & \beta \geq 0. (\alpha \geq 0) \end{aligned} \tag{3.23}$$

The Lagrangian formula is

$$L(\mathbf{x}, x_0, \beta, \mathbf{y}_a, \mathbf{y}_b, \alpha) = \beta + \mu \|\mathbf{x}\|^2 - \mathbf{y}_a^T \left(A^T \mathbf{x} + x_0 \mathbf{e} + \beta \mathbf{e} - \mathbf{e} \right) - \mathbf{y}_b^T \left(-B^T \mathbf{x} - x_0 \mathbf{e} + \beta \mathbf{e} - \mathbf{e} \right) - \alpha \beta$$

First order derivative condition

$$\begin{aligned} \nabla_{\mathbf{x}} L(\cdot) &= 2\mu \mathbf{x} - A \mathbf{y}_a + B \mathbf{y}_b = \mathbf{0}, (\text{ replace } \mathbf{x}) \\ \nabla_{x_0} L(\cdot) &= -\mathbf{e}^T \mathbf{y}_a + \mathbf{e}^T \mathbf{y}_b = 0, (\text{dual constraint}) \\ \nabla_{\beta} L(\cdot) &= 1 - \mathbf{e}^T \mathbf{y}_a - \mathbf{e}^T \mathbf{y}_b - \alpha = 0. (\text{dual constraint}) \end{aligned} \tag{3.24}$$

Then the dual objective is

$$\frac{-1}{4\mu} \|A \mathbf{y}_a - B \mathbf{y}_b\|^2 + \mathbf{e}^T \mathbf{y}_a + \mathbf{e}^T \mathbf{y}_b$$

In the dual side, we can understand the support vector machine. The fundamentation is to separate two sets.

In addition, some problem could not be straight-ward. We cannot use the linear algebra or first order derivative to substitute. The dual formulation structure depends different kind of primal problem. Sometimes the dual can be constructed by simple reasoning: consider

$$\begin{aligned} (LP) \quad \min \quad & \mathbf{c}^T \mathbf{x} \\ s.t. \quad & A\mathbf{x} = \mathbf{b}, -\mathbf{e} \leq \mathbf{x} \leq \mathbf{e} \\ & (\|\mathbf{x}\|_\infty \leq 1) \end{aligned}$$

Let the Lagrangian multipliers be y for equality constraints. Then the Lagrangian dual objective would be

$$\phi(\mathbf{y}) = \inf_{-\mathrm{e} \le \mathbf{x} \le \mathrm{e}} L(\mathbf{x}, \mathbf{y}) = \inf_{-\mathrm{e} \le \mathrm{x} \le \mathrm{e}} \left[\left(\mathbf{c} - A^T \mathbf{y}\right)^T \mathbf{x} + \mathbf{b}^T \mathbf{y} \right]$$

where if $\left(\mathbf{c} - A^T\mathbf{y}\right)_j \le 0, x_j = 1$; and otherwise, $x_j = -1$.

Therefore, the Lagrangian dual is

$$\begin{array}{lll} (LDP) & \max & \mathbf{b}^T\mathbf{y} - \left\|\mathbf{c} - A^T\mathbf{y}\right\|_1 \\ s.t. & & \mathbf{y} \in R^m \end{array}$$

3.10.4 Farkas Lemma for Nonlinear Constraints

Consider the convex constrained system:

$$\begin{array}{lll} \text{(CCS)} & \min & \mathbf{0}^T\mathbf{x} \\ & \text{s.t.} & c_i(\mathbf{x}) \ge 0, i = 1, \ldots, m, \end{array}$$

where $c_i(.)$ are concave functions and the Lagrangian Function is given by

$$L(\mathbf{x}, \mathbf{y}) = -\mathbf{y}^T\mathbf{c}(\mathbf{x}) = -\sum_{i=1}^{m} y_i c_i(\mathbf{x}), \mathbf{y} \ge \mathbf{0}$$

Again, let

$$\phi(\mathbf{y}) := \inf_{\mathbf{x}} L(\mathbf{x}, \mathbf{y}).$$

Theorem 3.26. *If there exists $y \ge 0$ such that $\phi(y) > 0$, then (CSS) is infeasible. The proof is directly from the dual objective function $\phi(\mathbf{y})$ is a homogeneous function and the dual has its objective value unbounded from above.*

3.10.5 The Conic Duality vs. Lagrangian Duality

Consider SOCP problem

$$\begin{array}{lll} (SOCP) & \min & \mathbf{c}^T\mathbf{x} \\ & s.t. & A\mathbf{x} = \mathbf{b} \\ & & x_1 - \|\mathbf{x}_{-1}\|_2 \ge 0 \end{array} \tag{3.25}$$

and it conic dual problem

$$\begin{array}{lll} (SOCD) & \max & \mathbf{b}^T\mathbf{y} \\ & s.t. & A^T\mathbf{y} + \mathbf{s} = \mathbf{c} \\ & & s_1 - \|\mathbf{s}_{-1}\|_2 \ge 0 \end{array} \tag{3.26}$$

Let the Lagrangian multipliers be y for equalities and scalar $s \ge 0$ for the single constraint $x_1 \ge \|\mathrm{x}_{-1}\|_2$. Then the Lagrangian function would be

$$L(\mathbf{x}, \mathbf{y}, s) = \mathbf{c}^T\mathbf{x} - \mathbf{y}^T(A\mathbf{x} - \mathbf{b}) - s\left(x_1 - \|\mathbf{x}_{-1}\|_2\right) = \left(\mathbf{c} - A^T\mathbf{y}\right)^T \mathbf{x} - s\left(x_1 - \|\mathbf{x}_{-1}\|_2\right) + \mathbf{b}^T\mathbf{y}$$

Now consider the Lagrangian dual objective

$$\phi(\mathbf{y}, s) = \inf_{\mathbf{x}\in R^n} L(\mathbf{x}, \mathbf{y}, \mathbf{s}) = \inf_{\mathbf{x}\in R^n} \left[\left(\mathbf{c} - A^T\mathbf{y}\right)^T \mathbf{x} - s\left(x_1 - \|\mathbf{x}_{-1}\|_2\right) + \mathbf{b}^T\mathbf{y} \right].$$

The objective function of the problem may not be differentiable so that the classical optimal condition theory do not apply. Consequently, it is difficult to write a clean/explicit form of the Lagrangian dual problem.

On the other hand, many nonlinear optimization problems, even they are convex, are difficult to transform them into structured CLP problems (especially to construct the dual cones). Therefore, each of the duality form, Conic or Lagrangian, has its own pros and cons.

3.11 Conic Duality

This section heavily from this course webpage. We will discuss the conic duality. The bullet points is when we deal with the conic program, 99% first step to try the conic duality.

The conic optimization problem in standard equality form is:

$$p^\star := \min_x : c^T x \;\;:\;\; Ax = b, x \in \mathcal{K}$$

where cal K is a proper cone, for example a direct product of cones that are one of the three types: positive orthant, second-order cone, or semidefinite cone. Let cal $\mathcal{K}^\star$ be the cone dual cal K, which we define as ($\mathcal{K}^\star = \{\lambda : \forall x \in \mathcal{K}, \lambda^T x \geq 0\}$.) All cones we mentioned (positive orthant, second-order cone, or semidefinite cone, and products thereof), are self-dual, in the sense that $\mathcal{K}^\star = \mathcal{K}$.

The Lagrangian of the problem is given by

$$\mathcal{K}^\star(x, \lambda, y) = c^T x + y^T(b - Ax) - \lambda^T x$$

The last term is added to take account of the constraint $x \in \mathcal{K}$x in cal K. From the very definition of the dual cone:

$$\max_{\lambda\in\mathcal{K}^\star} \quad -\lambda^T x = \begin{cases} 0 & \text{if} x \in \mathcal{K} \\ +\infty & \text{otherwise} \end{cases}$$

Thus, we have

$$\begin{aligned} p^\star \quad &= \min_x \max_{y,\lambda\in\mathcal{K}^\star} \mathcal{L}(x, \lambda, y) \\ &= \min_x \max_{y,\lambda\in\mathcal{K}^\star} c^T x + y^T(b - Ax) - \lambda^T x \\ &\geq d^\star := \max_{y,\lambda\in\mathcal{K}^\star} g(\lambda, y) \end{aligned}$$

where

$$g(\lambda, y) = \min_x c^T x + y^T(b - Ax) - \lambda^T x = \begin{cases} y^T b & \text{if} c - A^T y - \lambda = 0 \\ -\infty & \text{otherwise} \end{cases}$$

The dual for the problem is:

$$d^\star = \max y^T b : c - A^T y - \lambda = 0, \lambda \in \mathcal{K}^\star$$

Eliminating λ, we can simplify the dual as

$$d^\star = \max y^T b : c - A^T y \in \mathcal{K}^\star$$

Note that dual cone $\mathcal{K}^\star = \{y :< x, y >\geq 0, x \in \mathcal{K}\}$

3.11.1 Strong duality and KKT conditions

Conditions for Strong Duality

We now summarize the results stated before. Strong duality holds when either:

- The primal is strictly feasible, i.e. $\exists x : Ax = b, x \in \operatorname{int}(\mathcal{K})$. This also implies that the dual problem is attained
- The dual is strictly feasible, i.e. $\exists y : c - A^T y \in \operatorname{int}(\mathcal{K}^*)$. This also implies that the primal problem is attained. item If both the primal and dual are strictly feasible then both are attained (and $p^* = d^*$)

KKT Conditions

Assume $p^* = d^*$ and both the primal and dual are attained by some primal-dual triplet (x^*, λ^*, y^*) . Then,

$$\begin{aligned} p^* = c^T x^* = d^* &= g(\lambda^*, y^*) \\ &= \min_x \mathcal{L}(x, \lambda^*, y^*) \\ &\leq \mathcal{L}(x^*, \lambda^*, y^*) \\ &= c^T x^* - \lambda^{*T} x^* + y^{*T}(b - Ax^*) \\ &\leq c^T x^* = p^* \end{aligned}$$

The last term in the fourth line is equal to zero which implies $\lambda^{*T} x^* = 0$. Thus the KKT conditions are:

1. $x \in \mathcal{K}, Ax = b$
2. $\lambda \in \mathcal{K}^*$
3. $\lambda^T x = 0$
4. $c - A^T y - \lambda = 0$, that is, $\nabla_x \mathcal{L}(x, \lambda, y) = 0$

Eliminating λ from the above allows us to get rid of the Lagrangian stationarity condition, and gives us the following theorem.

Theorem 3.27. *The conic problem*

$$p^* := \min_x c^T x : Ax = b, \quad x \in \mathcal{K}.$$

admits the dual bound $p^* \geq d^*$, where

$$d^* = \max y^T b : c - A^T y \in \mathcal{K}^*.$$

If both problems are strictly feasible, then the duality gap is zero: $p^* = d^*$, and both values are attained. Then, a pair (x, y) is primal-dual optimal if and only if the KKT conditions

- *Primal feasibility: $x \in \mathcal{K}, Ax = b$*
- *Dual feasibility: $c - A^T y \in \mathcal{K}^*$*
- *Complementary slackness: $\left(c - A^T y\right)^T x = 0$, hold*

3.11.2 SDP Duality

In this section, we consider the SDP in standard form:

$$\begin{aligned} p^* := \max_X & \langle C, X \rangle \\ s.t. \quad & \langle A_i, X \rangle = b_i, \quad i = 1, \ldots, m \\ & X \succeq 0 \end{aligned} \tag{3.27}$$

where C, A_i are given symmetric matrices, $\langle A, B \rangle = \mathrm{Tr}(AB)$ denotes the scalar product between two symmetric matrices, and $b \in \mathbf{R}^m$ is given. Note that

$$tr(A) \sum_i a_{ii}$$

Conic Lagrangian

Consider the "conic" Lagrangian

$$\mathcal{L}(X, \nu, Y) := \langle C, X \rangle + \sum_{i=1}^m \nu_i \left(b_i - \langle A_i, X \rangle\right) + \langle Y, X \rangle,$$

where now we associate a matrix dual variable Y to the constraint $X \succeq 0$. Let us check that the Lagrangian above "works", in the sense that we can represent the constrained maximization problem as an unconstrained, maximin problem:

$$p^* = \max_X \min_{Y \succeq 0} \mathcal{L}(X, \nu, Y).$$

This is an immediate consequence of the following:

$$\min_{Y \succeq 0} \langle Y, X \rangle = \min_{t \geq 0} \min_{Y \succeq 0, \mathrm{Tr}\, Y = t} \langle Y, X \rangle = \min_{t \geq 0} t \lambda_{\min}(X),$$

where we have exploited the representation of the minimum eigenvalue. The geometric interpretation is that the cone of positive-semidefinite matrices has a 90° angle at the origin.

Dual Problem

The minimax inequality then implies

$$p^* \leq d^* := \min_{\nu, Y \succeq 0} \max_X \mathcal{L}(X, \nu, Y).$$

The corresponding dual function is

$$g(Y, \nu) = \max_X \mathcal{L}(X, \nu, Y) = \begin{cases} \nu^T b & \text{if } C - \sum_{i=1}^m \nu_i A_i + Y = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem then writes

$$d^* = \min_{\nu, Y \succeq 0} g(Y, \nu) = \min_{\nu, Y \succeq 0} \nu^T b : C - \sum_{i=1}^m \nu_i A_i = -Y \preceq 0.$$

After elimination of the variable Y, we find the dual

$$d^* = \min_\nu \nu^T b : C - \sum_{i=1}^m \nu_i A_i \preceq 0. \tag{3.28}$$

which is in standard inequality form.

Theorem 3.28. *(Strong duality in SDP) Consider the SDP*

$$p^* := \max_X \langle C, X \rangle : \langle A_i, X \rangle = b_i, \quad i = 1, \ldots, m, X \succeq 0$$

and its dual

$$d^* = \min_\nu \nu^T b : \sum_{i=1}^m \nu_i A_i \succeq C.$$

The following holds:

1. *Duality is symmetric, in the sense that the dual of the dual is the primal*
2. *Weak duality always holds: $p^* \leq d^*$, so that, for any primal-dual feasible pair (X, ν), we have $\nu^T b \geq \langle C, X \rangle$*
3. *If the primal (resp.dual) problem is bounded above (resp. below), and strictly feasible, then $p^* = d^*$ and the dual (resp. primal) is attained*
4. *If both problems are strictly feasible, then $p^* = d^*$ and both problems are attained*

3.11.3 SOCP Duality

We start from the second-order cone problem in inequality form:

$$\begin{aligned} p^* := \min_x \quad & c^T x \\ s.t. \quad & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m, \end{aligned}$$

where $c \in \mathbf{R}^n, A_i \in \mathbf{R}^{n_i \times n}, b_i \in \mathbf{R}^{n_i}, c_i \in \mathbf{R}^n, d_i \in \mathbf{R}, i = 1, \ldots, m$.

Conic Lagrangian

To build a Lagrangian for this problem, we use the fact that, for any pair (t, y) :

$$\max_{(u,\lambda):\|u\|_2\leq\lambda} -u^T y - t\lambda = \max_{\lambda\geq 0} \lambda\left(\|y\|_2 - t\right) = \begin{cases} 0 & \text{if } \|y\|_2 \leq t \\ +\infty & \text{otherwise.} \end{cases}$$

The above means that the second-order cone has a 90° angle at the origin. To see this, observe that

$$\max_{(u,\lambda):\|u\|_2\leq\lambda} -u^T y - t\lambda = -\min_{(u,\lambda):\|u\|_2\leq\lambda} \begin{pmatrix} u \\ \lambda \end{pmatrix}^T \begin{pmatrix} y \\ t \end{pmatrix}.$$

The objective in the right-hand side is proportional to the cosine of the angle between the vectors involved. The largest angle achievable between any two vectors in the second-order cone is 90°. If $\|y\|_2 > t$, then the cosine reaches negative values, and the maximum scalar product becomes infinite.

Geometric interpretation of the 90° angle at the origin property. The two orthogonal Dalt vectors in black form the maximum angle attainable by vector in the second-order cone. text The vector in red forms a greater angle with the vector on the left, and the corresponding scalar product is unbounded. Consider the following Lagrangian, with variables $x, \lambda \in \mathbf{R}^m, u_i \in \mathbf{R}^{n_i}, i = 1, \ldots, m$:

$$\mathcal{L}(x, \lambda, u_1, \ldots, u_m) = c^T x - \sum_{i=1}^m \left[u_i^T\left(A_i x + b_i\right) + \lambda_i\left(c_i^T x + d_i\right)\right].$$

Using the fact above leads to the following minimax representation of the primal problem:

$$p^* = \min_x \max_{\|u_i\|_2\leq\lambda_i, i=1,\ldots,m} \mathcal{L}(x, \lambda, u_1, \ldots, u_m).$$

Conic dual

Weak duality expresses as $p^* \geq d^*$, where

$$d^* := \max_{\|u_i\|_2\leq\lambda_i, i=1,\ldots,m} \min_x \mathcal{L}(x, \lambda, u_1, \ldots, u_m).$$

The inner problem, which corresponds to the dual function, is very easy to solve as the problem is unconstrained and the objective affine (in x). Setting the derivative with respect to x leads to the dual constraints

$$c = \sum_{i=1}^m \left[A_i^T u_i + \lambda_i c_i\right]$$

We obtain

$$\begin{aligned} d^* = \max_{\lambda, u_i, i=1,\ldots,m} & -\lambda^T d - \sum_{i=1}^m u_i^T b_i \\ s.t. \quad & c = \sum_{i=1}^m \left[A_i^T u_i + \lambda_i c_i\right], \quad i = 1, \ldots, m \\ & \|u_i\|_2 \leq \lambda_i, \quad i = 1, \ldots, m \end{aligned}$$

The above is an SOCP, just like the original one.

Theorem 3.29. *(Strong duality in SOCP) Consider the SOCP*

$$
\begin{aligned}
p^* := \min_x \quad & c^T x \\
s.t. \quad & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m,
\end{aligned}
\tag{3.29}
$$

and its dual

$$
\begin{aligned}
d^* = \max_{\lambda, u_i, i=1,\ldots,m} & - \lambda^T d - \sum_{i=1}^m u_i^T b_i \\
s.t. \quad & c = \sum_{i=1}^m \left[A_i^T u_i + \lambda_i c_i \right], \quad i = 1, \ldots, m \\
& \|u_i\|_2 \leq \lambda_i, \quad i = 1, \ldots, m
\end{aligned}
\tag{3.30}
$$

The following holds:

1. *Duality is symmetric, in the sense that the dual of the dual is the primal*
2. *Weak duality always holds: $p^* \geq d^*$, so that, for any primal-dual feasible pair $(x, (u_i, \lambda_i)_{i=}^m)$, we have $\lambda^T d + \sum_{i=1}^m u_i^T b_i \leq c^T x$.*
3. *If the primal (resp. dual) problem is bounded above (resp. below), and strictly feasible, then $p^* = d^*$ and the dual (resp. primal) is attained*
4. *If both problems are strictly feasible, then $p^* = d^*$ and both problems are attained*

Chapter 4

Optimization Algorithm

Ref Chapters 4.2, 8.4-5, 9.1-7, 12.3-6

4.1 Optimization Algorithm

Optimization algorithms tend to be iterative procedures. Starting from a given point x^0, they generate a sequence $\{\mathrm{x}^k\}$ of iterates (or trial solutions) that converge to a "solution" - or at least they are designed to be so.

Recall that scalars $\{x^k\}$ converges to 0 if and only if for all real numbers $\varepsilon > 0$ there exists a positive integer K such that

$$|x^k| < \varepsilon \quad \text{for all } k \geq K.$$

Then $\{\mathrm{x}^k\}$ converges to solution x^* if and only if $\{\|\mathrm{x}^k - \mathrm{x}^*\|\}$ converges to 0 . We study algorithms that produce iterates according to

- well determined rules-Deterministic Algorithm
- random selection process-Randomized Algorithm.

The rules to be followed and the procedures that can be applied depend to a large extent on the characteristics of the problem to be solved. Generally, they are several kinds of solution, such as local optimal, global optimal, first order derivative, and KKT point.

4.1.1 General Algorithm

A General Algorithm: A point to set mapping in a subspace of R^n.

Theorem 4.1. *(Page 222, Luenberger et al. (1984)) Let A be an "algorithmic mapping" defined over set X, and let sequence $\left\{\mathrm{x}^k\right\}$, starting from a given point $\mathbf{x}^0$, be generated from*

$$\mathbf{x}^{k+1} \in A\left(\mathbf{x}^k, \ldots\right)$$

Let a solution set $S \subset X$ be given, and suppose

1. all points $\left\{\mathrm{x}^k\right\}$ are in a compact set
2. there is a continuous (merit) function $z(\mathbf{x})$ such that if $\mathbf{x} \notin S$, then $z(\mathbf{y}) < z(\mathbf{x})$ for all $\mathbf{y} \in A(\mathbf{x})$; otherwise, $z(\mathbf{y}) \leq z(\mathbf{x})$ for all $\mathbf{y} \in A(\mathbf{x})$
3. the mapping A is closed at points outside S ($\mathrm{x}^k \rightarrow \overline{\mathrm{x}} \in X$ and $A\left(\mathrm{x}^k\right) = \mathrm{y}^k \rightarrow \overline{\mathbf{y}}$ imply $\overline{\mathbf{y}} \in A(\overline{\mathbf{x}})$

Then, the limit of any convergent subsequences of $\left\{\mathbf{x}^k\right\}$ is a solution in S. Note that compact set means the set is closed and bounded. And it is the general optimization idea, how to solve the specific problem is case by case. In the term of $\mathbf{x}^{k+1} \in A\left(\mathbf{x}^k, \ldots\right)$, it will be two cases. case 1 we use x^k. case 2 we use all previous information such as $x_0, ..., x_k$

Convergence Rate of Iterative Method

This section content is from Senning (2007).

Definition 4.2. *If a sequence $x_1, x_2, \ldots, x_n$ converges to a value r and if there exist real numbers $\lambda > 0$ and $\alpha \geq 1$ such that*

$$\lim_{n \rightarrow \infty} \frac{\left|x_{n+1} - r\right|}{\left|x_n - r\right|^\alpha} = \lambda \tag{4.1}$$

then we say that α is the rate of convergence of the sequence

When $\alpha = 1$ we say the sequence converges linearly and when $\alpha = 2$ we say the sequence converges quadratically. If $1 < \alpha < 2$ then the sequence exhibits superlinear convergence.

Many root-finding methods are fixed-point iterations. These iterations have this name because the desired root r is a fixed-point of a function $g(x)$, i.e., $g(r) \rightarrow r$. To be useful for finding roots, a fixed-point iteration should have the property that, for x in some neighborhood of $r, g(x)$ is closer to r than x is. This leads to the iteration

$$x_{n+1} = g\left(x_n\right), \quad n = 0, 1, 2, \ldots$$

Newton's method is an lexample of a fixed-point iteration since

$$x_{n+1} = g\left(x_n\right), \quad g(x) = x - \frac{f(x)}{f'(x)} \tag{4.2}$$

and clearly $g(r) = r$ since $f(r) = 0$.

Theorem 4.3. *Let r be a fixed-point of the iteration $x_{n+1} = g(x_n)$ and suppose that $g'(r) \neq 0$. Then the iteration will have a linear rate of convergence.*

Proof. Using Taylor's Theorem for an expansion about fixed-point r we find

$$g(x) = g(r) + g'(r)(x - r) + \frac{g''(\xi)}{2}(x - r)^2$$

where ξ is some value between x and r. Evaluating at x_n and noting that $x_{n+1} = g(x_n)$ and $g(r) = r$ we obtain

$$x_{n+1} = r + g'(r)(x_n - r) + \frac{g''(\xi)}{2}(x_n - r)^2.$$

Subtracting r from both sides and dividing by $x_n - r$ gives

$$\frac{x_{n+1} - r}{x_n - r} = g'(r) + \frac{g''(\xi)}{2}(x_n - r)$$

which, as $n \to \infty$, yields

$$\lim_{n\to\infty} \frac{|x_{n+1} - r|}{|x_n - r|} = |g'(r)|.$$

Comparing this with Equation 4.1 we see that $\alpha = 1$ and $\lambda = |g'(r)|$, indicating that the method converges linearly. $\square$

We use the Taylor expansion to approximate the original function, and $g''(\xi)$ is small term comparing with $g'(r)$. Next, consider the case when $g'(r) = 0$. This is important because it explains why Newton's method converges so quickly .

Theorem 4.4. *Let r be a fixed-point of the iteration $x_{n+1} = g(x_n)$ and suppose that $g'(r) = 0$ but $g''(r) \neq 0$. Then the iteration will have a quadratic rate of convergence.*

Proof. Using Taylor's Theorem once again, but including one more term, we have

$$g(x) = g(r) + g'(r)(x - r) + \frac{g''(r)}{2}(x - r)^2 + \frac{g'''(\xi)}{6}(x - r)^3.$$

As before, we substitute x_n for x and use the facts that $x_{n+1} = g(x_n), g(r) = r$, and $g'(r) = 0$ to obtain

$$x_{n+1} = r + \frac{g''(r)}{2}(x_n - r)^2 + \frac{g'''(\xi)}{6}(x_n - r)^3.$$

Subtracting r from both sides and dividing by $(x_n - r)^2$ gives

$$\frac{x_{n+1} - r}{(x_n - r)^2} = \frac{g''(r)}{2} + \frac{g'''(\xi)}{6}(x_n - r)$$

which, as $n \to \infty$, gives

$$\lim_{n\to\infty} \frac{|x_{n+1} - r|}{|x_n - r|^2} = \frac{|g''(r)|}{2}.$$

Observe that $\alpha = 2$, which shows the iteration will converge quadratically. $\square$

In most instances this situation applies to Newton's method. Computing $g'(x)$ from 4.2 we have

$$g'(x) = 1 - \frac{f'(x)f'(x) - f(x)f''(x)}{[f'(x)]^2} = \frac{f(x)f''(x)}{[f'(x)]^2}.$$

When this is evaluated at r, we find that $g'(r) = 0$ because $f(r) = 0$, provided $f'(r) \neq 0$, and so we expect Newton's method will converge quadratically. It is possible to show that

$$\lim_{x \to r} g'(x) = \frac{1}{2}$$

when $f'(r) = 0$, so in this case Newton's method exhibits only linear convergence.

4.1.2 Descent Direction Method

In this case, merit function $z(\mathbf{x}) = f(\mathbf{x})$, that is, just the objective itself.

1. Test for convergence If the termination conditions are satisfied at x^k, then it is taken (accepted) as a "solution." In practice, this may mean satisfying the desired conditions to within some tolerance. If so, stop. Otherwise, go to step 2
2. Compute a search direction, say $\mathrm{d}^k \neq 0$. This might be a direction in which the function value is known to decrease within the feasible region
3. Compute a step length, say α^k such that
$$f\left(\mathbf{x}^k + \alpha^k \mathbf{d}^k\right) < f\left(\mathbf{x}^k\right). \tag{4.3}$$
This may necessitate a one-dimensional (or line) search
4. Define the new iterate by setting
$$\mathbf{x}^{k+1} - \mathbf{x}^k + \alpha^k \mathbf{d}^k$$
and return to step 1

There are two points, we could let $\alpha^\star = \arg\min f\left(\mathbf{x}^k + \alpha^k \mathbf{d}^k\right)$. It is also the optimized direction. And the second formula is named the convergence speed/complexity

4.1.3 Algorithm Complexity and Speeds

The intrinsic computational cost/time of an algorithm depends on

- number of decision variables n : cost of the inner product of two vectors, cost of solving system of linear equations
- number of constraints m : cost of the product of a matrix and a vector, cost of the product of two matrices

- number of nonzero data entries NNZ : sparse matrix/data representation
- the desired accuracy $0 \leq \epsilon < 1$: the cost could be propotional to $\frac{1}{\epsilon^2}, \frac{1}{\epsilon}, \log\left(\frac{1}{\epsilon}\right), \log\log\left(\frac{1}{\epsilon}\right), \ldots$
- problem difficulty or complexity measures such as the Lipschiz constant β, the condition number of a matrix, etc
- Finite versus infinite convergence. For some classes of optimization problems there are algorithms that obtain an exact solution-or detect the unboundedness-in a finite number of iterations. If we determine the exact solution, the result will not depend on ϵ
- Polynomial-time versus exponential-time. The solution time grows, in the worst-case, as a function of problem sizes (number of variables, constraints, accuracy, etc.)
- Convergence order and rate. If there is a positive number γ such that

$$\left\|\mathrm{x}^k - \mathrm{x}^*\right\| \leq \frac{O(1)}{k^\gamma}\left\|\mathrm{x}^0 - \mathrm{x}^*\right\|,$$

then $\left\{\mathbf{x}^k\right\}$ converges arithmetically to $\mathbf{x}^*$ with power γ. If there exists a number $\gamma \in [0,1)$ such that

$$\left\|\mathrm{x}^{k+1} - \mathrm{x}^*\right\| \leq \gamma\left\|\mathrm{x}^k - \mathrm{x}^*\right\| \quad \left(\Rightarrow \left\|\mathrm{x}^k - \mathrm{x}^*\right\| \leq \gamma^k\left\|\mathrm{x}^0 - \mathrm{x}^*\right\|\right),$$

then $\left\{\mathbf{x}^k\right\}$ converges geometrically or linearly to $\mathbf{x}^*$ with rate γ. If there exists a number $\gamma \in [0,1)$

$$\left\|\mathbf{x}^{k+1} - \mathbf{x}^*\right\| \leq \gamma\left\|\mathbf{x}^k - \mathbf{x}^*\right\|^2 \text{ after } \gamma\left\|\mathbf{x}^k - \mathbf{x}^*\right\| < 1$$

then $\left\{\mathbf{x}^k\right\}$ converges quadratically to $\mathbf{x}^*$ (such as $\left\{\left(\frac{1}{2}\right)^{2^k}\right\}$)

It is a bit different in the computer science complexity, we also consider the the number of nonzero data entries. And if we set the desired accuracy, the computation cost will decrease from left side. Obviously, you can get the same conclusion in the below figure 4.1.3.

Regrading the convergence order, we attach detailed explanation from page 220 Luenberger et al. (1984). Consider a sequence of real numbers $\{r_k\}_{k=0}^{\infty}$ converging to the limit r^*. We define several notions related to the speed of convergence of such a sequence.

Definition 4.5. *Let the sequence $\{r_k\}$ converge to r^*. The order of convergence of $\{r_k\}$ is defined as the supremum of the nonnegative numbers p satisfying*

$$0 \leqslant \varlimsup_{k\to\infty} \frac{|r_{k+1} - r^*|}{|r_k - r^*|^p} < \infty.$$

To ensure that the definition is applicable to any sequence, it is stated in terms of limit superior rather than just limit and $0/0$ (which occurs if $r_k = r^*$ for all k) is regarded as finite. But these technicalities are rarely necessary in actual analysis, since the sequences generated by algorithms are generally quite well behaved.

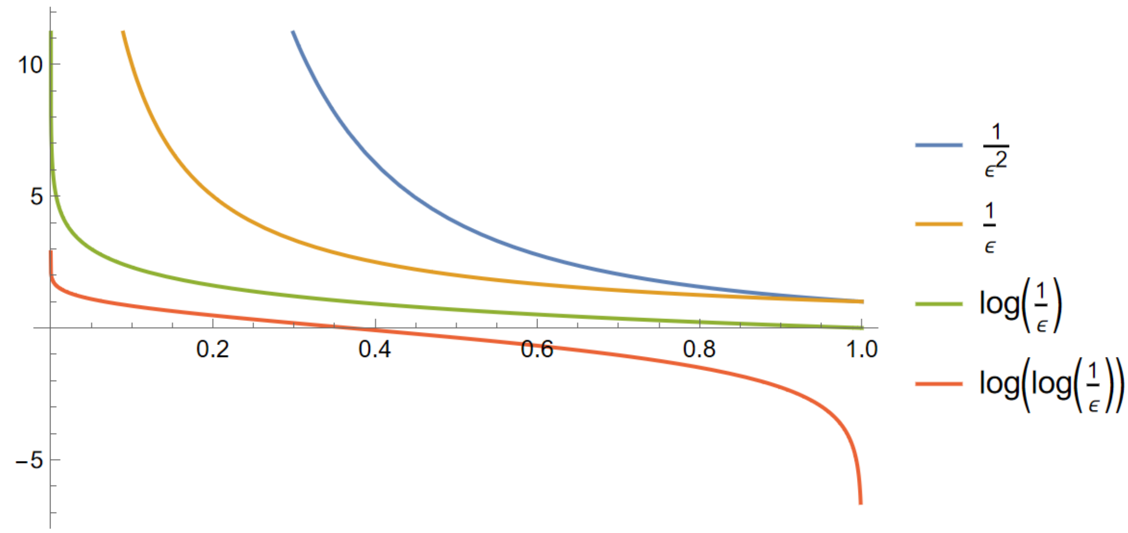

Figure 4.1: Cost vs Desired Accuracy

It should be noted that the order of convergence, as with all other notions related to speed of convergence that are introduced, is determined only by the properties of the sequence that hold as $k \to \infty$. Somewhat loosely but picturesquely, we are therefore led to refer to the tail of a sequence-that part of the sequence that is arbitrarily far out. In this language we might say that the order of convergence is a measure of how good the worst part of the tail is. Larger values of the order p imply, in a sense, faster convergence, since the distance from the limit r^* is reduced, at least in the tail, by the p th power in a single step. Indeed, if the sequence has order p and (as is the usual case) the limit

$$\beta = \lim_{k\to\infty} \frac{|r_{k+1} - r^*|}{|r_k - r^*|^p} \tag{4.4}$$

exists, then asymptotically we have

$$|r_{k+1} - r^*| = \beta \, |r_k - r^*|^p \, . \tag{4.5}$$

Example 4.6. *The sequence with $r_k = a^k$ where $0 < a < 1$ converges to zero with order unity, since $r_{k+1}/r_k = a$*

Example 4.7. *The sequence with $r_k = a^{(2^k)}$ for $0 < a < 1$ converges to zero with order two, since $r_{k+1}/r_k^2 = 1$*

Algorithm Classes

Depending on information of the problem being used to create a new iterate, we have

- Zero-order algorithms. Popular when the gradient and Hessian information are difficult to obtain/compute, e.g., no explicit function forms are given, functions are not differentiable, "black-box" optimization, etc

- First-order algorithms. Most popular now-days, suitable for large scale data optimization with low accuracy requirement, e.g., Data Science, Machine Learning, Statistical Estimation ...
- Second-order algorithms. Popular for optimization problems with high accuracy need, e.g., some scientific computing, etc

4.1.4 Zero Order Method

Golden Section Search

Assume that the one variable function $f(x)$ is **Unimodel** in interval $[ab]$, that is, for any point $x \in [a_r b_l]$ such that $a \leq a_r < b_l \leq b$, we have that $f(x) \leq \max\{f(a_r), f(b_l)\}$. How do we find x^* within an error tolerance ϵ. Answer is $|x_r - x_l| < \epsilon$

1. Initialization: let $x_l = a, x_r = b$, and choose a constant $0 < r < 0.5$
2. Let two other points $\hat{x}_l = x_l + r(x_r - x_l)$ and $\hat{x}_r = x_l + (1-r)(x_r - x_l)$, and evaluate their function values
3. Update the triple points $x_r = \hat{x}_r, \hat{x}_r = \hat{x}_l, x_l = x_l$ if $f(\hat{x}_l) < f(\hat{x}_r)$; otherwise update the triple points $x_l = \hat{x}_l, \hat{x}_l = \hat{x}_r, x_r = x_r$; and return to Step 2 .

In either cases, the length of the new interval after one golden section step is $(1-r)$. If we set $(1-2r)/(1-r) = r$, then only one point is new in each step and needs to be evaluated. This give $r = 0.382$ and the linear convergence rate is 0.618. This figure 4.1.4 from Wikipedia illustrate the idea of golden section search

4.1.5 First Order Method

Bisection/Binary Search Method

For a one variable problem, an KKT point is the root of $g(x) := f'(x) = 0$. Assume we know an interval $[ab]$ such that $a < b$, and $g(a)g(b) < 0$. Then we know there exists an x^*, $a < x^* < b$, such that $g(x^*) = 0$; that is, interval $[ab]$ contains a root of g. How do we find x within an error tolerance ϵ, that is, $|x - x^*| \leq \epsilon$

1. Initialization: let $x_l = a, x_r = b$
2. Let $x_m = (x_l + x_r)/2$, and evaluate $g(x_m)$
3. If $g(x_m) = 0$ or $x_r - x_l < \epsilon$ stop and output $x^* = x_m$. Otherwise, if $g(x_l)g(x_m) > 0$ set $x_l = x_m$; else set $x_r = x_m$; and return to Step 2

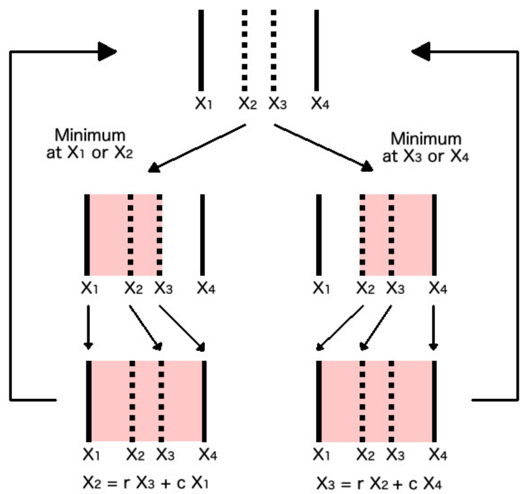

Figure 4.2: Diagram of Golden Section Search

The length of the new interval containing a root after one bisection step is 1/2 which gives the linear convergence rate is 1/2, and this establishes a linear convergence rate 0.5.

Steepest Descent Method (SDM)

Let f be a differentiable function and assume we can compute gradient (column) vector ∇f. We want to solve the unconstrained minimization problem

$$\min_{\mathbf{x}\in R^n} f(\mathbf{x}).$$

In the absence of further information, we seek a first-order KKT or stationary point of f, that is, a point x^* at which $\nabla f(\mathbf{x}^*) = 0$. Here we choose direction vector $\mathrm{d}^k = -\nabla f\left(\mathbf{x}^k\right)$ as the search direction at $\mathbf{x}^k$, which is the direction of steepest descent. The number $\alpha^k \geq 0$, called step-size, is chosen "appropriately" as

$$\alpha^k \in \arg\min f\left(\mathbf{x}^k - \alpha\nabla f\left(\mathbf{x}^k\right)\right). \tag{4.6}$$

Then the new iterate is defined as

$$\mathbf{x}^{k+1} = \mathrm{x}^k - \alpha^k \nabla f\left(\mathrm{x}^k\right) \tag{4.7}$$

In some implementations, step-size α^k is fixed through out the process - independent of iteration count k

Example 4.8. *Let $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x}$ where $Q \in R^{n\times n}$ is symmetric and positive definite. This implies that the eigenvalues of Q are all positive. The unique minimum x^* of $f(\mathbf{x})$ exists and is given by the solution of the system of linear equations*

$$\nabla f(\mathbf{x}) = Q\mathbf{x} + \mathbf{c} = \mathbf{0}$$

or equivalently

$$Q\mathbf{x} = -\mathbf{c}.$$

The iterative scheme becomes, from $\mathrm{d}^k = -\left(Q\mathbf{x}^k + \mathbf{c}\right)$

$$\mathrm{x}^{k+1} = \mathrm{x}^k + \alpha^k \mathbf{d}^k = \mathrm{x}^k - \alpha^k \left(Q\mathrm{x}^k + \mathbf{c}\right).$$

To compute the step size, α^k*, we consider*

$$\begin{aligned} & f\left(\mathbf{x}^k + \alpha\mathbf{d}^k\right) \\ =& \mathbf{c}^T\left(\mathbf{x}^k + \alpha\mathbf{d}^k\right) + \frac{1}{2}\left(\mathbf{x}^k + \alpha\mathbf{d}^k\right)^T Q\left(\mathbf{x}^k + \alpha\mathbf{d}^k\right) \\ =& \mathbf{c}^T\mathbf{x}^k + \alpha\mathbf{c}^T\mathbf{d}^k + \frac{1}{2}\left(\mathbf{x}^k\right)^{\mathrm{T}} Q\mathbf{x}^k + \alpha\left(\mathbf{x}^k\right)^{\mathrm{T}} Q\mathbf{d}^k + \frac{1}{2}\alpha^2\left(\mathbf{d}^k\right)^{\mathrm{T}} Q\mathbf{d}^k \end{aligned}$$

which is a strictly convex quadratic function of α*. Its minimizer* α^k *is the unique value of* α *where the derivative* $f'\left(\mathbf{x}^k + \alpha\mathbf{d}^k\right)$ *vanishes, i.e., where*

$$\mathbf{c}^T\mathbf{d}^k + \left(\mathbf{x}^k\right)^{\mathrm{T}} Q\mathbf{d}^k + \alpha\left(\mathbf{d}^k\right)^{\mathrm{T}} Q\mathbf{d}^k = 0.$$

Thus

$$\alpha^k = -\frac{\mathbf{c}^T\mathbf{d}^k + \left(\mathbf{x}^k\right)^{\mathrm{T}} Q\mathbf{d}^k}{\left(\mathbf{d}^k\right)^{\mathrm{T}} Q\mathbf{d}^k} = \frac{\left\|\mathbf{d}^k\right\|^2}{\left(\mathbf{d}^k\right)^{\mathrm{T}} Q\mathbf{d}^k}. \tag{4.8}$$

The recursion for the method of steepest descent now becomes

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \left(\frac{\left\|\mathbf{d}^k\right\|^2}{\left(\mathbf{d}^k\right)^{\mathrm{T}} Q\mathbf{d}^k}\right)\mathbf{d}^k. \tag{4.9}$$

Therefore, minimize a strictly convex quadratic function is equivalent to solve a system of equation with a positive definite matrix. The former may be ideal if the system only needs to be solved approximately.

Here we can determine the close form solution of α^k. We recall how to compute the optimal α^k, let $\alpha^\star = \arg\min f\left(\mathbf{x}^k + \alpha^k\mathbf{d}^k\right)$. The direct method is to take the first order derivative.

The following theorem gives some conditions under which the steepest descent method will generate a sequence of iterates that converge.

Theorem 4.9. *Let* $f: R^n \to R$ *be given. For some given point* $\mathrm{x}^0 \in R^n$*, let the level set*

$$X^0 = \left\{\mathbf{x} \in R^n : f(\mathbf{x}) \leq f\left(\mathbf{x}^0\right)\right\}$$

be bounded. Assume further that f *is continuously differentiable on the convex hull of* X^0*. Let* $\left\{\mathrm{x}^k\right\}$ *be the sequence of points generated by the steepest descent method initiated at* x^0*. Then every accumulation point of* $\left\{\mathbf{x}^k\right\}$ *is a stationary point of* f

Proof. Note that the assumptions imply the compactness of X^0. Since the iterates will all belong to X^0, the existence of at least one accumulation point of $\left\{\mathrm{x}^k\right\}$ is guaranteed by the Bolzano-Weierstrass Theorem. Let $\overline{\mathrm{x}}$ be such an accumulation point, and without losing generality, $\left\{\mathrm{x}^k\right\}$

converge to $\bar{\mathrm{x}}$. Assume $\nabla f(\overline{\mathbf{x}}) \neq 0$. Then there exists a value $\bar{\alpha} > 0$ and a $\delta > 0$ such that $f(\overline{\mathbf{x}} - \bar{\alpha}\nabla f(\overline{\mathbf{x}})) + \delta = f(\overline{\mathbf{x}})$. This means that $\overline{\mathbf{y}} := \overline{\mathbf{x}} - \bar{\alpha}\nabla f(\overline{\mathbf{x}})$ is an interior point of X^0 and

$$f(\overline{\mathbf{y}}) = f(\overline{\mathbf{x}}) - \delta$$

For an arbitrary iterate of the sequence, say x^k, the Mean-Value Theorem implies that we can write

$$f\left(\mathbf{x}^k - \bar{\alpha}\nabla f\left(\mathbf{x}^k\right)\right) = f(\overline{\mathbf{y}}) + \left(\nabla f\left(\mathbf{y}^k\right)\right)^T\left(\mathbf{x}^k - \bar{\alpha}\nabla f\left(\mathbf{x}^k\right) - \overline{\mathbf{y}}\right)$$

where $\mathbf{y}^k$ lies between $\mathbf{x}^k - \bar{\alpha}\nabla f\left(\mathbf{x}^k\right)$ and $\overline{\mathbf{y}}$. Then $\left\{\mathbf{y}^k\right\} \to \overline{\mathbf{y}}$ and $\left\{\nabla f\left(\mathbf{y}^k\right)\right\} \to \nabla f(\overline{\mathbf{y}})$ as $\left\{\mathrm{x}^k\right\} \to \bar{\mathrm{x}}$. Thus, for sufficiently large k,

$$f\left(\mathbf{x}^k - \bar{\alpha}\nabla f\left(\mathbf{x}^k\right)\right) \leq f(\overline{\mathbf{y}}) + \frac{\delta}{2} = f(\overline{\mathbf{x}}) - \frac{\delta}{2}$$

Since the sequence $\left\{f\left(\mathbf{x}^k\right)\right\}$ is monotonically decreasing and converges to $f(\overline{\mathbf{x}})$, hence

$$f(\overline{\mathbf{x}}) < f\left(\mathrm{x}^{k+1}\right) = f\left(\mathrm{x}^k - \alpha_k\nabla f\left(\mathrm{x}^k\right)\right) \leq f\left(\mathrm{x}^k - \bar{\alpha}\nabla f\left(\mathrm{x}^k\right)\right) \leq f(\overline{\mathbf{x}}) - \frac{\delta}{2} \tag{4.10}$$

which is a contradiction. Hence $\nabla f(\overline{\mathbf{x}}) = 0$. □

Remark 4.10. *According to this theorem, the steepest descent method initiated at any point of the level set X^0 will converge to a stationary point of f, which property is called global convergence.*

This proof can be viewed as a special form of Theorem 4.1: the SDM algorithm mapping is closed and the (merit) objective function is strictly decreasing if not optimal yet.

What if the SDM optimizes he β-Lipschitz function. Let $f(\mathbf{x})$ be differentiable every where and satisfy the (first-order) β-Lipschitz condition, that is, for any two points x and y

$$\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\| \leq \beta\|\mathbf{x} - \mathbf{y}\| \tag{4.11}$$

for a positive real constant β. Then, we have

Lemma 4.11. *Let f be a β-Lipschitz function. Then for any two points* x *and* y

$$f(\mathbf{x}) - f(\mathbf{y}) - \nabla f(\mathbf{y})^T(\mathbf{x} - \mathbf{y}) \leq \frac{\beta}{2}\|\mathbf{x} - \mathbf{y}\|^2. \tag{4.12}$$

At the kth step of SDM, we have

$$f(\mathbf{x}) - f\left(\mathbf{x}^k\right) \leq \nabla f\left(\mathbf{x}^k\right)^T\left(\mathbf{x} - \mathbf{x}^k\right) + \frac{\beta}{2}\left\|\mathbf{x} - \mathbf{x}^k\right\|^2.$$

The left hand strict convex quadratic function of x establishes a upper bound on the objective reduction. And then we can minimize the right hand term because of the boundness and we assume that $f(\mathbf{x}) < f\left(\mathbf{x}^k\right)$ if $\nabla f\left(\mathbf{x}^k\right) \neq 0$

Let us minimize the quadratic function

$$\mathbf{x}^{k+1} = \arg\min_{\mathbf{x}} \nabla f\left(\mathbf{x}^k\right)^T\left(\mathbf{x} - \mathbf{x}^k\right) + \frac{\beta}{2}\left\|\mathbf{x} - \mathbf{x}^k\right\|^2$$

and let the minimizer be the next iterate. Then it has a close form:

$$\frac{\partial}{\partial \mathbf{x}^k}\left(\nabla f\left(\mathbf{x}^k\right)^T\left(\mathbf{x}^{k+1}-\mathbf{x}^k\right)+\frac{\beta}{2}\left\|\mathbf{x}^{k+1}-\mathbf{x}^k\right\|^2\right)=0$$

$$\mathbf{x}^{k+1}=\mathbf{x}^k-\frac{1}{\beta}\nabla f\left(\mathbf{x}^k\right)$$

which is the SDM with the fixed step-size $\frac{1}{\beta}$. Then from 4.11

$$f\left(\mathbf{x}^{k+1}\right)-f\left(\mathbf{x}^k\right)\leq \nabla f\left(\mathbf{x}^k\right)^T\left(-\frac{1}{\beta}\nabla f\left(\mathbf{x}^k\right)\right)+\frac{\beta}{2}\|-\frac{1}{\beta}\nabla f\left(\mathbf{x}^k\right)\|^2$$

$$f\left(\mathbf{x}^{k+1}\right)-f\left(\mathbf{x}^k\right)\leq -\frac{1}{2\beta}\left\|\nabla f\left(\mathbf{x}^k\right)\right\|^2,\quad \text{or}\quad f\left(\mathbf{x}^k\right)-f\left(\mathbf{x}^{k+1}\right)\geq \frac{1}{2\beta}\left\|\nabla f\left(\mathbf{x}^k\right)\right\|^2$$

Then, after $K(\geq 1)$ steps, we must have

$$f\left(\mathbf{x}^0\right)-f\left(\mathbf{x}^K\right)\geq \frac{1}{2\beta}\sum_{k=0}^{K-1}\left\|\nabla f\left(\mathbf{x}^k\right)\right\|^2 \tag{4.13}$$

Theorem 4.12. *(Error Convergence Estimate Theorem) Let the objective function $p^*=\inf f(\mathbf{x})$ be finite and let us stop the SDM as soon as $\left\|\nabla f\left(\mathbf{x}^k\right)\right\|\leq f$ for a given tolerance $f\in(0,1)$. Then the SDM terminates in $\frac{2\beta\left(f\left(\mathbf{x}^0\right)-p^*\right)}{\epsilon^2}$ steps.*

Proof. From inequality 4.16, after $K=\frac{2\beta\left(f\left(\mathbf{x}^0\right)-p^*\right)}{\epsilon^2}$ steps

$$f\left(\mathbf{x}^0\right)-p^*\geq f\left(\mathbf{x}^0\right)-f\left(\mathbf{x}^K\right)\geq \frac{1}{2\beta}\sum_{k=0}^{K-1}\left\|\nabla f\left(\mathbf{x}^k\right)\right\|^2$$

If $\left\|\nabla f\left(\mathbf{x}^k\right)\right\|>\epsilon$ for all $k=0,\ldots,K-1$, then we have

$$f\left(\mathbf{x}^0\right)-p^*>\frac{K}{2\beta}\epsilon^2\geq f\left(\mathbf{x}^0\right)-p^*$$

which is a contradiction. □

Corollary 4.13. *If a minimizer x^* of f is attainable, then the SDM terminates in $\frac{\beta^2\left\|\mathrm{x}^0-\mathrm{x}^*\right\|^2}{\epsilon^2}$ steps. The proof is based on Lipschitz propriety with $\mathbf{x}=\mathbf{x}^0$ and $\mathbf{y}=\mathbf{x}^*$ and noting $\nabla f(\mathbf{y})=\nabla f\left(\mathbf{x}^*\right)=0$*

$$f\left(\mathbf{x}^0\right)-p^*=f\left(\mathbf{x}^0\right)-f\left(\mathbf{x}^*\right)\leq \frac{\beta}{2}\left\|\mathbf{x}^0-\mathbf{x}^*\right\|^2$$

Here we consider $f(\mathbf{x})$ being convex and differentiable everywhere and satisfying the (first-order) β-Lipschitz condition. Given the knowledge β, we again adopt the fixed step-size rule:

$$\mathrm{x}^{k+1}=\mathrm{x}^k-\frac{1}{\beta}\nabla f\left(\mathrm{x}^k\right). \tag{4.14}$$

Theorem 4.14. *For convex Lipschitz optimization the Steepest Descent Method generates a sequence of solutions such that*

$$f\left(\mathbf{x}^{k+1}\right)-f\left(\mathbf{x}^*\right)\leq \frac{\beta}{k+2}\left\|\mathbf{x}^0-\mathbf{x}^*\right\|^2 \ \textit{and}\ \min_{l=0,\ldots,k}\left\|\nabla f\left(\mathbf{x}^l\right)\right\|^2\leq \frac{4\beta^2}{(k+1)(k+2)}\left\|\mathbf{x}^0-\mathbf{x}^*\right\|^2, \tag{4.15}$$

where x^* is a minimizer of the problem.

Proof. For simplicity, we let $\delta^k = f\left(\mathbf{x}^k\right) - f\left(\mathbf{x}^*\right) (\geq 0), \mathbf{g}^k = \nabla f\left(\mathbf{x}^k\right)$, and $\Delta^k = \mathbf{x}^k - \mathbf{x}^*$ in the rest of proof. As we have proved for general Lipschitz optimization

$$\delta^{k+1} - \delta^k = f\left(\mathbf{x}^{k+1}\right) - f\left(\mathbf{x}^k\right) \leq -\frac{1}{2\beta}\left\|\mathbf{g}^k\right\|^2, \quad \text{that is} \quad \delta^k - \delta^{k+1} \geq \frac{1}{2\beta}\left\|\mathbf{g}^k\right\|^2 \tag{4.16}$$

Furthermore, from the convexity,

$$-\delta^k = f\left(\mathbf{x}^*\right) - f\left(\mathbf{x}^k\right) \geq \left(\mathbf{g}^k\right)^T\left(\mathbf{x}^* - \mathbf{x}^k\right) = -\left(\mathbf{g}^k\right)^T\Delta^k, \quad \text{that is} \quad \delta^k \leq \left(\mathbf{g}^k\right)^T\Delta^k \tag{4.17}$$

Thus, from 4.16 and 4.18

$$\begin{aligned}
\delta^{k+1} &= \delta^{k+1} - \delta^k + \delta^k \\
&\leq -\frac{1}{2\beta}\left\|\mathrm{g}^k\right\|^2 + \left(\mathrm{g}^k\right)^T\Delta^k \\
&= -\frac{\beta}{2}\left\|\mathbf{x}^{k+1} - \mathbf{x}^k\right\|^2 - \beta\left(\mathbf{x}^{k+1} - \mathrm{x}^k\right)\Delta^k, \quad (\text{ using (4)) }) \\
&= -\frac{\beta}{2}\left(\left\|\mathbf{x}^{k+1} - \mathbf{x}^k\right\|^2 + 2\left(\mathbf{x}^{k+1} - \mathbf{x}^k\right)^T\Delta^k\right) \\
&= -\frac{\beta}{2}\left(\left\|\Delta^{k+1} - \Delta^k\right\|^2 + 2\left(\Delta^{k+1} - \Delta^k\right)^T\Delta^k\right) \\
&= \frac{\beta}{2}\left(\left\|\Delta^k\right\|^2 - \left\|\Delta^{k+1}\right\|^2\right)
\end{aligned} \tag{4.18}$$

Sum up 4.18 from 1 to $k+1$, we have

$$\sum_{l=1}^{k+1}\delta^l \leq \frac{\beta}{2}\left(\left\|\Delta^0\right\|^2 - \left\|\Delta^{k+1}\right\|^2\right) \leq \frac{\beta}{2}\left\|\Delta^n\right\|^2$$

From the proof of the Corollary 4.13 of last lecture, we have $\delta^0 \leq \frac{\beta}{2}\left\|\Delta^0\right\|^2$. Thus,

$$\sum_{l=0}^{k+1}\delta^l \leq \beta\left\|\Delta^0\right\|^2,$$

and

$$\begin{aligned}
\sum_{l=0}^{k+1}\delta^l &= \sum_{l=0}^{k+1}(l+1-l)\delta^l \\
&= \sum_{l=0}^{k+1}(l+1)\delta^l - \sum_{l=0}^{k+1}l\delta^l \\
&= \sum_{l=1}^{k+2}l\delta^{l-1} - \sum_{l=1}^{k+1}l\delta^l \\
&= (k+2)\delta^{k+1} + \sum_{l=1}^{k+1}l\delta^{l-1} - \sum_{l=1}^{k+1}l\delta^l \\
&= (k+2)\delta^{k+1} + \sum_{l=1}^{k+1}l\left(\delta^{l-1} - \delta^l\right) \\
&\geq (k+2)\delta^{k+1} + \sum_{l=1}^{k+1}l\frac{1}{2\beta}\left\|\mathrm{g}^{l-1}\right\|^2
\end{aligned}$$

where the first inequality comes from 4.16. Let $\|g'\| = \min_{l=0,\ldots,k} \left\| g^l \right\|$. Then we finally have

$$(k+2)\delta^{k+1} + \frac{(k+1)(k+2)/2}{2\beta} \|g'\|^2 \leq \beta \left\|\Delta^0\right\|^2,$$

which completes the proof. □

4.1.6 Second Order Method

For functions of a single real variable x, the KKT condition is $g(x) := f'(x) = 0$. When f is twice continuously differentiable then g is once continuously differentiable, Newton's method can be a very effective way to solve such equations and hence to locate a root of g. Given a starting point x^0, Newton's method for solving the equation $g(x) = 0$ is to generate the sequence of iterates

$$x^{k+1} = x^k - \frac{g\left(x^k\right)}{g'\left(x^k\right)}. \tag{4.19}$$

The iteration is well defined provided that $g'\left(x^k\right) \neq 0$ at each step. For strictly convex function, Newton's method has a linear convergence rate and, when the point is "close" to the root, the convergence becomes quadratic, which lead to the iteration bound of $O(\log\log(1/\epsilon))$. From the zero order method to second order method. It indicates that high order information can guarantee high convergence speed.

Theorem 4.15. *(Smale (1986)) Let $g(x)$ be an analytic function. Then, if x in the domain of g satisfies*

$$\sup_{k>1} \left|\frac{g^{(k)}(x)}{k!g'(x)}\right|^{1/(k-1)} \leq (1/8) \left|\frac{g'(x)}{g(x)}\right|.$$

Then, x is an approximate root of g. In the following, for simplicity, let the root be in interval $\begin{bmatrix} 0 & R \end{bmatrix}$.

Corollary 4.16. *(Ye (2011)) Let $g(x)$ be an analytic function in R^{++} and let g be convex and monotonically decreasing. Furthermore, for $x \in R^{++}$ and $k > 1$ let*

$$\left|\frac{g^{(k)}(x)}{k!g'(x)}\right|^{1/(k-1)} \leq \frac{\alpha}{8} \cdot \frac{1}{x}$$

for some constant $\alpha > 0$. Then, if the root $\bar{x} \in [\hat{x}, (1+1/\alpha)\hat{x}] \subset R^{++}, \hat{x}$ is an approximate root of g. **Ref Chapter 4.2, 8.4-5, 9.1-7, 12.3-6**

4.2 More First-Order Optimization Algorithms

4.2.1 Double-Directions: Heavy-Ball Method

Polyak (1964) proposes this method named Heavy-Ball Method

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \frac{4}{\left(\sqrt{\lambda_n} + \sqrt{\lambda_1}\right)^2} \nabla f\left(\mathbf{x}^k\right) + \left(\frac{\sqrt{\lambda_n} - \sqrt{\lambda_1}}{\sqrt{\lambda_n} + \sqrt{\lambda_1}}\right)\left(\mathbf{x}^k - \mathbf{x}^{k-1}\right) \tag{4.20}$$

where the convergence rate can be improved to

$$\left(\frac{\sqrt{\lambda_n} - \sqrt{\lambda_1}}{\sqrt{\lambda_n} + \sqrt{\lambda_1}}\right)^2$$

This is also called the Parallel-Tangent or Conjugate Direction method, where the second direction-term in the formula is nowadays called "acceleration" or"momentum" direction. For minimizing general functions, we can let

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha^g \nabla f\left(\mathbf{x}^k\right) + \alpha^m\left(\mathbf{x}^k - \mathbf{x}^{k-1}\right) = \mathbf{x}^k + \mathbf{d}\left(\alpha^g, \alpha^m\right)$$

where the pair of step-sizes (α^g, α^m) can be chosen to

$$\min_{(\alpha^g, \alpha^d)} \nabla f\left(\mathbf{x}^k\right) \mathbf{d}\left(\alpha^g, \alpha^m\right) + \frac{1}{2} \mathbf{d}\left(\alpha^g, \alpha^m\right) \nabla^2 f\left(\mathbf{x}^k\right) \mathbf{d}\left(\alpha^g, \alpha^m\right)$$

where x^1 can be computed from the SDM step. If $\nabla f\left(\mathbf{x}^k\right) \leq 0$ and $\nabla^2 f\left(\mathbf{x}^k\right)$ is SDP matrix, the objective function is convex.

4.2.2 Accelerated Steepest Descent Method (ASDM)

Nesterov (????) propose an accelerated steepest descent method that works as follows:

$$\lambda^0 = 0, \lambda^{k+1} = \frac{1 + \sqrt{1 + 4\left(\lambda^k\right)^2}}{2}, \alpha^k = \frac{1 - \lambda^k}{\lambda^{k+1}},$$
$$\tilde{\mathbf{x}}^{k+1} = \mathbf{x}^k - \frac{1}{\beta} \nabla f\left(\mathbf{x}^k\right), \mathbf{x}^{k+1} = \left(1 - \alpha^k\right) \tilde{\mathbf{x}}^{k+1} + \alpha^k \tilde{\mathbf{x}}^k.$$

Note that $\left(\lambda^k\right)^2 = \lambda^{k+1}\left(\lambda^{k+1} - 1\right), \lambda^k > k/2$ and $\alpha^k \leq 0$.

Theorem 4.17. *(Accelerated Steepest Descent) Let $f(\mathbf{x})$ be convex and differentiable everywhere, satisfies the (first-order) β-Lipschitz condition, and admits a minimizer $\mathbf{x}^*$. Then, the method of accelerated steepest descent generates a sequence of solutions such that*

$$f\left(\tilde{\mathbf{x}}_{k+1}\right) - f\left(\mathbf{x}^*\right) \leq \frac{2\beta}{k^2}\left|\mathbf{x}^0 - \mathbf{x}^*\right|^2, \forall k \geq 1.$$

The proof we can check out in the page 269 of textbook Luenberger et al. (1984)

4.2.3 First-Order Algorithms for Conic Constrained Optimization (CCO)

Consider the conic nonlinear optimization problem:

$$\begin{array}{rl} \min & f(\mathbf{x}) \\ s.t. & \mathbf{x} \in K \end{array}$$

Nonnegative Linear Regression: given data $A \in R^{m \times n}$ and $\mathbf{b} \in R^m$

$$\begin{array}{rl} \min f(\mathbf{x}) = \frac{1}{2}\|A\mathbf{x} - \mathbf{b}\|^2 & \\ \text{s.t.} \quad \mathbf{x} \geq \mathbf{0}; & \end{array}$$

where $\nabla f(\mathbf{x}) = A^T(A\mathbf{x} - \mathbf{b})$

Semidefinite Linear Regression: given data $A_i \in S^n$ for $i = 1, \ldots, m$ and $\mathbf{b} \in R^m$

$$\begin{array}{rl} \min f(X) = \frac{1}{2}\|\mathcal{A}X - \mathbf{b}\|^2 & \\ \text{s.t.} \quad X \succeq \mathbf{0} & \end{array}$$

where $\nabla f(X) = \mathcal{A}^T(\mathcal{A}X - \mathbf{b})$.

$$\mathcal{A}X = \begin{pmatrix} A_1 \bullet X \\ \cdots \\ A_m \bullet X \end{pmatrix} \quad \text{and} \mathcal{A}^T\mathbf{y} = \sum_{i=1} y_i A_i$$

Suppose we start from a feasible solution $\mathbf{x}^0$ or X^0.

4.2.4 SDM with Feasible Region Projection

$$\begin{aligned} \hat{\mathbf{x}}^{k+1} &= \mathbf{x}^k - \frac{1}{\beta}\nabla f\left(\mathbf{x}^k\right) \\ \mathbf{x}^{k+1} &= \mathrm{Proj}_K\left(\hat{\mathbf{x}}^{k+1}\right) : \text{Solve} \quad \min_{\mathbf{x} \in K} \left\|\mathbf{x} - \hat{\mathbf{x}}^{k+1}\right\|^2 \end{aligned} \tag{4.21}$$

For examples:

- if $K = \{\mathbf{x} : \mathbf{x} \geq 0\}$, then

$$\mathbf{x}^{k+1} = \mathrm{Proj}_K\left(\hat{\mathbf{x}}^{k+1}\right) = \max\left\{\mathbf{0}, \hat{\mathbf{x}}^{k+1}\right\}$$

- If $K = \{X : X \succeq \mathbf{0}\}$, then factorize $\hat{X}^{k+1} = \sum_{j=1}^n \lambda_j \mathbf{v}_j \mathbf{v}_j^T$ and let

$$X^{k+1} = \mathrm{Proj}_K\left(\hat{X}^{k+1}\right) = \sum_{j:\lambda_j > 0} \lambda_j \mathbf{v}_j \mathbf{v}_j^T$$

The drawback is that the total eigenvalue-factorization may be costly. But the convergence rate is linear

Consider the conic nonlinear optimization problem:

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ s.t. \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \in \mathcal{K} \end{aligned}$$

The projection method becomes, starting from a feasible solution x^0 and let direction

$$\mathbf{d}^k = -\left(I - A^T \left(AA^T\right)^{-1} A\right) \nabla f\left(\mathbf{x}^k\right)$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k \mathbf{d}^k$$

we could hold above formula

$$\begin{aligned} A\left(\mathbf{x}^k + \alpha^k \mathbf{d}^k\right) &= \mathbf{b} \\ A\mathbf{d}^k &= \mathbf{0} \\ A\mathbf{d}^k + A\nabla f\left(\mathbf{x}^k\right) &= A\nabla f\left(\mathbf{x}^k\right) \\ A\mathbf{d}^k + A\nabla f\left(\mathbf{x}^k\right) &= AA^T\left(AA^T\right)^{-1} A\nabla f\left(\mathbf{x}^k\right) \\ \mathbf{d}^k + \nabla f\left(\mathbf{x}^k\right) &= A^T\left(AA^T\right)^{-1} A\nabla f\left(\mathbf{x}^k\right) \\ \mathbf{d}^k &= \left(A^T\left(AA^T\right)^{-1} A - I\right) \nabla f\left(\mathbf{x}^k\right) \end{aligned}$$

where the stepsize can be chosen from line-search or again simply let

$$\alpha^k = \frac{1}{\beta}$$

and β is the (global) Lipschitz constant. This method also convergence linearly.

- $K \subset R^n$ whose support size is no more than $d(< n) : \mathbf{x} = \mathrm{Proj}_K(\hat{\mathbf{x}})$ contains the largest d absolute entries of $\hat{\mathrm{x}}$ and set the rest of them to zeros
- $K \subset R^n_+$ and its support size is no more than $d(< n) : \mathbf{x} = \mathrm{Proj}_K(\hat{\mathbf{x}})$ contains the largest no mor than d positive entries of $\hat{\mathrm{x}}$ and set the rest of them to zeros
- $K \subset S^n$ whose rank is no more than $d(< n)$: factorize $\hat{X} = \sum_{j=1}^n \lambda_j \mathbf{v}_j \mathbf{v}_j^T$ with $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$ then $\mathrm{Proj}_K(\hat{X}) = \sum_{j=1}^d \lambda_j \mathbf{v}_j \mathbf{v}_j^T$
- $K \subset S^n_+$ whose rank is no more than $d(< n)$: factorize $\hat{X} = \sum_{j=1}^n \lambda_j \mathbf{v}_j \mathbf{v}_j^T$ with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ then $\mathrm{Proj}_K(\hat{X}) = \sum_{j=1}^d \max\{0, \lambda_j\} \mathbf{v}_j \mathbf{v}_j^T$

4.2.5 Multiplicative-Update

At the k th iterate with $\mathrm{x}^k > 0$:

$$\mathbf{x}^{k+1} = \mathbf{x}^k \exp\left(-\frac{1}{\beta} \nabla f\left(\mathbf{x}^k\right)\right)$$

Then

$$\log\left(x^{k+1}\right) = \log\left(x^{k}\right) - \frac{1}{\beta} \bigtriangledown f(x^k)$$

Note that x^{k+1} remains positive in the updating process. The classical Projected SDM update can be viewed as

$$\mathbf{x}^{k+1} = \arg\min_{\mathbf{x}\geq\mathbf{0}} \nabla f\left(\mathbf{x}^k\right)^T \mathbf{x} + \frac{\beta}{2}\left\|\mathbf{x}-\mathbf{x}^k\right\|^2.$$

Definition 4.18. *A strongly convex function is also strictly convex, but not vice versa. A differentiable function f is called strongly convex with parameter $m > 0$ if the following inequality holds for all points x, y in its domain:*

$$(\nabla f(x) - \nabla f(y))^T (x-y) \geq m\|x-y\|_2^2$$

or, more generally,

$$\langle \nabla f(x) - \nabla f(y), x-y\rangle \geq m\|x-y\|^2 \tag{4.22}$$

where $\|\cdot\|$ is any norm.

One can choose any strongly convex function $h(\cdot)$ and define

$$\mathcal{D}_h(\mathbf{x},\mathbf{y}) = h(\mathbf{x}) - h(\mathbf{y}) - \nabla h(\mathbf{y})^T(\mathbf{x}-\mathbf{y})$$

and define the update as

$$\mathbf{x}^{k+1} = \arg\min_{\mathbf{x}\geq 0} \nabla f\left(\mathbf{x}^k\right)^T \mathbf{x} + \beta\mathcal{D}_h\left(\mathbf{x},\mathbf{x}^k\right).$$

The update above is the result of choosing (negative) entropy function $h(\mathbf{x}) = \sum_j x_j \log(x_j)$. (https://math.stackexchange.com/users/474939/gerhard s) is the proof of strong convex negative entropy function.

4.2.6 Reduced Gradient Method

. Simplex method is important to solve the linear program. we can use the simplex to reduce gradient. Think this linear program

$$\begin{aligned} \min \quad & \mathbf{c}^T\mathbf{x} \\ s.t. \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{4.23}$$

where $A \in R^{m\times n}$ has a full row rank m.

Theorem 4.19. *(The Fundamental Theorem of LP in Algebraic form) Given (LP) and (LD) where A has full row rank m,*

- *if there is a feasible solution, there is a basic feasible solution (Carathéodory's theorem)*

- *if there is an optimal solution, there is an optimal basic solution*

Main idea of procedure

1. Initialization Start at a BSF or corner point of the feasible polyhedron
2. Test for Optimality. Compute the reduced gradient vector at the corner. If no descent and feasible direction can be found, stop and claim optimality at the current corner point; otherwise, select a new corner point and go to Step 2

The reduced gradient is equivalent with reduced cost in Bertsimas and Tsitsiklis (1997)

Suppose the basis of a basic feasible solution is A_B and the rest is A_N. One can transform the equality constraint to

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ A_B^{-1}A\mathbf{x} &= A_B^{-1}\mathbf{b} \\ A_B^{-1}\left(A_B\mathbf{x}_B + A_N\mathbf{x}_N\right) &= A_B^{-1}\mathbf{b} \\ \mathbf{x}_B &= A_B^{-1}\mathbf{b} - A_B^{-1}A_N\mathbf{x}_N \end{aligned} \tag{4.24}$$

That is, we express $\mathbf{x}_B$ in terms of $\mathbf{x}_N$, the non-basic variables are are active for constraints $\mathbf{x} \geq \mathbf{0}$. Then the objective function equivalently becomes

$$\begin{aligned} \mathbf{c}^T\mathbf{x} &= \mathbf{c}_B^T\mathbf{x}_B + \mathbf{c}_N^T\mathbf{x}_N \\ &= \mathbf{c}_B^TA_B^{-1}\mathbf{b} - \mathbf{c}_B^TA_B^{-1}A_N\mathbf{x}_N + \mathbf{c}_N^T\mathbf{x}_N \\ &= \mathbf{c}_B^TA_B^{-1}\mathbf{b} + \left(\mathbf{c}_N^T - \mathbf{c}_B^TA_B^{-1}A_N\right)\mathbf{x}_N \end{aligned}$$

Vector $\mathbf{r}^T = \mathbf{c}^T - \mathbf{c}_B^TA_B^{-1}A$ is called the Reduced Gradient/Cost Vector where $\mathbf{r}_B = \mathbf{0}$ always.

Theorem 4.20. *If Reduced Gradient Vector $\mathbf{r}^T = \mathbf{c}^T - \mathbf{c}_B^TA_B^{-1}A \geq \mathbf{0}$, then the BFS is optimal.*

Proof. Let $\mathbf{y}^T = \mathbf{c}_B^TA_B^{-1}$ (called Shadow Price Vector), then $\mathbf{y}$ is a dual feasible solution $\left(\mathbf{r} = \mathbf{c} - A^T\mathbf{y} \geq \mathbf{0}\right)$ and $\mathbf{c}^T\mathbf{x} = \mathbf{c}_B^T\mathbf{x}_B = \mathbf{c}_B^TA_B^{-1}\mathbf{b} = \mathbf{y}^T\mathbf{b}$, that is, the duality gap is zero □

And the simplex procedure shows below

1. Initialize Start a BFS with basic index set B and let N denote the complementary index set
2. Test for Optimality: Compute the Reduced Gradient Vector r at the current BFS and let

$$r_e = \min_{j \in N}\{r_j\}.$$

If $r_e \geq 0$, stop - the current BFS is optimal

3. Determine the Replacement: Increase x_e while keep all other non-basic variables at the zero value (inactive) and maintain the equality constraints:

$$\mathbf{x}_B = A_B^{-1}\mathbf{b} - A_B^{-1}A_{.e}x_e(\geq \mathbf{0}).$$

   If x_e can be increased to ∞, stop - the problem is unbounded below. Otherwise, let the basic variable x_0 be the one first becoming 0

4. Update basis: update B with x_o being replaced by x_e, and return to Step 2

4.2.7 Frank-Wolf Algorithm

$$\begin{aligned} \min \quad & f(\mathrm{x}) \\ s.t. \quad & \mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

where $A \in R^{m\times n}$ has a full row rank m. Start with a feasible solution x^0, and at the kth iterate do:

1. Solve the LP problem

$$\min \nabla f\left(\mathbf{x}^k\right)^T \mathbf{x} \text{ s.t. } A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$$

   and let $\tilde{\mathbf{x}}^{k+1}$ be an optimal solution

2. Choose a step-size $0 < \alpha^k \leq 1$ and let

$$\mathrm{x}^{k+1} = \mathrm{x}^k + \alpha^k\left(\tilde{\mathrm{x}}^{k+1} - \mathrm{x}^k\right).$$

This is also called sequential linear programming (SLP) method.

We summary the first-order method

- Good global convergence property (e.g. starting from any (feasible) solution under mild technical assumption...)
- Simple to implement and the computation cost is mainly compute the numerical gradient
- Maybe difficult to decide step-size: simple back-track is popular in practice
- The convergence speed can be slow: not suitable for high accuracy computation, certain accelerations available
- Can only guarantee converging to a first-order KKT solution

Ref Chapter 14.1-6

4.3 Dual/Lagrangian Methods for Constrained Optimization

4.3.1 Local Lagrangian Method

We consider the nonlinear program

$$\begin{aligned} f^* := \min \quad & f(\mathbf{x}) \\ s.t. \quad & \mathbf{h}(\mathbf{x}) = \mathbf{0} \\ & \mathbf{x} \in X \end{aligned} \tag{4.25}$$

Recall that the Lagrangian function:

$$L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - \mathbf{y}^T \mathbf{h}(\mathbf{x}).$$

and the dual function:

$$\phi(\mathbf{y}) = \min_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{y})$$

and the dual problem

$$f^* \geq \phi^* := \max \phi(\mathbf{y}).$$

In many cases, we can find $\mathbf{y}^*$ of dual problem. Suppose x^* is a local minimizer, and consider the localized (convex) problem and the last constraint means the x belongs to $x^\star$ neighborhood

$$\begin{aligned} f\left(\mathbf{x}^*\right) := \min \quad & f(\mathbf{x}) \\ s.t. \quad & \mathbf{h}(\mathbf{x}) = \mathbf{0} \\ & \mathbf{x} \in X \\ & \left\|\mathbf{x} - \mathbf{x}^*\right\|^2 \leq \epsilon \end{aligned} \tag{4.26}$$

Then, the localized Lagrangian function:

$$L_{\mathbf{x}^*}(\mathbf{x}, \mathbf{y}, \mu) = f(\mathbf{x}) - \mathbf{y}^T \mathbf{h}(\mathbf{x}) - \mu\left(\left\|\mathbf{x} - \mathbf{x}^*\right\|^2 - \epsilon\right) \quad \text{s.t.} \mu \leq 0$$

and the localized dual function:

$$\phi_{\mathbf{x}^*}(\mathbf{y}, \mu) = \min_{\mathbf{x} \in X, \|\mathbf{x} - \mathbf{x}^*\|^2 \leq \epsilon} L_{\mathbf{x}^*}(\mathbf{x}, \mathbf{y}, \mu)$$

and the localized dual problem

$$\max \phi(\mathbf{y}, \mu)$$

Under certain constraint qualification and local convexity conditions, we must have $f\left(\mathrm{x}^*\right) = \phi\left(\mathbf{y}^*, \mu^* = 0\right)$ where the localization constraint becomes inactive.

Let x(y) be a minimizer and minimizer corresponds to the multiplier y. Then

$$\phi(\mathbf{y}) = f(\mathbf{x}(\mathbf{y})) - \mathbf{y}^T \mathbf{h}(\mathbf{x}(\mathbf{y}))$$

Thus

$$\begin{aligned}\nabla\phi(\mathbf{y}) &= \nabla f(\mathbf{x}(\mathbf{y}))^T\nabla\mathbf{x}(\mathbf{y}) - \mathbf{y}^T\nabla\mathbf{h}(\mathbf{x}(\mathbf{y}))\nabla\mathbf{x}(\mathbf{y}) - \mathbf{h}(\mathbf{x}(\mathbf{y})) \\ &= \left(\nabla f(\mathbf{x}(\mathbf{y}))^T - \mathbf{y}^T\nabla\mathbf{h}(\mathbf{x}(\mathbf{y}))\right)\nabla\mathbf{x}(\mathbf{y}) - \mathbf{h}(\mathbf{x}(\mathbf{y})) \\ &= -\mathbf{h}(\mathbf{x}(\mathbf{y}))\end{aligned} \tag{4.27}$$

According to the definition$\mathbf{x}(\mathbf{y})$, we vanish the right-hand first term.

By the definition, we have

$$\nabla f(\mathbf{x}(\mathbf{y}))^T - \mathbf{y}^T\nabla\mathbf{h}(\mathbf{x}(\mathbf{y})) = 0$$

And the Hessian is the derivative of gradient

$$\nabla^2\phi(\mathbf{y}) = -\nabla\mathbf{h}(\mathbf{x}(\mathbf{y}))\nabla\mathbf{x}(\mathbf{y})$$

and differentiate this with respect to $\mathbf{y}$

$$\begin{aligned}\nabla^2 f(\mathbf{x}(\mathbf{y}))^T\nabla\mathbf{x}(\mathbf{y}) - \nabla\mathbf{h}(\mathbf{x}(\mathbf{y}))^T - \mathbf{y}^T\nabla^2\mathbf{h}(\mathbf{x}(\mathbf{y}))\nabla\mathbf{x}(\mathbf{y}) = 0 \\ \nabla\mathbf{x}(\mathbf{y} = \left(\nabla^2_{\mathbf{x}}L(\mathbf{x}(\mathbf{y}),\mathbf{y})\right)^{-1}\nabla\mathbf{h}(\mathbf{x}(\mathbf{y}))^T\end{aligned}$$

Similarly, we can substitute above Hessian and derive

$$\nabla^2\phi(\mathbf{y}) = -\nabla\mathbf{h}(\mathbf{x}(\mathbf{y}))\left(\nabla^2_{\mathbf{x}}L(\mathbf{x}(\mathbf{y}),\mathbf{y})\right)^{-1}\nabla\mathbf{h}(\mathbf{x}(\mathbf{y}))^T \tag{4.28}$$

where $\nabla^2_{\mathbf{x}}L(\mathbf{x}(\mathbf{y}),\mathbf{y})$ is the Hessian of the Lagrangian function that is assumed to be positive definite at any (local) minimizer.

Lemma 4.21. *the dual function ϕ has gradient*

$$\nabla\phi(\mathbf{y}) = -\mathbf{h}(\mathbf{x}(\mathbf{y}))$$

Lemma 4.22. *The Hessian of the dual function is*

$$\nabla^2\phi(\mathbf{y}) = -\nabla\mathbf{h}(\mathbf{x}(\mathbf{y}))\left(\nabla^2_{\mathbf{x}}L(\mathbf{x}(\mathbf{y}),\mathbf{y})\right)^{-1}\nabla\mathbf{h}(\mathbf{x}(\mathbf{y}))^T$$

There two lemmas indicate the localization primal and dual problem has the local solution $\mathbf{x}^\star$ and $\mathbf{y}^\star$ respectively

Example 4.23. *we have toy example*

$$\begin{aligned}\min \quad & (x_1-1)^2 + (x_2-1)^2 \\ s.t. \quad & x_1 + 2x_2 - 1 = 0 \\ & 2x_1 + x_2 - 1 = 0\end{aligned}$$

Lagrangian formula

$$L(\mathbf{x},\mathbf{y}) = (x_1-1)^2 + (x_2-1)^2 - y_1(x_1+2x_2-1) - y_2(2x_1+x_2-1)$$

x_1 and x_2 have the closed form by differentiating with respect to x_1, x_2

$$x_1 = 0.5y_1 + y_2 + 1, \quad x_2 = y_1 + 0.5y_2 + 1$$

substitute to the Lagrangian formula

$$\phi(\mathbf{y}) = -1.25y_1^2 - 1.25y_2^2 - 2y_1y_2 - 2y_1 - 2y_2$$

The gradient of dual function

$$\nabla\phi(\mathbf{y}) = \begin{pmatrix} 2.5y_1 + 2y_2 + 2 \\ 2y_1 + 2.5y_2 1 + 2 \end{pmatrix}$$

The Hessian of the dual function is

$$\nabla^2\phi(\mathbf{y}) = -\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}^{-1}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}^T = -\begin{pmatrix} 2.5 & 2 \\ 2 & 2.5 \end{pmatrix}$$

4.3.2 Augmented Lagrangian Function (ALF)

The augmented Lagrangian function is an efficient method for local optimization problems. And this method combines the primal-dual method and penalty function. It will discard each drawback of the method. Let consider this toy example

Example 4.24. *(Fisher Example)*

$$\begin{aligned} \min \quad & 5\log(2x_1 + x_2) + 8\log(3x_3 + x_4) \\ s.t. \quad & x_1 + x_3 = 1 \\ & x_2 + x_4 = 1 \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

The Lagrangian function

$$L(\mathbf{x}, \mathbf{y}) = 5\log(2x_1 + x_2) + 8\log(3x_3 + x_4) - y_1(x_1 + x_3 - 1) - y_2(x_2 + x_4 - 1)$$

Start from $\mathbf{y}^0 > 0$*, at the kth step, compute* x^{k+1} *from*

$$\mathbf{x}^{k+1} = \arg\max_{\mathbf{x}\geq 0} L\left(\mathbf{x}, \mathbf{y}^k\right),$$

then let

$$\mathbf{y}^{k+1} = \mathbf{y}^k + \frac{1}{\beta}\left(A\mathbf{x}^{k+1} - \mathbf{b}\right)$$

In both theory and practice, we actually consider an augmented Lagrangian function (ALF)

$$L_a(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - \mathbf{y}^T\mathbf{h}(\mathbf{x}) + \frac{\beta}{2}\|\mathbf{h}(\mathbf{x})\|^2,$$

which corresponds to an equivalent problem above:

$$\begin{aligned} f^* := \min \quad & f(\mathbf{x}) + \frac{\beta}{2}\|\mathbf{h}(\mathbf{x})\|^2 \\ s.t. \quad & \mathbf{h}(\mathbf{x}) = \mathbf{0} \\ & \mathbf{x} \in X \end{aligned} \tag{4.29}$$

Note that, although at feasibility the additional square term in objective is redundant, it helps to improve strict convexity of the Lagrangian function. The Fisher example:

$$L_a(\mathbf{x}, \mathbf{y}) = 5\log(2x_1 + x_2) + 8\log(3x_3 + x_4) - y_1(x_1 + x_3 - 1) - y_2(x_2 + x_4 - 1)$$
$$+ \frac{\beta}{2}\left((x_1 + x_3 - 1)^2 + (x_2 + x_4 - 1)^2\right)$$

Now the dual function:

$$\phi_a(\mathbf{y}) = \min_{\mathbf{x}\in X} L_a(\mathbf{x}, \mathbf{y})$$

and the dual problem

$$f^* \geq \phi_a^* := \max \quad \phi_a(\mathbf{y}).$$

Note that the dual function approximately satisfies $\frac{1}{\beta}$-Lipschitz condition (see Chapter 14 of Luenberger et al. (1984)). For the convex optimization case, say $\mathbf{h}(\mathbf{x}) = A\mathbf{x} - \mathbf{b}$, we have

$$\nabla^2 L_a(\mathbf{x}, \mathbf{y}) = \nabla^2 f(\mathbf{x}) + \beta\left(A^T A\right).$$

The augmented Lagrangian method (ALM) is: Start from any $\left(\mathbf{x}^0 \in X, \mathbf{y}^0\right)$, we compute a new iterate pair

$$\mathbf{x}^{k+1} = \arg\min_{\mathbf{x}\in X} L_a\left(\mathbf{x}, \mathbf{y}^k\right), \text{ and } \mathbf{y}^{k+1} = \mathbf{y}^k - \beta\mathbf{h}\left(\mathbf{x}^{k+1}\right)$$

The calculation of $\mathbf{x}$ is used to compute the gradient vector of $\phi_a(\mathbf{y})$, which is a steepest ascent direction. The method converges just like the SDM, because the dual function satisfies $\frac{1}{\beta}$-Lipschitz condition. Other SDM strategies may be adapted to update y (the BB, ASDM, Conjugate, Quasi-Newton ...).

Consider the convex optimization case $\mathbf{h}(\mathbf{x}) = A\mathbf{x} - \mathbf{b}$. Since x^{k+1} makes KKT condition and take the derivative:

$$\begin{aligned}
0 &= \nabla f\left(\mathbf{x}^{k+1}\right) - A^T\mathbf{y}^k + \beta A^T\left(A\mathbf{x}^{k+1} - \mathbf{b}\right) \\
&= \nabla f\left(\mathbf{x}^{k+1}\right) - A^T\left(\mathbf{y}^k - \beta\left(A\mathbf{x}^{k+1} - \mathbf{b}\right)\right) \\
&= \nabla f\left(\mathbf{x}^{k+1}\right) - A^T\mathbf{y}^{k+1}
\end{aligned}$$

we only need to be concerned about whether or not $\left\|A\mathbf{x}^k - \mathbf{b}\right\|$ converges to zero and how fast it converges. First, from the convexity of $f(\mathbf{x})$, we have

$$\begin{aligned}
\mathbf{0} &\leq \left(\nabla f\left(\mathbf{x}^{k+1}\right) - \nabla f\left(\mathbf{x}^k\right)\right)^T\left(\mathbf{x}^{k+1} - \mathbf{x}^k\right) \\
&= \left(-A^T\mathbf{y}^{k+1} + A^T\mathbf{y}^k\right)^T\left(\mathbf{x}^{k+1} - \mathbf{x}^k\right) \\
&= \left(\mathbf{y}^{k+1} - \mathbf{y}^k\right)^T\left(A\mathbf{x}^{k+1} - A\mathbf{x}^k\right) \\
&= -\beta\left(A\mathbf{x}^{k+1} - \mathbf{b}\right)\left(A\mathbf{x}^{k+1} - \mathbf{b} - \left(A\mathbf{x}^k - \mathbf{b}\right)\right)
\end{aligned}$$

which implies that $\left\|A\mathbf{x}^{k+1} - \mathbf{b}\right\| \leq \left\|A\mathbf{x}^k - \mathbf{b}\right\|$, that is, the error is non-increasing. Note that it is easy to get $\mathbf{0} \leq \left(\nabla f\left(\mathbf{x}^{k+1}\right) - \nabla f\left(\mathbf{x}^k\right)\right)^T\left(\mathbf{x}^{k+1} - \mathbf{x}^k\right)$ from convex function first order condition.

Proof. From convex function first order condition

$$\begin{aligned}
-\left(f(\mathbf{y}) - (\mathbf{x})\right) &\leq -\nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \\
f(\mathbf{y}) - (\mathbf{x}) &\leq \nabla f(\mathbf{y})^T(\mathbf{y} - \mathbf{x})
\end{aligned}$$

Then

$$\mathbf{0} \leq (\nabla f(\mathbf{y}) - \nabla f(\mathbf{x}))^T (\mathbf{y} - \mathbf{x})$$

□

Again, from the convexity, we have

$$\begin{aligned}
\mathbf{0} &\leq \left(\nabla f\left(\mathbf{x}^{k+1}\right) - \nabla f\left(\mathbf{x}^*\right)\right)^T \left(\mathbf{x}^{k+1} - \mathbf{x}^*\right) \\
&= \left(A^T\mathbf{y}^{k+1} - A^T\mathbf{y}^*\right)^T \left(\mathbf{x}^{k+1} - \mathbf{x}^*\right) \\
&= \left(\mathbf{y}^{k+1} - \mathbf{y}^*\right)^T \left(A\mathbf{x}^{k+1} - A\mathbf{x}^*\right) \\
&= \left(\mathbf{y}^{k+1} - \mathbf{y}^*\right)^T \left(A\mathbf{x}^{k+1} - \mathbf{b}\right) \\
&= \frac{1}{\beta}\left(\mathbf{y}^{k+1} - \mathbf{y}^*\right)^T \left(\mathbf{y}^{k} - \mathbf{y}^{k+1}\right)
\end{aligned}$$

Thus, from the positivity of the cross product, we have

$$\begin{aligned}
\left\|\mathbf{y}^k - \mathbf{y}^*\right\|^2 &= \left\|\mathbf{y}^k - \mathbf{y}^{k+1} + \mathbf{y}^{k+1} - \mathbf{y}^*\right\|^2 \\
&\geq \left\|\mathbf{y}^k - \mathbf{y}^{k+1}\right\|^2 + \left\|\mathbf{y}^{k+1} - \mathbf{y}^*\right\|^2 \\
&= \beta\left\|A\mathbf{x}^{k+1} - \mathbf{b}\right\|^2 + \left\|\mathbf{y}^{k+1} - \mathbf{y}^*\right\|^2
\end{aligned}$$

Sum up from 0 to k of the inequality we have

$$\begin{aligned}
\left\|\mathbf{y}^0 - \mathbf{y}^*\right\|^2 &\geq \left\|\mathbf{y}^{k+1} - \mathbf{y}^*\right\|^2 + \beta\sum_{l=0}^{k}\left\|A\mathbf{x}^{l+1} - \mathbf{b}\right\|^2 \\
&\geq \beta\sum_{l=0}^{k}\left\|A\mathbf{x}^{l+1} - \mathbf{b}\right\|^2 \\
&\geq (k+1)\beta\left\|A\mathbf{x}^{k+1} - \mathbf{b}\right\|^2
\end{aligned}$$

4.4 Augmented Direction Method with Multipliers (ADMM)

4.4.1 Two-Blocks ADMM

For the ADMM method, we consider two block structured problem

$$\min f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2) \quad \text{s.t.} \quad A_1\mathbf{x}_1 + A_2\mathbf{x}_2 = \mathbf{b}, \mathbf{x}_1 \in X_1, \mathbf{x}_2 \in X_2.$$

Consider

$$L(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2) - \mathbf{y}^T(A_1\mathbf{x}_1 + A_2\mathbf{x}_2 - \mathbf{b}) + \frac{\beta}{2}\|A_1\mathbf{x}_1 + A_2\mathbf{x}_2 - \mathbf{b}\|^2.$$

Then, for any given $\left(\mathrm{x}_1^k, \mathrm{x}_2^k, \mathrm{y}^k\right)$, we compute a new iterate

$$\begin{aligned}
\mathbf{x}_1^{k+1} &= \arg\min_{\mathbf{x}_1 \in X_1} L\left(\mathbf{x}_1, \mathbf{x}_2^k, \mathbf{y}^k\right) \\
\mathbf{x}_2^{k+1} &= \arg\min_{\mathbf{x}_2 \in X_2} L\left(\mathbf{x}_1^{k+1}, \mathbf{x}_2, \mathbf{y}^k\right) \\
\mathbf{y}^{k+1} &= \mathbf{y}^k \quad \beta\left(A_1\mathbf{x}_1^{k+1} + A_2\mathbf{x}_2^{k+1} \quad \mathbf{b}\right).
\end{aligned}$$

Again, we can prove that the iterates converge with the same speed. The ADMM method resembles the Block Coordinate Descent (BCD) Method

Consider the standard-form LP

$$
\begin{array}{ll}
\min & \mathbf{c}^T \mathbf{x} \\
s.t. & A\mathrm{x} = \mathbf{b} \\
 & \mathrm{x} \geq 0
\end{array}
\qquad
\begin{array}{ll}
\min & \mathbf{c}^T \mathbf{x}_1 \\
s.t. & A\mathbf{x}_1 = \mathbf{b} \\
 & \mathrm{x}_1 - \mathrm{x}_2 = \mathbf{0} \\
 & \mathrm{x}_2 \geq \mathbf{0}
\end{array}
$$

$$
L\left(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}\right) = \mathbf{c}^T \mathbf{x}_1 - \mathbf{y}^T\left(A\mathbf{x}_1 - \mathbf{b}\right) - \mathbf{s}^T\left(\mathbf{x}_1 - \mathbf{x}_2\right) + \frac{\beta}{2}\left(\left\|A\mathbf{x}_1 - \mathbf{b}\right\|^2 + \left\|\mathbf{x}_1 - \mathbf{x}_2\right\|^2\right)
$$

where y and s are the multiplier vectors of first and second equality constraints in the reformulation. The advantage of such splitting reformulation is that the update of either x_1 or x_2 has a simple close form solution.

Consider the dual LP

$$
\begin{array}{ll}
\max_{(\mathbf{y}, \mathbf{s})} & \mathbf{b}^T \mathbf{y} \\
\text{s.t.} & A^T \mathbf{y} + \mathbf{s} = \mathbf{c}, \mathbf{s} \geq \mathbf{0}
\end{array}
$$

The augmented Lagrangian function would be

$$
L(\mathbf{y}, \mathbf{s}, \mathbf{x}) = -\mathbf{b}^T \mathbf{y} - \mathbf{x}^T\left(A^T \mathbf{y} + \mathbf{s} - \mathbf{c}\right) + \frac{\beta}{2}\left\|A^T \mathbf{y} + \mathbf{s} - \mathbf{c}\right\|^3,
$$

where β is a positive parameter, and x is the multiplier vector.

The ADMM for the dual is straightforward: starting from any y^0, $\mathrm{s}^0 \geq 0$, and multiplier x^0, - Update variable y:

$$
\mathbf{y}^{k+1} = \arg\min_{\mathbf{y}} L\left(\mathbf{y}, \mathbf{s}^k, \mathbf{x}^k\right)
$$

- Update slack variable s:

$$
\mathbf{s}^{k+1} = \arg\min_{\mathbf{s} \geq 0} L\left(\mathbf{y}^{k+1}, \mathbf{s}, \mathbf{x}^k\right)
$$

- Update multipliers x :

$$
\mathbf{x}^{k+1} = \mathbf{x}^k - \beta\left(A^T \mathbf{y}^{k+1} + \mathbf{s}^{k+1} - \mathbf{c}\right).
$$

Note that the updates of y is a least-squares problem with constant matrix, and the update of s has a simple close form. (Also note that x would be non-positive at the end, since we changed maximization to minimization of the dual.) To split $\mathbf{y}$ into multi blocks and update cyclically in random order? The answer is it could work well.

we provide the toy example

$$\begin{aligned}
\min \quad & 5\log(2x_1+x_2)+8\log(3x_3+x_3) \\
s.t. \quad & x_1+x_3-1=0 \\
& x_2+x_4-1=0 \\
& 2x_1+x_2-u_1=0 \\
& 3x_3+x_4-u_2=0 \\
& \mathrm{x}-\mathrm{s}=\mathbf{0} \\
& \mathrm{s}\geq\mathbf{0}
\end{aligned}$$

Based on the Lagrangian method, let the first block primal variables be x and the second be $(\mathbf{u},\mathrm{s})$. Then start from y^0 repeat the ADMM steps. Note that all primal variables have close-form solutions.

4.4.2 Three-Blocks ADMM

What about ADMM for

$$\min f_1(\mathbf{x}_1)+f_2(\mathbf{x}_2)+f_3(\mathbf{x}_3) \quad \text{s.t.} \quad A_1\mathbf{x}_1+A_2\mathbf{x}_2+A_3\mathbf{x}_3=\mathbf{b},$$

where the Lagrangian function

$$\begin{aligned}
L(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{y}) = & f_1(\mathbf{x}_1)+f_2(\mathbf{x}_2)+f_3(\mathbf{x}_3)-\mathbf{y}^T(A_1\mathbf{x}_1+A_2\mathbf{x}_2+A_3\mathbf{x}_3-\mathbf{b}) \\
& +\frac{\beta}{2}\|A_1\mathbf{x}_1+A_2\mathbf{x}_2+A_3\mathbf{x}_3-\mathbf{b}\|^2
\end{aligned}$$

Then, for any given $(\mathbf{x}_1^k,\mathbf{x}_2^k,\mathbf{x}_3^k,\mathbf{y}^k)$, we compute a new iterate

$$\begin{aligned}
\mathbf{x}_1^{k+1} &= \arg\min_{\mathbf{x}_1} L(\mathbf{x}_1,\mathbf{x}_2^k,\mathbf{x}_3^k,\mathbf{y}^k) \\
\mathbf{x}_2^{k+1} &= \arg\min_{\mathbf{x}_2} L(\mathbf{x}_1^{k+1},\mathbf{x}_2,\mathbf{x}_3^k,\mathbf{y}^k) \\
\mathbf{x}_3^{k+1} &= \arg\min_{\mathbf{x}_3} L(\mathbf{x}_1^{k+1},\mathbf{x}_2^{k+1},\mathbf{x}_3,\mathbf{y}^k) \\
\mathbf{y}^{k+1} &= \mathbf{y}^k-\beta(A_1\mathbf{x}^{k+1}+A_2\mathbf{x}_2^{k+1}+A_3\mathbf{x}_3^{k+1}-\mathbf{b})
\end{aligned}$$

Not easy to analyze the convergence: the operator theory for the ADMM cannot be directly extended to the ADMM with three blocks, since the proof for two blocks breaks down for three blocks. Existing results for convergence:

- Strong convexity; plus carefully select β in a specific range.
- Other restricted conditions on the problem, and take a sufficiently smaller step-size factor $1>\gamma>0$ in dual update.

$$\mathbf{y}^{k+1}=\mathbf{y}^k-\gamma\beta(A_1\mathbf{x}_1^{k+1}+A_2\mathbf{x}_2^{k+1}+A_3\mathbf{x}_3^{k+1}-\mathbf{b}).$$

- Various post correction steps, which are costly. But, these did not answer the open question whether or not the direct extension of multi-block ADMM converges under the original simple convexity assumption.

Theorem 4.25. *There existing an example where the direct extension of ADMM of three blocks is not necessarily convergent for any choice of β. Moreover, for any randomly generated initial point, ADMM diverges with probability one.*

The problem with unique solution $\mathrm{x}^* = 0$:

$$\min 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 \quad \text{s.t.} \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0},$$

Does the smaller step-size ($1 > \gamma > 0$) dual update work? Answer: it remains divergent when solving

$$\min 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 \quad \text{s.t.} \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1+\gamma \\ 1 & 1+\gamma & 1+\gamma \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0},$$

The ADMM with $\beta = 1$ is a linear matrix mapping

$$\begin{pmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 4 & 6 & 0 & 0 & 0 & 0 \\ 5 & 7 & 9 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 2 & 0 & 0 & 1 \end{pmatrix} (\mathrm{x}^{k+1}\mathbf{y}^{k+1}) = \begin{pmatrix} 0 & -4 & -5 & 1 & 1 & 1 \\ 0 & 0 & -7 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} (x^k\mathbf{y}^k)$$

we can give some comments to matrix M, we can simplify the good structure. Even if the we iterate by many steps, the matrix M eigenvalue still is very large. We need try to reduce the eigenvalue during iterating. which can be reduced to

$$\begin{pmatrix} x_2^{k+1} \\ x_3^{k+1} \\ \mathbf{y}^{k+1} \end{pmatrix} = M \begin{pmatrix} x_2^k \\ x_3^k \\ \mathbf{y}^k \end{pmatrix}$$

where and it is the linear mapping

$$M = \frac{1}{162} \begin{pmatrix} 144 & -9 & -9 & -9 & 18 \\ 8 & 157 & -5 & 13 & -8 \\ 64 & 122 & 122 & -58 & -64 \\ 56 & -35 & -35 & 91 & -56 \\ -88 & -26 & -26 & -62 & 88 \end{pmatrix}$$

But the spectral radius of the matrix, $\rho(M) = 1.0087 > 1$, which implies that the mapping is not a contraction.

Definition 4.26. *Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of a matrix $A \in \mathbf{C}^{n \times n}$. The spectral radius of A is defined as*

$$\rho(A) = \max\{|\lambda_1|, \ldots, |\lambda_n|\}. \tag{4.30}$$

the trick we use is to take the average. $\rho(M_i) > 1, \forall i$, then $\rho(\frac{\sum_i^n M_i}{n}) < 1$. In the sense, spectral radius equals $\rho(\mathbb{E}[M_i]) < 1$

In general, consider a convex optimization problem

$$\begin{aligned} &\min_{\mathbf{x} \in R^N} f_1(\mathbf{x}_1) + \ldots + f_n(\mathbf{x}_n) \\ &\text{s.t. } A\mathbf{x} := A_1\mathbf{x}_1 + \cdots + A_n\mathbf{x}_n = \mathbf{b} \\ &\mathbf{x}_i \in \mathcal{X}_i \subset R^{d_i}, i = 1, \ldots, n \end{aligned}$$

$$L(\mathbf{x}_1, \ldots, \mathbf{x}_n; \mathbf{y}) = \sum_i f_i(x_i) - \mathbf{y}^T\left(\sum_i A_i\mathbf{x}_i - \mathbf{b}\right) + \frac{\beta}{2}\left\|\sum_i A_i\mathbf{x}_i - \mathbf{b}\right\|^2$$

The direct Cyclic Extension Multi-block ADMM:

$$\begin{aligned} \mathbf{x}_1 &\longleftarrow \arg\min_{\mathbf{x}_1 \in \mathcal{X}_1} L(\mathbf{x}_1, \ldots, \mathbf{x}_n; \mathbf{y}) \\ &\vdots \\ \mathbf{x}_n &\longleftarrow \arg\min_{\mathbf{x}_n \in \mathcal{X}_n} L(\mathbf{x}_1, \ldots, \mathbf{x}_n; \mathbf{y}) \\ \mathbf{y} &\longleftarrow \mathbf{y} - \beta(A\mathbf{x} - \mathbf{h}) \end{aligned}$$

we can use mathematics induction to prove why multi-block could work

4.4.3 Random-Permuted ADMM (RP-ADMM)

In each round, draw a random permutation $\sigma = (\sigma(1), \ldots, \sigma(n))$ of $\{1, \ldots, n\}$, and use the

Update Order : $\mathbf{x}_{\sigma(1)} \rightarrow \mathbf{x}_{\sigma(2)} \rightarrow \ldots \rightarrow \mathbf{x}_{\sigma(n)} \rightarrow \mathbf{y}$

- This is equivalent to a random sample without replacement so it costs nothing.
- Interpretation: Force "absolute fairness" among blocks.
- Simulation Test Result on solving linear equations: always converges!

We produced a positive result for ADMM on solving the system of linear equations.

Consider solving a nonsingular square system of linear equations ($f_i = 0, \forall i$). $\min_{\mathbf{x} \in R^N} \quad {}_0 \quad 0$ s.t. $A_1\mathbf{x}_1 + \cdots + A_n\mathbf{x}_n = \mathbf{b}$, RP-ADMM generates z^k, an r.v., depending on

$$\xi_k = (\sigma_1, \ldots, \sigma_k), \quad \mathbf{z}^i = M_{\sigma_i}\mathbf{z}^{i-1}, i = 1, \ldots, k,$$

where σ_i is the random permutation at i-th round. Denote the expected iterate $\phi^k := E_{\xi_k}(\mathbf{z}^k)$

Theorem 4.27. *The expected output converges to the unique solution of the linear system equations any integer $N \geq 1$.*

Remark 4.28. *Expected convergence $\neq$ convergence, but is a strong evidence for convergence for solving most problems, e.g., when iterates are bounded.*

- The update equation of RP-ADMM is

$$\mathbf{z}^{k+1} = M_\sigma \mathbf{z}^k$$

  where $M_\sigma \in R^{2N \times 2N}$ depend on σ.

- Define the expected update matrix as

$$M = E_\sigma \left(M_\sigma\right) = \frac{1}{n!} \sum_\sigma M_\sigma$$

Theorem 4.29. *The spectral radius of $M, \rho(M)$, is strictly less than 1 for any integer $N \geq 1$.*

Remark 4.30. *For A in the divergence example, $\rho\left(M_\sigma\right) > 1$ for any σ. Averaging Helps*

In general, consider a convex quadratic optimization problem

$$\begin{aligned} \min_{\mathbf{x} \in R^N} & \mathbf{c}_1^T \mathbf{x}_1 + \ldots + \mathbf{c}_n^T \mathbf{x}_n + \frac{1}{2} \mathbf{x}^T Q \mathbf{x} \\ \text{s.t. } & A\mathbf{x} := A_1 \mathbf{x}_1 + \cdots + A_n \mathbf{x}_n = \mathbf{b} \end{aligned}$$

Theorem 4.31. *Under some technical assumptions, the expected output of randomly permuted ADMM converges to the solution of the original problem for any integer $N \geq 1$.*

Here could we add the constraint $\mathbf{x}_i \geq 0$. In fact, it does not work by ADMM. Because in the iteration step, it is linear mapping. This specific constraint will block the linear mapping. Experiment shows it could work but proof we do not know.

- Non-square system of linear equations - ”yes”
- Non-separable convex quadratic minimization with linear equality constraints - ”yes”
- Convergence w.h.p.??
- Generalize to inequality systems or convex optimization at large??
- Generalize to non-convex optimization??
- ADMM where, in every iteration, each block are randomly assembled without replacement??

Software based on ADMM

- SCS: http://www. stanford. edu/”boyd/cvx for CLP

- ABIP: https: / github. com/sepvar/ABIP for solving LP
- RACQP: https://github. com/kmihic/RACQP for quadratic minimization with mixed continuous and integer decision variables.

Ref Chapter 8.6-7,9.1-9.5,10.1-4

4.5 Second Order Optimization Algorithms

4.5.1 1.5-Order Algorithm

Why the the method is 1.5-order? Because in the whole process, the Hessian matrix is involved. However, we approximate the second order into first order information. And some paper could provide the theoretical guarantee that it preform well.

Dimension Reduced Second Order Method

We expand to use the double direction method. If the problem is the convex function and the Hessian matrix is PSD and the coefficient matrix is invertible

Let $\mathrm{d}^k = \mathrm{x}^k - \mathrm{x}^{k-1}$, $\mathrm{g}^k = \nabla f\left(\mathrm{x}^k\right)$ and $H^k = \nabla^2 f\left(\mathrm{x}^k\right)$, then the step-sizes can be chosen from

$$\begin{pmatrix} \left(\mathbf{g}^k\right)^T H^k \mathbf{g}^k & -\left(\mathbf{d}^k\right)^T H^k \mathbf{g}^k \\ -\left(\mathbf{d}^k\right)^T H^k \mathbf{g}^k & \left(\mathbf{d}^k\right)^T H^k \mathbf{d}^k \end{pmatrix} \begin{pmatrix} \alpha^g \\ \alpha^m \end{pmatrix} = \begin{pmatrix} \left\|\mathbf{g}^k\right\|^2 \\ -\left(\mathbf{g}^k\right)^T \mathbf{d}^k \end{pmatrix} \tag{4.31}$$

If the Hessian $\nabla^2 f\left(\mathbf{x}^k\right)$ is not available, one can approximate. That's why this is dimension reduced method.

$$H^k \ \mathrm{g}^k \sim \nabla\left(\mathrm{x}^k + \mathrm{g}^k\right) - \mathrm{g}^k \text{ and } H^k \ \mathrm{d}^k \sim \nabla\left(\mathrm{x}^k + \mathrm{d}^k\right) - \mathrm{g}^k \sim -\left(\mathrm{g}^{k-1} - \mathrm{g}^k\right)$$

or for some small $\epsilon > 0$:

$$H^k \mathbf{g}^k \sim \frac{1}{\epsilon}\left(\nabla\left(\mathbf{x}^k + \epsilon \mathbf{g}^k\right) - \mathbf{g}^k\right) \text{ and } H^k \mathbf{d}^k \sim \frac{1}{\epsilon}\left(\nabla\left(\mathbf{x}^k + \epsilon \mathbf{d}^k\right) - \mathbf{g}^k\right).$$

Then we will introduce the Newton method

4.5.2 Ellipsoid Method

This is the first method to solve the linear program

Ellipsoids are just sets of the form

$$E = \left\{\mathbf{y} \in \mathbf{R}^m : (\mathbf{y} - \overline{\mathbf{y}})^T B^{-1} (\mathbf{y} - \overline{\mathbf{y}}) \leq 1\right\}$$

where $\overline{\mathbf{y}} \in R^m$ is a given point (called the center) and B is a symmetric positive definite matrix of dimension m. We can use the notation ell $(\overline{\mathbf{y}}, B)$ to specify the ellipsoid E defined above. Note that

$$\operatorname{vol}(E) = (\det B)^{1/2} \operatorname{vol}(S(\mathbf{0}, 1))$$

where $S(\mathbf{0}, 1)$ is the unit sphere in $\mathbf{R}^m$. and *vol* is the volume. How to iterate? We will use the half ellipsoid to generate the new one.

By a Half-Ellipsoid of E, we mean the set

$$\frac{1}{2} E_a := \left\{ \mathbf{y} \in E : \mathbf{a}^T \mathbf{y} \leq \mathbf{a}^T \overline{\mathbf{y}} \right\}$$

for a given non-zero vector $\mathbf{a} \in \mathbf{R}^m$ where $\overline{\mathbf{y}}$ is the center of E - the intersection of the ellipsoid and a plane cutting through the center. We are interested in finding a new ellipsoid containing $\frac{1}{2} E_a$ with the least volume.

The new ellipsoid $E^+ = \text{ell}\left(\overline{\mathbf{y}}^+, B^+\right)$can be constructed as follows. Define

$$\tau := \frac{1}{m+1}, \quad \delta := \frac{m^2}{m^2-1}, \quad \sigma := 2\tau$$

And let

$$\begin{aligned} \overline{\mathbf{y}}^+ &:= \overline{\mathbf{y}} - \frac{\tau}{\left(\mathbf{a}^{\mathrm{T}} B \mathbf{a}\right)^{1/2}} B \mathbf{a} \\ B^+ &:= \delta \left(B - \sigma \frac{B \mathbf{a} \mathbf{a}^{\mathrm{T}} B}{\mathbf{a}^{\mathrm{T}} B \mathbf{a}} \right) \end{aligned}$$

Theorem 4.32. *Ellipsoid E^+ =ell $\left(\overline{\mathbf{y}}^+, B^+\right)$defined as above is the ellipsoid of least volume containing $\frac{1}{2} E_a$. Moreover,*

$$\frac{\operatorname{vol}\left(E^+\right)}{\operatorname{vol}(E)} \leq \exp\left(-\frac{1}{2(m+1)}\right) \tag{4.32}$$

This theorem indicate the ellipsoid strictly decreases. consider a standard linear program

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ s.t. \quad & A\mathbf{x} \leq \mathbf{b} \end{aligned}$$

we could use penalty function to relax

$$\min \quad \mathbf{c}^T \mathbf{x} + M I(A\mathbf{x} - \mathbf{b})$$

Where M is big value

The ellipsoid method also work in the non-smooth objective function. In addition, we introduce the average case analysis.

In computational complexity theory, the average-case complexity of an algorithm is the amount of some computational resource (typically time) used by the algorithm, averaged over all possible inputs. It is frequently contrasted with worst-case complexity which considers the maximal complexity of the algorithm over all possible inputs.

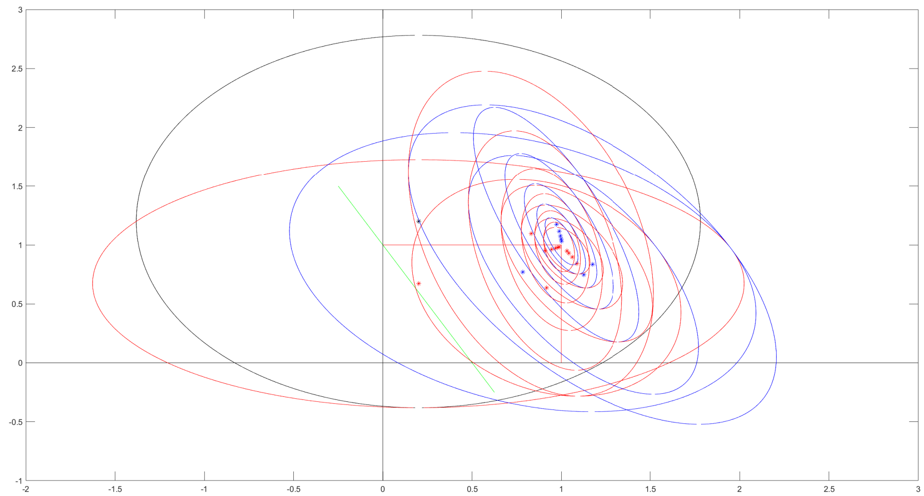

Figure 4.3: Ellipsoid Method

There are three primary motivations for studying average-case complexity. First, although some problems may be intractable in the worst-case, the inputs which elicit this behavior may rarely occur in practice, so the average-case complexity may be a more accurate measure of an algorithm's performance. Second, average-case complexity analysis provides tools and techniques to generate hard instances of problems which can be utilized in areas such as cryptography and derandomization. Third, average-case complexity allows discriminating the most efficient algorithm in practice among algorithms of equivalent best case complexity (for instance Quicksort).

Average-case analysis requires a notion of an "average" input to an algorithm, which leads to the problem of devising a probability distribution over inputs. Alternatively, a randomized algorithm can be used. The analysis of such algorithms leads to the related notion of an expected complexity

Example 4.33. *Qicksort has a worst-case running time of $O(n^2)$, but an average-case running time of $O(n \log n)$, where n is the length of the input to be sorted.*

4.5.3 Newton Method

Consider the two order Taylor expansion

$$f(\mathbf{x}) \approx f(\mathbf{x}_k) + \nabla f(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k)^T + \frac{1}{2}(\mathbf{x} - \mathbf{x}_k)^T \nabla^2 f(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k)$$

The right hand could be minimized by derivative

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [\nabla^2 f(\mathbf{x_k})]^{-1} \nabla f(\mathbf{x}_k)^T$$

the first modification is that usually a search parameter α

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha[\nabla^2 f(\mathbf{x_k})]^{-1}\nabla f(\mathbf{x}_k)^T$$

We now introduce the second-order β-Lipschitz condition: for any point x and direction vector d

$$\left\|\nabla f(\mathbf{x}+\mathbf{d}) - \nabla f(\mathbf{x}) - \nabla^2 f(\mathbf{x})\mathbf{d}\right\| \leq \beta\|\mathbf{d}\|^2,$$

which implies

$$f(\mathbf{x}+\mathbf{d}) - f(\mathbf{x}) \leq \nabla f(\mathbf{x})^T\mathbf{d} + \frac{1}{2}\mathbf{d}^T\nabla^2 f(\mathbf{x})\mathbf{d} + \frac{\beta}{3}\|\mathbf{d}\|^3.$$

In the following, for notation simplicity, we use $\mathbf{g}(\mathbf{x}) = \nabla f(\mathbf{x})$ and $\nabla g(\mathbf{x}) = \nabla^2 f(\mathbf{x})$. Thus,

$$\mathrm{x}^{k+1} = \mathrm{x}^k - \left(\nabla \mathrm{g}\left(\mathrm{x}^k\right)\right)^{-1}\nabla f\left(\mathrm{x}^k\right), \text{ or } \nabla \mathrm{g}\left(\mathrm{x}^k\right)\left(\mathrm{x}^{k+1} - \mathrm{x}^k\right) = -\mathrm{g}\left(\mathrm{x}^k\right)$$

Indeed, Newton's method was initially developed for solving a system of nonlinear equations in the form g(x) = 0

Theorem 4.34. *Let $f(\mathbf{x})$ be β-Lipschitz and the smallest absolute eigenvalue of its Hessian uniformly bounded below by $\lambda_{\min} > 0$. Then, provided that $\left\|\mathrm{x}^0 - \mathrm{x}^*\right\|$ is sufficiently small, the sequence generated by Newton's method converges quadratically to x^* that is a KKT solution with $\mathrm{g}\left(\mathrm{x}^*\right) = 0$.*

$$\begin{aligned}
\left\|\mathrm{x}^{k+1} - \mathrm{x}^*\right\| &= \left\|\mathrm{x}^k - \mathrm{x}^* - \nabla \mathrm{g}\left(\mathrm{x}^k\right)^{-1}\ \mathrm{g}\left(\mathrm{x}^k\right)\right\| \\
&= \left\|\nabla \mathbf{g}\left(\mathrm{x}^k\right)^{-1}\left(\ \mathrm{g}\left(\mathrm{x}^k\right) - \nabla \mathrm{g}\left(\mathrm{x}^k\right)\left(\mathrm{x}^k - \mathrm{x}^*\right)\right)\right\| \\
&= \left\|\nabla \mathbf{g}\left(\mathrm{x}^k\right)^{-1}\left(\ \mathrm{g}\left(\mathrm{x}^k\right) - \mathrm{g}\left(\mathrm{x}^*\right) - \nabla \mathrm{g}\left(\mathrm{x}^k\right)\left(\mathrm{x}^k - \mathrm{x}^*\right)\right)\right\| \\
&\leq \left\|\nabla \mathbf{g}\left(\mathrm{x}^k\right)^{-1}\right\|\left\|\mathrm{g}\left(\mathrm{x}^k\right) - \mathrm{g}\left(\mathrm{x}^*\right) - \nabla \mathrm{g}\left(\mathrm{x}^k\right)\left(\mathrm{x}^k - \mathrm{x}^*\right)\right\| \\
&\leq \left\|\nabla \mathrm{g}\left(\mathrm{x}^k\right)^{-1}\right\|\beta\left\|\mathrm{x}^k - \mathrm{x}^*\right\|^2 \leq \frac{\beta}{\lambda_{\min}}\left\|\mathrm{x}^k - \mathrm{x}^*\right\|^2
\end{aligned}$$

Thus, when $\frac{\beta}{\lambda_{\min}}\left\|\mathrm{x}^0 - \mathrm{x}^*\right\| < 1$, the quadratic convergence takes place:

$$\frac{\beta}{\lambda_{\min}}\left\|\mathbf{x}^{k+1} - \mathbf{x}^*\right\| \leq \left(\frac{\beta}{\lambda_{\min}}\left\|\mathbf{x}^k - \mathbf{x}^*\right\|\right)^2. \tag{4.33}$$

Such a starting solution x^0 is called an approximate root of g(x).

Bibliography

Ahuja, Ravindra K, James B Orlin. 2001. Inverse optimization. *Operations Research* **49**(5) 771–783.

Bertsimas, Dimitris, John N Tsitsiklis. 1997. *Introduction to linear optimization*, vol. 6. Athena Scientific Belmont, MA.

Boyd, Stephen, Stephen P Boyd, Lieven Vandenberghe. 2004. *Convex optimization*. Cambridge university press.

Golub, Gene H, Charles F Van Loan. 2013. *Matrix computations*. JHU press.

(https://math.stackexchange.com/users/474939/gerhard s), Gerhard S. ???? How to show negative entropy function $f(x) = x\log(x)$ is strongly convex? Mathematics Stack Exchange. URL `https://math.stackexchange.com/q/3077306`. URL:https://math.stackexchange.com/q/3077306 (version: 2019-01-17).

Luenberger, David G, Yinyu Ye, et al. 1984. *Linear and nonlinear programming*, vol. 2. Springer.

Nesterov, Y. ???? A method of solving a convex programming problem with convergence rate {O}(1/kˆ{2}). *Sov. Math. Dokl*, vol. 27.

Polyak, Boris T. 1964. Some methods of speeding up the convergence of iteration methods. *Ussr computational mathematics and mathematical physics* **4**(5) 1–17.

Senning, Jonathan R. 2007. Computing and estimating the rate of convergence.

Smale, Steve. 1986. Newton's method estimates from data at one point. *The merging of disciplines: new directions in pure, applied, and computational mathematics*. Springer, 185–196.

Ye, Yinyu. 2011. *Interior point algorithms: theory and analysis*. John Wiley & Sons.