Mathematical Optimization Model/Theory Review II

Yinyu Ye
Department of Management Science and Engineering
Stanford University
Stanford, CA 94305, U.S.A.

http://www.stanford.edu/~yyye

Chapters 1, 2.1-2, 6.1-2, 7.2, 11.3, 11.6

Unstructured Optimization

Now consider the general (constrained) optimization (GCO) problem:

$$
\begin{array}{lll}
\text{(P)} & \text{minimize} & f(\mathbf{x}) \\
 & \text{subject to} & c_i(\mathbf{x})\ (\leq, = \geq)\ 0 \quad i = 1, ..., m
\end{array}
$$

Optimality Conditions help to identify and verify when a solution is optimal.

First-Order Necessary Conditions for Constrained Optimization I

Consider constraints $\{\mathbf{x}: \ \mathbf{h}(\mathbf{x}) = \mathbf{0}, \ \mathbf{c}(\mathbf{x}) \geq \mathbf{0}.\}$

Lemma 1 *Let $\bar{\mathbf{x}}$ be a feasible solution and a regular point of the hypersurface of*

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

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

$$\begin{aligned} \nabla f(\bar{\mathbf{x}})\mathbf{d} & < 0 \\ \nabla \mathbf{h}(\bar{\mathbf{x}})\mathbf{d} & = \mathbf{0} \in R^m, \\ \nabla c_i(\bar{\mathbf{x}})\mathbf{d} & \geq 0, \ \forall i \in \mathcal{A}_{\bar{\mathbf{x}}}. \end{aligned} \tag{1}$$

This lemma was proved when constraints are linear in which case $\mathbf{d}$ is a feasible direction, but needs more work otherwise since there is no feasible direction when constraints are nonlinear.

$\bar{\mathbf{x}}$ being a regular point is often referred as a Constraint Qualification condition.

First-Order Necessary Conditions for Constrained Optimization II

Theorem 1 *(First-Order or KKT Optimality Condition) Let* $\bar{\mathbf{x}}$ *be a (local) minimizer of (GCO) and it is a regular point of* $\{\mathbf{x} : \ \mathbf{h}(\mathbf{x}) = \mathbf{0}, \ c_i(\mathbf{x}) = 0, i \in \mathcal{A}_{\bar{\mathbf{x}}}\}$. *Then, for some multipliers* $(\bar{\mathbf{y}}, \bar{\mathbf{s}} \geq \mathbf{0})$

$$\nabla f(\bar{\mathbf{x}}) = \bar{\mathbf{y}}^T \nabla \mathbf{h}(\bar{\mathbf{x}}) + \bar{\mathbf{s}}^T \nabla \mathbf{c}(\bar{\mathbf{x}}) \tag{2}$$

and (complementarity slackness)

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

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 $\bar{\mathbf{x}}$, if it is also a regular point, must be an KKT solution; but the reverse may not be true.

Constraint Qualification and the KKT Theorem

One condition for a local minimizer $\bar{\mathbf{x}}$ that must always be an KKT solution is the constraint qualification: $\bar{\mathbf{x}}$ is a regular point of the constraints. Otherwise, a local minimizer may not be an KKT solution: Consider $\bar{\mathbf{x}} = (0; 0)$ of a convex nonlinearly-constrained problem

$$\min \ x_1, \quad \text{s.t.} \quad x_1^2 + (x_2 - 1)^2 - 1 \leq 0, \ x_1^2 + (x_2 + 1)^2 - 1 \leq 0\}.$$

On the other hand, even the regular point condition does not hold, the KKT theorem may still true:

$$\min \ x_2, \quad \text{s.t.} \quad x_1^2 + (x_2 - 1)^2 - 1 \leq 0, \ x_1^2 + (x_2 + 1)^2 - 1 \leq 0\},$$

that is, $\bar{\mathbf{x}} = (0; 0)$ is an KKT solution of the latter problem.

Therefore, finding an KKT solution is a plausible way to find a local minimizer.

KKT via the Lagrangian Function

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

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

where multipliers $\mathbf{y}$ of the equality constraints are “free” and $\mathbf{s} \geq \mathbf{0}$ for the “greater or equal to” inequality constraints, so that the KKT condition (2) can be written as

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

Lagrangian Function can be viewed as a “penalty” function aggregated with the original objective function plus the penalized terms on constraint violations.

In theory, one can adjust the penalty multipliers $(\mathbf{y}, \mathbf{s} \geq \mathbf{0})$ to repeatedly solve the following so-called Lagrangian Relaxation Problem:

$$(LRP) \quad \min_{\mathbf{x}} \quad L(\mathbf{x}, \mathbf{y}, \mathbf{s}).$$

Summary of KKT Conditions for More General GCO

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

the Lagrangian Function is given by

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

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 $\bar{\mathbf{x}}$ be a local minimizer for (GCO) and $\bar{\mathbf{x}}$ is a regular point on the hypersurface of the active constraints Then there exist multipliers $\bar{\mathbf{y}}$ such that

$$\begin{array}{rcll} \nabla f(\bar{\mathbf{x}}) & = & \bar{\mathbf{y}}^T \nabla \mathbf{c}(\bar{\mathbf{x}}) & \text{(Lagrangian Derivative Conditions (LDC))} \\ \bar{y}_i & (\leq, '\text{free}', \geq) & 0, \; i = 1, ..., m, & \text{(Multiplier Sign Constraints (MSC))} \\ \bar{y}_i c_i(\bar{\mathbf{x}}) & = & 0, & \text{(Complementarity Slackness Conditions (CSC))}. \end{array}$$

Second-Order Necessary Conditions for Constrained Optimization

Now in addition we assume all functions are in C^2, that is, twice continuously differentiable. Recall the tangent linear sub-space at $\bar{\mathbf{x}}$:

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

Theorem 2 *Let $\bar{\mathbf{x}}$ be a (local) minimizer of (GCO) and a regular point of hypersurface $\{\mathbf{x} :\ \mathbf{h}(\mathbf{x}) = \mathbf{0},\ c_i(\mathbf{x}) = 0, i \in \mathcal{A}_{\bar{\mathbf{x}}}\}$, and let $\bar{\mathbf{y}}, \bar{\mathbf{s}}$ denote Lagrange multipliers such that $(\bar{\mathbf{x}}, \bar{\mathbf{y}}, \bar{\mathbf{s}})$ satisfies the (first-order) KKT conditions of (GCO). Then, it is necessary to have*

$$\mathbf{d}^T \, \nabla^2_{\mathbf{x}} L(\bar{\mathbf{x}}, \bar{\mathbf{y}}, \bar{\mathbf{s}})\mathbf{d} \geq 0 \qquad \forall\, \mathbf{d} \in T_{\bar{\mathbf{x}}}.$$

The Hessian of the Lagrangian function need to be positive semidefinite on the tangent-space.

Second-Order Sufficient Conditions for GCO

Theorem 3 *Let* $\bar{\mathbf{x}}$ *be a regular point of (GCO) with* **equality constraints only** *and let* $\bar{\mathbf{y}}$ *be the Lagrange multipliers such that* $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ *satisfies the (first-order) KKT conditions of (GCO). Then, if in addition*

$$\mathbf{d}^T \nabla^2_{\mathbf{x}} L(\bar{\mathbf{x}}, \bar{\mathbf{y}})\mathbf{d} > 0 \qquad \forall\, \mathbf{0} \neq \mathbf{d} \in T_{\bar{\mathbf{x}}},$$

then $\bar{\mathbf{x}}$ *is a local minimizer of (GCO).*

See the proof in Chapter 11.5 of LY.

The SOSC for general (GCO) is proved in Chapter 11.8 of LY.

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

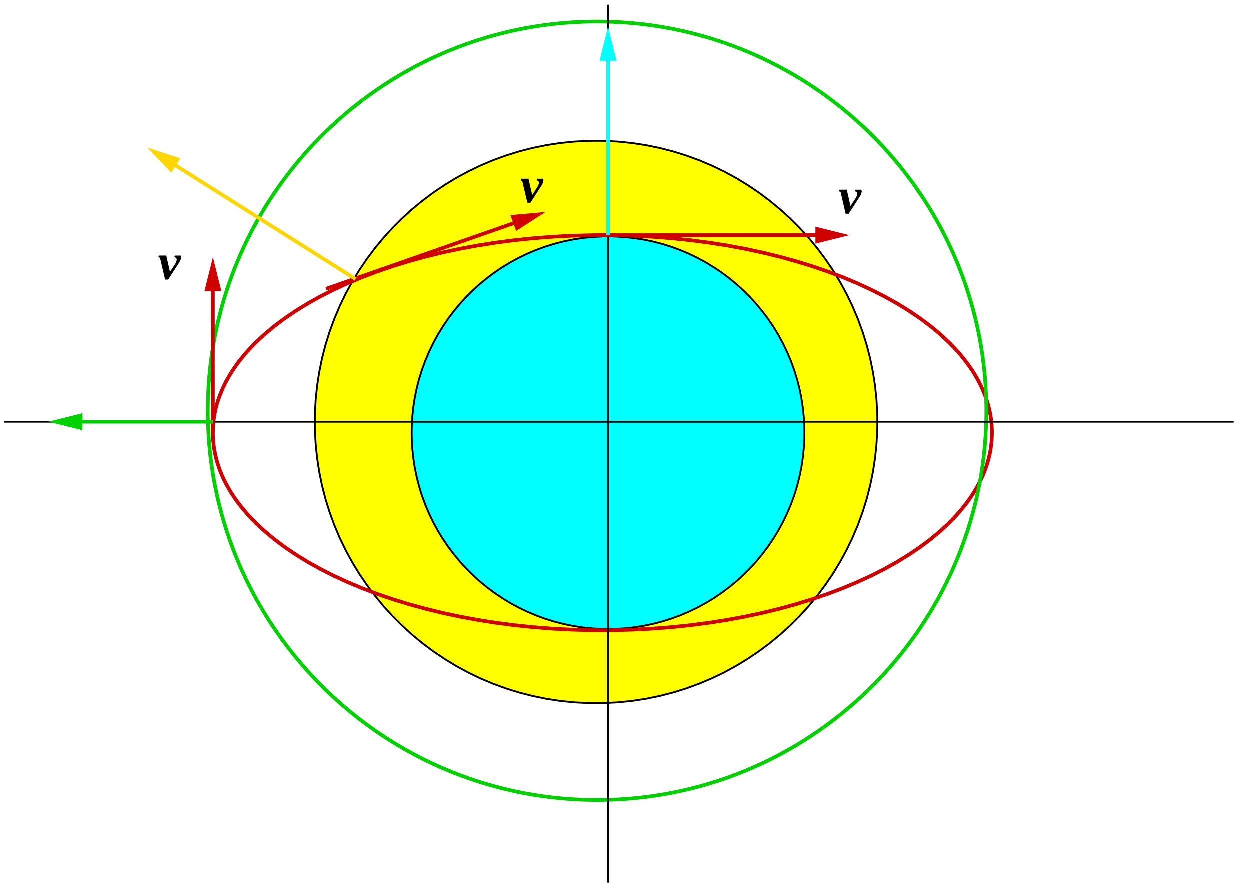

Figure 1: FONC and SONC for Constrained Minimization

KKT Conditions: Fisher's Equilibrium Price

Player $i \in B$'s optimization problem for given prices p_j, $j \in G$.

$$\begin{array}{ll} \text{maximize} & \mathbf{u}_i^T \mathbf{x}_i := \sum_{j\in G} u_{ij} x_{ij} \\ \text{subject to} & \mathbf{p}^T \mathbf{x}_i := \sum_{j\in G} p_j x_{ij} \le w_i, \\ & x_{ij} \ge 0, \quad \forall j, \end{array}$$

Assume that the amount of each good is $\bar{s}_j$. The equilinitum price vector is the one that for all $j \in G$

$$\sum_{i\in B} x(\mathbf{p})_{ij} = \bar{s}_j$$

Example of Fisher's Equilibrium Price

There two goods, x and y, each with 1 unit on the market. Buyer $1, 2$'s optimization problems for given prices p_x, p_y.

$$
\begin{array}{ll}
\text{maximize} & 2x_1 + y_1 \\
\text{subject to} & p_x \cdot x_1 + p_y \cdot y_1 \le 5, \\
& x_1, y_1 \ge 0; \\
\text{maximize} & 3x_2 + y_2 \\
\text{subject to} & p_x \cdot x_2 + p_y \cdot y_2 \le 8, \\
& x_2, y_2 \ge 0.
\end{array}
$$

$$p_x = \frac{26}{3},\ p_y = \frac{13}{3},\ x_1 = \frac{1}{13},\ y_1 = 1,\ x_2 = \frac{12}{13},\ y_2 = 0$$

Equilibrium Price Conditions

Player $i \in B$'s dual problem for given prices p_j, $j \in G$.

$$\begin{array}{ll} \text{minimize} & w_i y_i \\ \text{subject to} & \mathbf{p} y_i \ge \mathbf{u}_i,\ y_i \ge 0 \end{array}$$

The necessary and sufficient conditions for an equilibrium point $\mathbf{x}_i, \mathbf{p}$ are:

$$\begin{array}{cc} \mathbf{p}^T\mathbf{x}_i = w_i,\ \mathbf{x}_i \ge \mathbf{0}, & \forall i, \\ p_j y_i \ge u_{ij},\ y_i \ge 0, & \forall i, j, \\ \mathbf{u}_i^T\mathbf{x}_i = w_i y_i, & \forall i, \\ \sum_i x_{ij} = \bar{s}_j, & \forall j. \end{array} \quad <=> \quad \begin{array}{cc} \mathbf{p}^T\mathbf{x}_i = w_i,\ \mathbf{x}_i \ge \mathbf{0}, & \forall i, \\ p_j \ge w_i \frac{u_{ij}}{\mathbf{u}_i^T\mathbf{x}_i}, & \forall i, j, \\ \sum_i x_{ij} = \bar{s}_j, & \forall j. \end{array}$$

Equilibrium Price Conditions (continued)

These conditions can be equivalently represented by

$$\begin{aligned} \textstyle\sum_j \bar{s}_j p_j \le \sum_i w_i,\ \mathbf{x}_i \ge \mathbf{0}, &\quad \forall i, \\ p_j \ge w_i \frac{u_{ij}}{\mathbf{u}_i^T \mathbf{x}_i}, &\quad \forall i, j, \\ \textstyle\sum_i x_{ij} = \bar{s}_j, &\quad \forall j. \end{aligned}$$

since from the second inequality (after multiplying x_{ij} to both sides and take sum over j) we have

$$\mathbf{p}^T \mathbf{x}_i \ge w_i,\ \forall i.$$

Then, from the rest conditions

$$\sum_i w_i \ge \sum_j \bar{s}_j p_j = \sum_i \mathbf{p}^T \mathbf{x}_i \ge \sum_i w_i.$$

Thus, every inequality in the sequel has to be equal, that is, $\mathbf{p}^T \mathbf{x}_i = w_i,\ \forall i$ and $p_j x_{ij} = w_i \frac{u_{ij} x_{ij}}{\mathbf{u}_i^T \mathbf{x}_i},\ \forall i, j$.

Equilibrium Price Property

If u_{ij} has at least one positive coefficient for every j, then we must have $p_j > 0$ for every j at every equilibrium. Moreover, The second inequality can be rewritten as

$$\log(\mathbf{u}_i^T \mathbf{x}_i) + \log(p_j) \geq \log(w_i) + \log(u_{ij}), \ \forall i, j, \ u_{ij} > 0.$$

The function on the left is (strictly) concave in $\mathbf{x}_i$ and p_j. Thus,

Theorem 4 *The equilibrium set of the Fisher Market is convex, and the equilibrium price vector is unique.*

Aggregate Social Optimization

$$\begin{array}{ll} \text{maximize} & \sum_{i\in B} w_i \log(\mathbf{u}_i^T \mathbf{x}_i) \\ \text{subject to} & \sum_{i\in B} x_{ij} = \bar{s}_j, \quad \forall j \in G, \; x_{ij} \geq 0, \quad \forall i, j. \end{array}$$

Theorem 5 *(Eisenberg and Gale 1959) Optimal dual (Lagrange) multiplier vector of equality constraints is an equilibrium price vector.*

The proof is from Optimality Conditions of the Aggregate Social Problem:

$$\begin{array}{rcl} w_i \frac{u_{ij}}{\mathbf{u}_i^T \mathbf{x}_i} & \leq & p_j, \quad \forall i, j \\ w_i \frac{u_{ij} x_{ij}}{\mathbf{u}_i^T \mathbf{x}_i} & = & p_j x_{ij}, \quad \forall i, j \quad \text{(complementarity)} \\ \sum_i x_{ij} & = & \bar{s}_j, \quad \forall j \\ \mathbf{x}_i & \geq & \mathbf{0}, \; \forall i, \end{array}$$

which is identical to the equilibrium conditions described earlier.

Rewrite Aggregate Social Optimization

$$\begin{array}{ll} \text{maximize} & \sum_{i \in B} w_i \log u_i \\ \text{subject to} & \sum_{j \in G} u_{ij}^T x_{ij} - u_i = 0, \quad \forall i \in B \\ & \sum_{i \in B} x_{ij} = \bar{s}_j, \quad \forall j \in G \\ & x_{ij} \geq 0,\ u_i \geq 0, \quad \forall i, j, \end{array}$$

This is called the weighted analytic center problem.

Question: Is the price vector $\mathbf{p}$ unique when at least one $u_{ij} > 0$ among $i \in B$ and $u_{ij} > 0$ among $j \in G$.

Aggregate Example:

$$\begin{array}{ll} \text{maximize} & 5\log(2x_1 + y_1) + 8\log(3x_2 + y_2) \\ \text{subject to} & x_1 + x_2 = 1, \\ & y_1 + y_2 = 1, \\ & x_1, x_2, y_1, y_2 \geq 0. \end{array}$$

Lagrangian Function and Dual

Consider the general constrained optimization again:

$$\text{(GCO)} \qquad \begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & c_i(\mathbf{x}) \quad (\leq, =, \geq) \quad 0, \; i = 1, ..., m, \end{array}$$

For Lagrange Multipliers.

$$Y := \{y_i \quad (\leq, '\text{free}', \geq) \quad 0, \; i = 1, ..., m\},$$

the Lagrangian Function is again given by

$$L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - \mathbf{y}^T \mathbf{c}(\mathbf{x}) = f(\mathbf{x}) - \sum_{i=1}^{m} y_i c_i(\mathbf{x}), \; \mathbf{y} \in Y.$$

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.

Toy Example Again

$$\begin{array}{ll} \text{minimize} & (x_1-1)^2+(x_2-1)^2 \\ \text{subject to} & x_1+2x_2-1 \le 0, \\ & 2x_1+x_2-1 \le 0. \end{array}$$

$$L(\mathbf{x},\mathbf{y}) = f(\mathbf{x}) - \mathbf{y}^T\mathbf{c}(\mathbf{x}) = f(\mathbf{x}) - \sum_{i=1}^{2} y_i c_i(\mathbf{x}) =$$

$$= (x_1-1)^2+(x_2-1)^2 - y_1(x_1+2x_2-1) - y_2(2x_1+x_2-1),\ (y_1;y_2) \le \mathbf{0}$$

where

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

Lagrangian Relaxation Problem

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

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

Again, $\mathbf{y}_i$ can be viewed as a penalty parameter to penalize constraint violation $c_i(\mathbf{x}), i = 1, ..., m$.

In the toy example, for given $(y_1; y_2) \leq \mathbf{0}$, the LRP is:

$$\begin{array}{ll} \inf & (x_1 - 1)^2 + (x_2 - 1)^2 - y_1(x_1 + 2x_2 - 1) - y_2(2x_1 + x_2 - 1) \\ \text{s.t.} & (x_1; x_2) \in R^2, \end{array}$$

and it has a close form solution $\mathbf{x}$ for any given $\mathbf{y}$:

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

with the minimal or infimum value function $= -1.25y_1^2 - 1.25y_2^2 - 2y_1y_2 - 2y_1 - 2y_2$.

Inf-Value Function as the Dual Objective

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

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

Theorem 6 *The Lagrangian dual objective* $\phi(\mathbf{y})$ *is a concave function.*

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

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

Dual Objective Establishes a Lower Bound

Theorem 7 *(Weak duality theorem) For every $\mathbf{y} \in Y$, the Lagrangian dual function $\phi(\mathbf{y})$ is less or equal to the infimum value of the original GCO problem.*

Proof:

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

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 $-\mathbf{y}^T \mathbf{c}(\mathbf{x}) \leq 0$.

The Lagrangian Dual of Classical LP I

Consider LP problem

$$(LP) \quad \text{minimize} \quad \mathbf{c}^T\mathbf{x}$$
$$\text{subject to} \quad A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0};$$

and its conic dual problem is given by

$$(LD) \quad \text{maximize} \quad \mathbf{b}^T\mathbf{y}$$
$$\text{subject to} \quad A^T\mathbf{y} + \mathbf{s} = \mathbf{c}, \ \mathbf{s} \geq \mathbf{0}.$$

We now derive the Lagrangian Dual of (LP). Let the Lagrangian multipliers be $\mathbf{y}('\text{free}')$ for equalities and $\mathbf{s} \geq \mathbf{0}$ for constraints $\mathbf{x} \geq \mathbf{0}$. Then the Lagrangian function would be

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

where $\mathbf{x}$ is “free”.

The Lagrangian Dual of Classical LP II

Now consider the Lagrangian dual objective

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

If $(\mathbf{c} - A^T\mathbf{y} - \mathbf{s}) \neq \mathbf{0}$, then $\phi(\mathbf{y},\mathbf{s}) = -\infty$. Thus, in order to maximize $\phi(\mathbf{y},\mathbf{s})$, the dual must choose its variables $(\mathbf{y},\mathbf{s} \geq \mathbf{0})$ such that $(\mathbf{c} - A^T\mathbf{y} - \mathbf{s}) = \mathbf{0}$.

This constraint, together with the sign constraint $\mathbf{s} \geq \mathbf{0}$, establish the Lagrangian dual problem:

$$\begin{array}{lll} (LDP) & \text{maximize} & \mathbf{b}^T\mathbf{y} \\ & \text{subject to} & A^T\mathbf{y} + \mathbf{s} = \mathbf{c},\ \mathbf{s} \geq \mathbf{0}. \end{array}$$

which is consistent with the conic dual of LP.

Lagrangian Strong Duality Theorem

Theorem 8 *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 $\hat{\mathbf{x}}$ such that all inequality constraints are strictly held. Then, $f^* = \phi^*$, and (LDP) admits a maximizer $\mathbf{y}^*$ such that*

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

Furthermore, if (GCO) admits a minimizer $\mathbf{x}^*$, then

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

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.

Note that the problem would be a convex minimization problem if all equality constraints are hyperplane or affine functions $c_i(\mathbf{x}) = \mathbf{a}_i \mathbf{x} - b_i$, all other level sets are convex.

More on Lagrangian Duality

Consider the constrained problem with additional constraints

$$
\begin{array}{lll}
(GCO) & \inf & f(\mathbf{x}) \\
 & \text{s.t.} & \mathbf{c}_i(\mathbf{x})\ (\leq, =, \geq)\ 0,\ i = 1, ..., m, \\
 & & \mathbf{x} \in \Omega \subset R^n.
\end{array}
$$

Typically, Ω has a simple form such as the cone

$$\Omega = R^n_+ = \{\mathbf{x} : \ \mathbf{x} \geq \mathbf{0}\}$$

or the box

$$\Omega := \{\mathbf{x} : \ -\mathbf{e} \leq \mathbf{x} \leq \mathbf{e}.\}$$

Then, when derive the Lagrangian dual, there is not need to introduce multipliers for Ω constraints.

Lagrangian Relaxation Problem

Consider again the (partial) Lagrangian Function:

$$L(\mathbf{x},\mathbf{y}) = f(\mathbf{x}) - \mathbf{y}^T\mathbf{c}(\mathbf{x}),\ \mathbf{y} \in Y;$$

and define the dual objective function of $\mathbf{y}$ be

$$\begin{array}{rll} \phi(\mathbf{y}) := & \inf_{\mathbf{x}} & L(\mathbf{x},\mathbf{y}) \\ & \text{s.t.} & \mathbf{x} \in \Omega. \end{array}$$

Theorem 9 *The Lagrangian dual function $\phi(\mathbf{y})$ is a concave function.*

Theorem 10 *(Weak duality theorem) For every $\mathbf{y} \in Y$, the Lagrangian dual function value $\phi(\mathbf{y})$ is less or equal to the infimum value of the original GCO problem.*

The Lagrangian Dual Problem

$$(LDP) \quad \sup \quad \phi(\mathbf{y})$$
$$\text{s.t.} \quad \mathbf{y} \in Y.$$

would called the Lagrangian dual of the original GCO problem:

Theorem 11 *(Strong duality theorem) Let (GCO) be a convex minimization problem, 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 $\hat{\mathbf{x}}$ such that all inequality constraints are strictly held. Then, $f^* = \phi^*$, and (LDP) admits a maximizer $\mathbf{y}^*$ such that*

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

Furthermore, if (GCO) admits a minimizer $\mathbf{x}^*$, then

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

Rules to Construct the Lagrangian Dual

$$\text{(GCO)} \quad \begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & c_i\mathbf{x}) \quad (\le, =, \ge) \quad 0, \; i = 1, ..., m, \end{array}$$

- All multipliers are dual variables.
- Derive the LDC

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

If no $\mathbf{x}$ appeared in an equation, set it as an equality constraint for the dual; otherwise, express $\mathbf{x}$ in terms of $\mathbf{y}$ and replace $\mathbf{x}$ in the Lagrange function, which becomes the Dual objective. (This may be very difficult ...)

- Add the MSC as dual constraints.

The Conic Duality vs. Lagrangian Duality I

Consider SOCP problem

$$(SOCP) \quad \text{minimize} \quad \mathbf{c}^T\mathbf{x}$$
$$\text{subject to} \quad A\mathbf{x} = \mathbf{b},\ x_1 - \|\mathbf{x}_{-1}\|_2 \geq 0;$$

and it conic dual problem

$$(SOCD) \quad \text{maximize} \quad \mathbf{b}^T\mathbf{y}$$
$$\text{subject to} \quad A^T\mathbf{y} + \mathbf{s} = \mathbf{c},\ s_1 - \|\mathbf{s}_{-1}\|_2 \geq 0.$$

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

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

The Conic Duality vs. Lagrangian Duality II

Now consider the Lagrangian dual objective

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

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.