Optimality Conditions for General Constrained Optimization

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

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

Chapter 11.1-11.8

General Constrained Optimization

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

We have dealt the cases when the feasible region is a convex polyhedron and/or the feasible can be represented by nonlinear convex cones intersect linear equality constraints.

We now study the case that the only assumption is that all functions are in C^1, and C^2 later, either convex or nonconvex.

We again establish optimality conditions to qualify/verify any local optimizers. These conditions give us qualitative structures of (local) optimizers and lead to quantitative algorithms to numerically find a local optimizer or an KKT solution.

The main proof idea is that if $\bar{\mathbf{x}}$ is a local minimier of (GCO), then it must be a local minimizer of the problem where the constraints are linearlized using the First-Order Taylor expansion.

Hypersurface and Implicit Function Theorem

Consider the (intersection) of Hypersurfaces (vs. Hyperplanes):

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

When functions $h_i(\mathbf{x})$s are C^1 functions, we say the surface is smooth.

For a point $\bar{\mathbf{x}}$ on the surface, we call it a regular point if $\nabla \mathbf{h}(\bar{\mathbf{x}})$ have rank m or the rows, or the gradient vector of each $h_i(\cdot)$ at $\bar{\mathbf{x}}$, are linearly independent. For example, $(0; \ 0)$ is not a regular point of $\{(x_1; \ x_2) \in R^2 : \ x_1^2 + (x_2 - 1)^2 - 1 = 0, \ x_1^2 + (x_2 + 1)^2 - 1 = 0\}$.

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

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

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

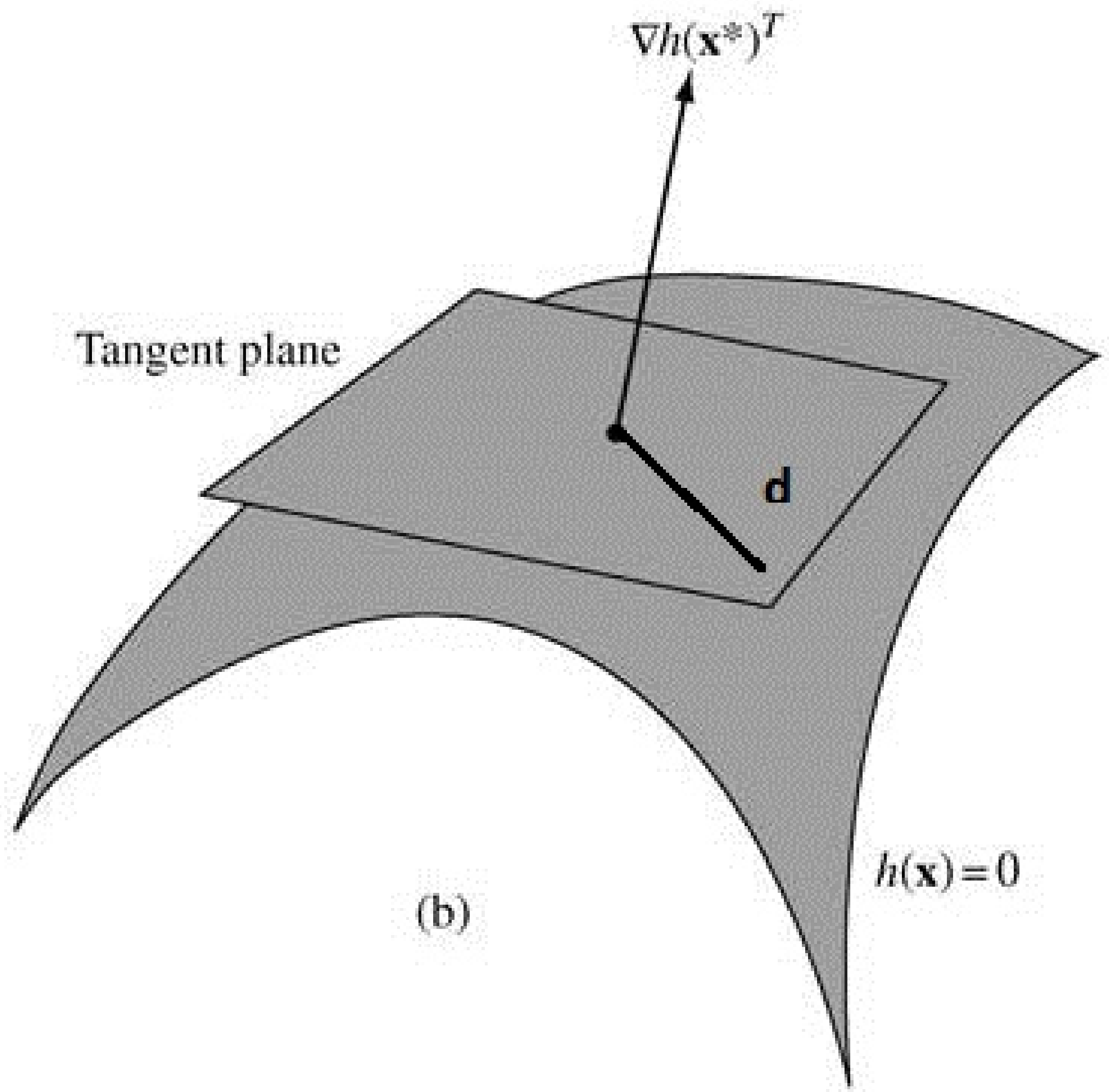

Figure 1: Tangent Plane on a Hypersurface at Point $\mathbf{x}^*$

First-Order Necessary Conditions for Constrained Optimization I

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:

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

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.

Proof

Suppose we have a $\bar{\mathbf{d}}$ satisfies all linear constraints. Then $\nabla f(\bar{\mathbf{x}})\bar{\mathbf{d}} < 0$ so that $\bar{\mathbf{d}}$ is a descent-direction vector. Denote the active-constraint set at $\bar{\mathbf{d}}$ among the linear inequalities in (1) by $\mathcal{A}^d_{\bar{\mathbf{x}}}$ $(\subset \mathcal{A}_{\bar{\mathbf{x}}})$. Then, $\bar{\mathbf{x}}$ remains a regular point of hypersurface of

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

Thus, there is a curve $\mathbf{x}(t)$ such that

$$\mathbf{h}(\mathbf{x}(t)) = \mathbf{0}, \quad c_i(\mathbf{x}(t)) = 0, \ i \in \mathcal{A}^d_{\bar{\mathbf{x}}}, \quad \mathbf{x}(0) = \bar{\mathbf{x}}, \quad \dot{\mathbf{x}}(0) = \bar{\mathbf{d}},$$

for $t \in [0 \quad a]$ of a sufficiently small positive constant a.

Also, $\nabla c_i(\bar{\mathbf{x}})\bar{\mathbf{d}} > 0, \ \forall i \notin \mathcal{A}^d_{\bar{\mathbf{x}}}$ but $i \in \mathcal{A}_{\bar{\mathbf{x}}}$; and $c_i(\bar{\mathbf{x}}) > 0, \ \forall i \notin \mathcal{A}_{\bar{\mathbf{x}}}$. Then, from Taylor's theorem, $c_i(\mathbf{x}(t)) > 0$ for all $i \notin \mathcal{A}^d_{\bar{\mathbf{x}}}$ so that $\mathbf{x}(t)$ is a feasible curve to the original (GCO) problem for $t \in [0 \quad a]$. Thus, $\bar{\mathbf{x}}$ must be also a local minimizer among all local solutions on the curve $\mathbf{x}(t)$.

Let $\phi(t) = f(\mathbf{x}(t))$. Then, $t = 0$ must be a local minimizer of $\phi(t)$ for $0 \le t \le a$ so that

$$0 \le \phi'(0) = \nabla f(\mathbf{x}(0))\dot{\mathbf{x}}(0) = \nabla f(\bar{\mathbf{x}})\bar{\mathbf{d}} < 0, \ \Rightarrow \text{a contradiction.}$$

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.

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 function aggregated 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:

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

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.

Summary Theorem of KKT Conditions for GCO

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

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

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 $\bar{\mathbf{x}}$ 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

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

The complete First-Order KKT Conditions consist of these four parts!

Recall SOCP Relaxation of Sensor Network Localization

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

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

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

$$\begin{array}{rcll} \mathbf{0} & = & \sum_{k=1}^{3} y_k(\mathbf{x} - \mathbf{a}_k) & \text{(LDC)} \\ y_k & \le & 0,\ k = 1, 2, 3, & \text{(MSC)} \\ y_k(\|\mathbf{x} - \mathbf{a}_k\|^2 - d_k^2) & = & 0. & \text{(CSC)}. \end{array}$$

Arrow-Debreu's Exchange Market with Linear Economy

Each trader i, equipped with a good bundle vector $\mathbf{w}_i$, trade with others to maximize its individual utility function. The equilibrium price is an assignment of prices to goods so as when every producer sells his/her own good bundle and buys a maximal bundle of goods then the market clears. Thus, trader i's optimization problem, for given prices p_j, $j \in G$, is

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

Then, the equilibrium price vector is the one such that there are maximizers $\mathbf{x}(\mathbf{p})_i$s

$$\sum_i x(\mathbf{p})_{ij} = \sum_i w_{ij}, \ \forall j.$$

Example of Arrow-Debreu's Model

Traders $1, 2$ have good bundle

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Their optimization problems for given prices p_x, p_y are:

$$\begin{array}{rl} \max & 2x_1 + y_1 \\ \text{s.t.} & p_x \cdot x_1 + p_y \cdot y_1 \le p_x, \\ & x_1, y_1 \ge 0 \end{array} \qquad \begin{array}{rl} \max & 3x_2 + y_2 \\ \text{s.t.} & p_x \cdot x_2 + p_y \cdot y_2 \le p_y \\ & x_2, y_2 \ge 0. \end{array}$$

One can normalize the prices $\mathbf{p}$ such that one of them equals 1. This would be one of the problems in HW2.

Equilibrium conditions of the Arrow-Debreu market

Similarly, the necessary and sufficient equilibrium conditions of the Arrow-Debreu market are

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

where the budget for trader i is replaced by $\mathbf{p}^T\mathbf{w}_i$. Again, the nonlinear inequality can be rewritten as

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

Let $y_j = \log(p_j)$ or $p_j = e^{y_j}$ for all j. Then, these inequalities become

$$\log(\mathbf{u}_i^T\mathbf{x}_i) + y_j - \log(\sum_j w_{ij}e^{y_j}) \geq \log(u_{ij}),\ \forall i, j,\ u_{ij} > 0.$$

Note that the function on the left is concave in $\mathbf{x}_i$ and y_j.

Theorem 2 *The equilibrium set of the Arrow-Debreu Market is convex in allocations and the logarithmic of prices.*

Exchange Markets with Other Economies

Cobb-Douglas Utility:

$$u_i(\mathbf{x}_i) = \prod_{j \in G} x_{ij}^{u_{ij}}, \; x_{ij} \geq 0.$$

Leontief Utility:

$$u_i(\mathbf{x}_i) = \min_{j \in G} \{ \frac{x_{ij}}{u_{ij}}, \; x_{ij} \geq 0. \}.$$

Again, the equilibrium price vector is the one such that there are maximizers to clear the market.

Example of Geometric Optimization

Consider the Geometric Optimization Problem

$$\begin{array}{rl} \min_{\mathbf{x}} & \sum_{i=1}^{m} \left(a_i \prod_{j=1}^{n} x_j^{u_{ij}} \right) \\ \text{s.t.} & \prod_{j=1}^{n} x_j^{c_{kj}} = b_k,\ k = 1, ..., K \\ & x_j > 0,\ \forall j, \end{array}$$

where the coefficients $a_i \geq 0\ \forall i$ and $b_k > 0\ \forall k$.

$$\begin{array}{rl} \min_{x,y,z} & xy + yz + zx \\ \text{s.t.} & xyz = 1 \\ & (x, y, x) > \mathbf{0}. \end{array}$$

Convexification of Geometric Optimization

Let $y_j = \log(x_j)$ so that $x_j = e^{y_j}$. Then the problem becomes

$$\begin{array}{ll} \min_{\mathbf{x}} & \sum_{i=1}^{m} \left(a_i e^{\sum_{j=1}^{n} u_{ij} y_j} \right) \\ \text{s.t.} & \sum_{j=1}^{n} c_{kj} y_j = \log(b_k), \; k = 1, ..., K \\ & y_j \text{ free } \forall j. \end{array}$$

This is a convex objective function with linear constraints!

$$\begin{array}{ll} \min_{u,v,w} & e^{u+v} + e^{v+w} + e^{w+u} \\ \text{s.t.} & u + v + w = 0 \\ & (u, v, w) \text{ free.} \end{array}$$

Now the KKT solution suffices!

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 3 *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.

Proof

The proof reduces to one-dimensional case by considering the objective function $\phi(t) = f(\mathbf{x}(t))$ 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 $[-a\ a]$ for a sufficiently small $a > 0$, we must have $\phi'(0) = 0$ so that

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

Let all active constraints (including the equality ones) be $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ 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

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

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

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

Note that this inequality holds for every $\mathbf{d} \in T_{\bar{\mathbf{x}}}$.

Second-Order Sufficient Conditions for GCO

Theorem 4 *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.

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

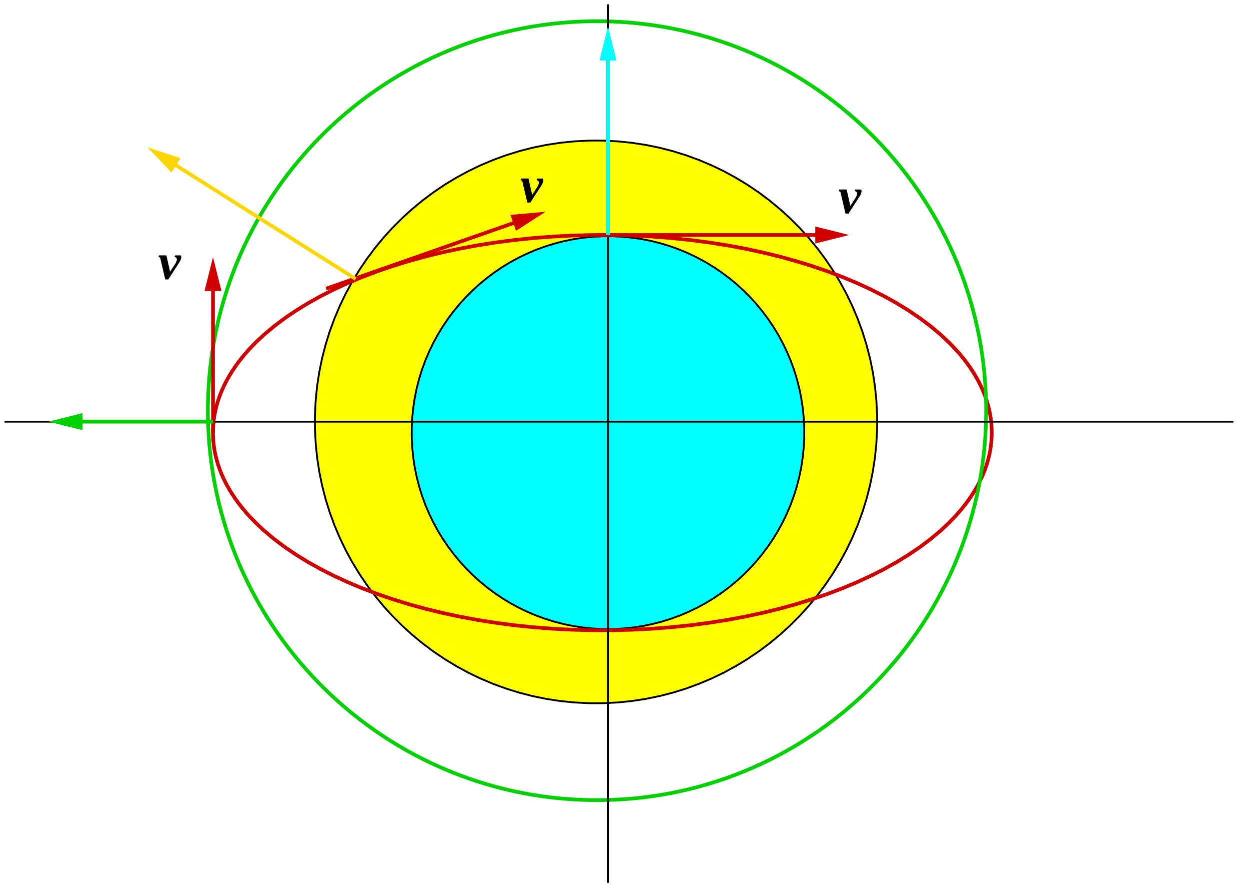

Figure 2: FONC and SONC for Constrained Minimization

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

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

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

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

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:

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

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

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

Consider the third KKT point:

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

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

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

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

Test Positive Semidefiniteness in a Subspace

In the second-order test, we typically like to know whether or not

$$\mathbf{d}^T Q\mathbf{d} \geq 0,\ \forall \mathbf{d},\ \text{s.t. } A\mathbf{d} = \mathbf{0}$$

for a given symmetric matrix Q and a rectangle matrix A. (In this case, the subspace is the null space of matrix A.) This test itself might be a nonconvex optimization problem.

But it is known that $\mathbf{d}$ is in the null space of matrix A if and only if

$$\mathbf{d} = (I - A^T(AA^T)^{-1}A)\mathbf{u} = P_A\mathbf{u}$$

for some vector $\mathbf{u} \in R^n$, where P_A is called the projection matrix of A. Thus, the test becomes whether or not

$$\mathbf{u}^T P_A Q P_A \mathbf{u} \geq 0,\ \forall \mathbf{u} \in R^n,$$

that is, we just need to test positive semidefiniteness of $P_A Q P_A$ as usual.

Spherical Constrained Nonconvex Quadratic Optimization

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

Theorem 5 *The FONC and SONC, that is, the following conditions on* $\mathbf{x}$*, together with the multiplier* y*,*

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

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