Modern Optimization Theory: 
Optimality Conditions and Perturbation Analysis

Defeng Sun
Department of Mathematics
National University of Singapore
Republic of Singapore

July 5-10, 2006
3 Perturbation Analysis

Let us consider

\[
\min_{x \in X} f(x) \quad \text{s.t.} \quad G(x) \in K,
\]

where \( f : X \to \mathbb{R} \) and \( G : X \to Y \) are \( C^2 \) (twice continuously differentiable), \( X, Y \) finite-dimensional real Hilbert vector spaces each equipped with a scalar product \( \langle \cdot, \cdot \rangle \) and its induced norm \( \| \cdot \| \), and \( K \) is a closed convex set in \( Y \).
The Lagrangian function $L : X \times Y \to \mathbb{R}$ for (OP) is defined by

$$L(x, \mu) := f(x) + \langle \mu, G(x) \rangle, \quad (x, \mu) \in X \times Y.$$ 

If $\bar{x}$ is a locally optimal solution to (OP) and the following Robinson’s CQ holds at $\bar{x}$:

$$0 \in \text{int}\{G(\bar{x}) + JG(\bar{x})X - K\},$$

(or $JG(\bar{x})X + TK(G(\bar{x})) = Y$),
then there exists a Lagrangian multiplier $\bar{\mu} \in Y$, together with $\bar{x}$, satisfying the KKT condition:

$$\nabla_x L(\bar{x}, \bar{\mu}) = 0 \quad \text{and} \quad \bar{\mu} \in N_K(G(\bar{x})),$$

(or

$$\nabla_x L(\bar{x}, \bar{\mu}) = 0 \quad \text{and} \quad G(\bar{x}) = \Pi_K(G(\bar{x}) + \bar{\mu})$$

and equivalently if $K$ is a closed convex cone

$$\nabla f(\bar{x}) + \nabla G(\bar{x})\bar{\mu} = 0 \quad \text{and} \quad K \ni G(\bar{x}) \perp (-\bar{\mu}) \in K^*.$$ 

Let $\mathcal{M}(\bar{x})$ denote the set of Lagrangian multipliers.
• Tremendous progress achieved in stability analysis in \((OP)\) subject to data perturbation.

• \(K\) is a polyhedral set, the theory quite complete. Especially for

\[
\begin{aligned}
\min_{x \in \mathbb{R}^n} & \quad f(x) \\
\text{s.t.} & \quad h(x) = 0, \\
& \quad g(x) \leq 0.
\end{aligned}
\]
For (NLP), Robinson’s CQ reduces to the Mangasarian-Fromovitz constraint qualification (MFCQ):

\[
\begin{cases}
Jh_i(\bar{x}), & i = 1, \ldots, m, \text{ are linearly independent,} \\
\exists d \in X : Jh_i(\bar{x})d = 0, i = 1, \ldots, m, Jg_j(\bar{x})d < 0, j \in \mathcal{I}(\bar{x}),
\end{cases}
\]

where

\[
\mathcal{I}(\bar{x}) := \{j : g_j(\bar{x}) = 0, j = 1, \ldots, p\}.
\]
A stronger notion than the MFCQ in (NLP) is the linear independence constraint qualification (LICQ):

\[ \{\mathcal{J}h_i(\bar{x})\}_{i=1}^m \text{ and } \{\mathcal{J}g_j(\bar{x})\}_{j\in\mathcal{I}(\bar{x})} \text{ are linearly independent.} \]

\(\mathcal{M}(\bar{x})\) is nonempty and bounded if and only if the MFCQ holds at \(\bar{x}\) while the LICQ implies that \(\mathcal{M}(\bar{x})\) is a singleton.
In 1980, Robinson\(^a\) introduced the far-reaching concept of **strong regularity** for generalized equations (KKT system is a special case) and **the strong second order sufficient condition (SSOSC)** for \((NLP)\) (the later is also developed by Luenberger\(^b\)).

Robinson proved for \((NLP)\):

\[
SSOSC + LICQ \implies \text{Strong Regularity.}
\]


\(^b\)D.G. Luenberger. *Introduction to Linear and Nonlinear Programming*, Addison-Wesley (London, 1973.)
Jongen, Mobert, Rückmann, and Tammer\textsuperscript{a}; Bonnans and Sulem\textsuperscript{b};
Dontchev and Rockafellar\textsuperscript{c} proved:

\textbf{SSOSC + LICQ} \iff \text{Strong Regularity.}


In the above characterizations, $K$ is a polyhedral set. Here we focus on the nonlinear semidefinite programming

$$(NLSDP)$$

$$\min_{x \in X} f(x)$$

s.t. $h(x) = 0,$

$g(x) \in S_+^p.$

Difficulty:

$S_+^p$ is not a polyhedral set.
Let $A \in S^p$ have the following spectral decomposition

$$A = P\Lambda P^T,$$

where $\Lambda$ is the diagonal matrix of eigenvalues of $A$ and $P$ is a corresponding orthogonal matrix of orthonormal eigenvectors. Then

$$A_+ := \Pi_{S_+^p}(A) = P\Lambda_+ P^T.$$
Define \( \alpha := \{ i : \lambda_i > 0 \} \), \( \beta := \{ i : \lambda_i = 0 \} \), \( \gamma := \{ i : \lambda_i < 0 \} \).

Write

\[
\Lambda = \begin{bmatrix}
\Lambda_\alpha & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & \Lambda_\gamma \\
\end{bmatrix}
\text{ and } P = \begin{bmatrix}
P_\alpha & P_\beta & P_\gamma \\
\end{bmatrix}.
\]

Define \( U \in \mathcal{S}^p \):

\[
U_{ij} := \frac{\max\{\lambda_i, 0\} + \max\{\lambda_j, 0\}}{|\lambda_i| + |\lambda_j|}, \quad i, j = 1, \ldots, p,
\]

where \( 0/0 \) is defined to be 1.
The tangent cone of $S^+_p$ at $A_+ = \Pi_{S^+_p}(A)$ is:

$$T_{S^+_p}(A_+) = \{ B \in S^p : P_{\bar{\alpha}}^T BP_{\bar{\alpha}} \succeq 0 \}.$$ 

and the lineality space of $T_{S^+_p}(A_+)$, i.e., the largest linear space in $T_{S^+_p}(A_+)$,

$$\text{lin} \left( T_{S^+_p}(A_+) \right) = \{ B \in S^n : P_{\bar{\alpha}}^T BP_{\bar{\alpha}} = 0 \},$$

where $\bar{\alpha} := \{1, \ldots, p\} \setminus \alpha$ and $P_{\bar{\alpha}} := [P_\beta \ P_\gamma]$.

The critical cone of $S_+^p$ at $A \in S^p$, is defined as

$$C(A; S_+^p) := T_{S_+^p}(A_+) \cap (A_+ - A)\perp,$$

$$= \left\{ B \in S^p : P^T_\beta BP_\beta \geq 0, \; P^T_\beta BP_\gamma = 0, \; P^T_\gamma BP_\gamma = 0 \right\}.$$

The affine hull of $C(A; S_+^p)$, $\text{aff}(C(A; S_+^p))$, can be written as

$$\text{aff} \left( C(A; S_+^p) \right) = \left\{ B \in S^p : P^T_\beta BP_\gamma = 0, \; P^T_\gamma BP_\gamma = 0 \right\}.$$
**Definition 3.1** For any $B \in S^p$, define the linear-quadratic function $\Upsilon_B : S^p \times S^p \to \mathbb{R}$ by

$$\Upsilon_B(\Gamma, A) := 2 \langle \Gamma, AB^\dagger A \rangle, \quad (\Gamma, A) \in S^p \times S^p,$$

where $B^\dagger$ is the Moore-Penrose pseudo-inverse of $B$.

**Proposition 3.1** Suppose that $B \in S^p_+$ and $\Gamma \in \mathcal{N}_{S^p_+}(B)$, i.e.,

$$B = \Pi_{S^p_+}(B + \Gamma).$$

Then for any $V \in \partial \Pi_{S^p_+}(B + \Gamma)$ and $\Delta B, \Delta \Gamma \in S^p$ such that $\Delta B = V(\Delta B + \Delta \Gamma)$, it holds that

$$\langle \Delta B, \Delta \Gamma \rangle \geq -\Upsilon_B(\Gamma, \Delta B).$$
Let $\bar{x}$ be a stationary point of $(NLSDP)$. Let $(\bar{\zeta}, \bar{\Gamma}) \in \mathcal{M}(\bar{x})$ such that

$$\nabla_x L(\bar{x}, \bar{\zeta}, \bar{\Gamma}) = 0, \quad -h(\bar{x}) = 0, \quad \text{and} \quad \bar{\Gamma} \in \mathcal{N}_{S^+}(g(\bar{x})).$$

Let $A := g(\bar{x}) + \bar{\Gamma}$ and\(^a\)

\[
g(\bar{x}) = P \begin{bmatrix} \Lambda_\alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} P^T, \quad \text{and} \quad \bar{\Gamma} = P \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Lambda_\gamma \end{bmatrix} P^T.
\]

\(^a\)Since $g(\bar{x})$ and $\bar{\Gamma}$ commute, we can simultaneously diagonalize them.
The critical cone $C(\bar{x})$ of $(NLSDP)$ at $\bar{x}$ is

$$C(\bar{x}) = \left\{ d : Jh(\bar{x})d = 0, Jg(\bar{x})d \in T_{S^+_x}(g(\bar{x})), Jf(\bar{x})d = 0 \right\}$$

$$= \left\{ d : Jh(\bar{x})d = 0, P^T_\beta (Jg(\bar{x})d)P_\beta \succeq 0, P^T_\beta (Jg(\bar{x})d)P_\gamma = 0, P^T_\gamma (Jg(\bar{x})d)P_\gamma = 0 \right\}.$$

The difficulty is that the affine hull of $C(\bar{x})$, $\text{aff}(C(\bar{x}))$, has no explicit formula.
Define the following outer approximation set to $\text{aff}(C(\bar{x}))$ with respect to $(\bar{\zeta}, \Gamma)$ by

$$\text{app}(\bar{\zeta}, \Gamma) := \left\{ d : \mathcal{J} h(\bar{x}) d = 0, \quad \mathcal{J} g(\bar{x}) d \in \text{aff} \left( C(A; S^p_+) \right) \right\}.$$ 

It holds that

$$\text{app}(\bar{\zeta}, \Gamma) = \left\{ d : \mathcal{J} h(\bar{x}) d = 0, \quad P^T_\beta (\mathcal{J} g(\bar{x}) d) P_\gamma = 0, \quad P^T_\gamma (\mathcal{J} g(\bar{x}) d) P_\gamma = 0 \right\}.$$
Then by the definition of \( \text{aff}(C(\bar{x})) \), we have for any \((\bar{\zeta}, \bar{\Gamma}) \in \mathcal{M}(\bar{x})\) that

\[
\text{aff}(C(\bar{x})) \subseteq \text{app}(\bar{\zeta}, \bar{\Gamma}).
\]

The two sets \( \text{aff}(C(\bar{x})) \) and \( \text{app}(\bar{\zeta}, \bar{\Gamma}) \) coincide if the strict complementary condition holds at \((\bar{x}, \bar{\zeta}, \bar{\Gamma})\):

\[
\text{rank}(g(\bar{x})) + \text{rank}(\bar{\Gamma}) = p,
\]

where “rank” denotes the rank of a square matrix.

In general, these two sets may be different even if \( \mathcal{M}(\bar{x}) \) is a singleton as in the case for \((NLP)\).
Proposition 3.2 Suppose that \((\tilde{\zeta}, \tilde{\Gamma})\) satisfies the following strict constraint qualification:

\[
\begin{pmatrix}
    \mathcal{J} h(\bar{x}) \\
    \mathcal{J} g(\bar{x})
\end{pmatrix} X + \begin{pmatrix}
    0 \\
    \mathcal{T}_{S^p}(g(\bar{x})) \cap \bar{\Gamma}^\perp
\end{pmatrix} = \begin{pmatrix}
    \mathbb{R}^m \\
    S^p
\end{pmatrix}.
\]

Then \(\mathcal{M}(\bar{x})\) is a singleton, i.e., \(\mathcal{M}(\bar{x}) = \{ (\tilde{\zeta}, \tilde{\Gamma}) \}\), and \(\text{aff}(C(\bar{x})) = \text{app}(\tilde{\zeta}, \tilde{\Gamma})\).
Recall that the “no-gap” second order necessary condition and the second order sufficient condition for \((NLSDP)\) can be stated as follows:

**Theorem 3.1** Let \(K = \{0\} \times \mathcal{S}_+^p \subset \mathbb{R}^m \times \mathcal{S}^p\). Suppose that \(\bar{x}\) is a locally optimal solution to \((NLSDP)\) and Robinson’s CQ holds at \(\bar{x}\). Then

\[
\sup_{\mu \in \mathcal{M}(\bar{x})} \left\{ \langle d, \nabla^2_{xx} L(\bar{x}, \mu)d \rangle - \sigma \left( \mu, T^2_K(G(\bar{x}), JG(\bar{x})d) \right) \right\} \geq 0
\]

for all \(d \in C(\bar{x})\).
(continued)

Conversely, let $\bar{x}$ be a feasible solution to \((\text{NLSDP})\) such that $\mathcal{M}(\bar{x})$ is nonempty. Suppose that Robinson's CQ holds at $\bar{x}$. Then the following condition

$$\sup_{\mu \in \mathcal{M}(\bar{x})} \left\{ \langle d, \nabla^2_{xx} L(\bar{x}, \mu) d \rangle - \sigma \left( \mu, T^2_K(G(\bar{x}), J G(\bar{x}) d) \right) \right\} > 0$$

for all $d \in C(\bar{x}) \setminus \{0\}$ is necessary and sufficient for the quadratic growth condition at the point $\bar{x}$:

$$f(x) \geq f(\bar{x}) + c \|x - \bar{x}\|^2 \quad \forall x \in \hat{N} \text{ such that } G(x) \in K$$

for some constant $c > 0$ and a neighborhood $\hat{N}$ of $\bar{x}$ in $X$. 
Proposition 3.3 Let $\bar{x}$ be a feasible solution to (NLSDP) such that $\mathcal{M}(\bar{x})$ is nonempty. Then for any $(\zeta, \Gamma) \in \mathcal{M}(\bar{x})$ with $\zeta \in \mathbb{R}^m$ and $\Gamma \in \mathcal{S}^p$, one has

$$
\Upsilon_{g(\bar{x})}(\Gamma, \mathcal{J}g(\bar{x})d) = \sigma \left( \Gamma, T^2_{\mathcal{S}^p_+}(g(\bar{x}), \mathcal{J}g(\bar{x})d) \right) \quad \forall d \in C(\bar{x}),
$$

where

$$
\Upsilon_B(\Gamma, A) = 2 \langle \Gamma, AB^\dagger A \rangle, \quad (\Gamma, A) \in \mathcal{S}^p \times \mathcal{S}^p.
$$
**Definition 3.2** Let $\bar{x}$ be a stationary point of (NLSDP). We say that the strong second order sufficient condition (SSOSC) holds at $\bar{x}$ if

$$\sup_{(\zeta, \Gamma) \in \mathcal{M}(\bar{x})} \{ \langle d, \nabla^2_{xx} L(\bar{x}, \zeta, \Gamma) \rangle - \Upsilon_{g(\bar{x})}(\Gamma, \mathcal{J}g(\bar{x})d) \} > 0$$

for all $d \in \hat{C}(\bar{x}) \setminus \{0\}$, where for any $(\zeta, \Gamma) \in \mathcal{M}(\bar{x})$, $(\zeta, \Gamma) \in \mathbb{R}^m \times \mathcal{S}^p$ and

$$\hat{C}(\bar{x}) := \bigcap_{(\zeta, \Gamma) \in \mathcal{M}(\bar{x})} \text{app}(\zeta, \Gamma).$$
Next, we define a **nondegeneracy condition** for \((NLSDP)\), which is an analogue of the **LICQ** for \((NLP)\). The concept of nondegeneracy originally appeared in Robinson\(^a\) for \((OP)\).

**Definition 3.3** We say that a feasible point \(\bar{x}\) to \((OP)\) is **constraint nondegenerate** if

\[
\mathcal{J}G(\bar{x})X + \text{lin}(\mathcal{T}_K(\bar{y})) = Y,
\]

where \(\bar{y} := G(\bar{x})\).

---

Write down the KKT condition as

\[
\begin{bmatrix}
\nabla L(x, \zeta, \Gamma) \\
-h(x) \\
-g(x) + \Pi_{S_p^+}(g(x) + \Gamma)
\end{bmatrix}
= \begin{bmatrix}
\nabla_x L(x, \zeta, \Gamma) \\
-h(x) \\
\Gamma - \Pi_{S_p^-}(\Gamma + g(x))
\end{bmatrix}
= 0,
\]

which is equivalent to the following generalized equation:

\[
0 \in \phi(z) + \mathcal{N}_D(z),
\]

where \( \phi \) is \( C^1 \) and \( D \) is a closed convex set in \( Z \).
Definition 3.4 [Robinson’80] Let $\bar{z}$ be a solution of the generalized equation. We say that $\bar{z}$ is a strongly regular solution if there exist neighborhoods $\mathcal{B}$ of the origin $0 \in Z$ and $\mathcal{V}$ of $\bar{z}$ such that for every $\delta \in \mathcal{B}$, the following linearized generalized equation

$$\delta \in \phi(\bar{z}) + J\phi(\bar{z})(z - \bar{z}) + \mathcal{N}_D(z)$$

has a unique solution in $\mathcal{V}$, denoted by $z_{\mathcal{V}}(\delta)$, and the mapping $z_{\mathcal{V}} : \mathcal{B} \to \mathcal{V}$ is Lipschitz continuous.
Let $U$ be a Banach space and $f : X \times U \to \mathbb{R}$ and $G : X \times U \to Y$.

We say that $(f(x,u), G(x,u))$, with $u \in U$, is a \textit{\textbf{C}^2\text{-smooth parameterization}} of $(OP)$ if $f(\cdot, \cdot)$ and $g(\cdot, \cdot)$ are $C^2$ and there exists a $\bar{u} \in U$ such that $f(\cdot, \bar{u}) = f(\cdot)$ and $G(\cdot, \bar{u}) = G(\cdot)$.

The corresponding parameterized problem takes the form:

\[ (OP_u) \]

\[
\begin{align*}
\min_{x \in X} & \quad f(x, u) \\
\text{s.t.} & \quad G(x, u) \in K.
\end{align*}
\]

We say that a parameterization is \textit{canonical} if $U := X \times Y$, $\bar{u} = (0, 0) \in X \times Y$, and

\[ (f(x,u), G(x,u)) := (f(x) - \langle u_1, x \rangle, G(x) + u_2), \quad x \in X. \]
Definition 3.5 [Bonnans and Shapiro’00] Let $\bar{x}$ be a stationary point of $(OP)$. We say that the \textit{uniform second order (quadratic) growth condition} holds at $\bar{x}$ with respect to a $C^2$-smooth parameterization $(f(x,u), G(x,u))$ if there exist $c > 0$ and neighborhoods $\mathcal{V}_X$ of $\bar{x}$ and $\mathcal{V}_U$ of $\bar{u}$ such that for any $u \in \mathcal{V}_U$ and any stationary point $x(u) \in \mathcal{V}_X$ of $(OP_u)$, the following holds:

$$f(x,u) \geq f(x(u),u) + c\|x-x(u)\|^2 \quad \forall x \in \mathcal{V}_X \text{ such that } G(x,u) \in K.$$ 

We say that the uniform second order growth condition holds at $\bar{x}$ if the above inequality holds for every $C^2$-smooth parameterization of $(OP)$. 
Definition 3.6 [Kojima\textsuperscript{a} and Bonnans and Shapiro’00]

Let $\tilde{x}$ be a stationary point of $(OP)$. We say that $\tilde{x}$ is strongly stable with respect to a $C^2$-smooth parameterization $(f(x,u), G(x,u))$ if there exist neighborhoods $\mathcal{V}_X$ of $\tilde{x}$ and $\mathcal{V}_U$ of $\tilde{u}$ such that for any $u \in \mathcal{V}_U$, $(OP_u)$ has a unique stationary point $x(u) \in \mathcal{V}_X$ and $x(\cdot)$ is continuous on $\mathcal{V}_U$.

If this holds for any $C^2$-smooth parameterization, we say that $\tilde{x}$ is strongly stable.

Let

$$\Phi(\delta) := F'(\bar{x}, \bar{\zeta}, \bar{\Gamma}; \delta).$$

Let \( \text{ind}(\phi, \bar{z}) \) denote the index of a continuous function \( \phi : \mathbb{Z} \to \mathbb{Z} \) at an isolated zero \( \bar{z} \in \mathbb{Z} \) used in degree theory.

Based on several recent results of Bonnans and Shapiro’00; Gowda\(^a\); Pang, Sun and Sun\(^b\); Sun and Sun’02, we get


Theorem 2\textsuperscript{a}. Let $\bar{x}$ be a locally optimal solution to \((NLSDP)\). Suppose that Robinson’s CQ holds at $\bar{x}$ so that $\bar{x}$ is necessarily a stationary point of \((NLSDP)\). Let $(\bar{\zeta}, \bar{\Gamma}) \in \mathbb{R}^m \times S^p$ be such that $(\bar{x}, \bar{\zeta}, \bar{\Gamma})$ is a KKT point of \((NLSDP)\). Then the following TEN statements are equivalent:

(a) The SSOSC holds at $\bar{x}$ and $\bar{x}$ is constraint nondegenerate.

(b) Any element in $\partial F(\bar{x}, \bar{\zeta}, \bar{\Gamma})$ is nonsingular.

(c) The KKT point $(\bar{x}, \bar{\zeta}, \bar{\Gamma})$ is strongly regular.

(d) The uniform second order growth condition holds at $\bar{x}$ and $\bar{x}$ is constraint nondegenerate.

(e) The point $\bar{x}$ is strongly stable and $\bar{x}$ is constraint nondegenerate.

(continued)

(f) $F$ is a locally Lipschitz homeomorphism near $(\bar{x}, \bar{\zeta}, \Gamma)$.

(g) For every $V \in \partial_B F(\bar{x}, \bar{\zeta}, \Gamma)$, $\text{sgn det}V = \text{ind}(F, (\bar{x}, \bar{\zeta}, \Gamma)) = \pm 1$.

(h) $\Phi$ is a globally Lipschitz homeomorphism.

(i) For every $V \in \partial_B \Phi(0)$, $\text{sgn det}V = \text{ind}(\Phi, 0) = \pm 1$.

(j) Any element in $\partial \Phi(0)$ is nonsingular.

Note that many more equivalent statements can be added by looking at statements (b) and (g).
Some unsolved problems:

(Q1) How far can we go beyond the SDP cone? Symmetric cone (SOC is fine)? Homogeneous cone? Hyperbolic cone?

(Q2) What can we say about the equivalent conditions in Theorem 2 if $\bar{x}$ is assumed to be a stationary point only? Or more generally

(Q3) How can we characterize the strong regularity for the conic complementarity problems?