University Paris-Saclay
Advanced Continuous Optimization I (J.Ch. Gilbert)
Exercises

The difficulty (or length?) of an exercise is sometimes loosely estimated by a number surrounded by braces at the beginning of its statement. The difficulty/length scale ranges from 1 to 5 (the higher the more difficult/long).

Background

0.1. {1} *Affine hull.* Let A be an affine subspace and $O \subset A$ be relatively open in A (i.e., open for the relative/induced/subspace topology of A). Then $\operatorname{aff} O = A$.

0.2. {1} *Relative interior*. Let C be a nonempty convex set of a finite dimensional vector space. Then

$$2\,(\operatorname{ri} C) = C + \operatorname{ri} C = \overline{C} + \operatorname{ri} C.$$

0.3. {3} *Decomposition in a subspace and a convex cone*. Let $\mathbb{E}_0$ be a subspace of a vector space $\mathbb{E}$ and K be a convex cone of $\mathbb{E}$. We denote by $\operatorname{vect} K$ the smallest vector space containing K (since K is a cone, $\operatorname{vect} K = \operatorname{aff} K$ actually). Then,

$$\mathbb{E}_0 + K = \mathbb{E} \qquad \Longleftrightarrow \qquad \begin{cases} \mathbb{E}_0 + \operatorname{vect} K = \mathbb{E} \\ \mathbb{E}_0 \cap (\operatorname{ri} K) \neq \varnothing. \end{cases}$$

1 Optimality conditions

First order optimality conditions for (P_G)

1.1. {1} *Convex* (P_{EI}). Let $G = \{0_{\mathbb{R}^{m_E}}\} \times \mathbb{R}^{m_I}_-$ and $T : \mathbb{R}^n \multimap \mathbb{R}^m : x \mapsto c(x) - G$ be the multifunction associated with the constraint of problem (P_{EI}). Show that T is convex if and only if c_E is affine and c_I is componentwise convex.

1.2. {1} *Examples of use of Robinson's constraint qualification condition.* Consider the following sets of the form $X_G := \{x \in \mathbb{E} : c(x) \in G\}$, with $G \subset \mathbb{F}$, and points $x_0 \in X_G$:

1. $\mathbb{E} = \mathbb{F} = \mathbb{R}^2$, and c and G are defined by

$$c(x) = (x_1^2 + (x_2-1)^2 - 1,\ x_1^2 + (x_2+1)^2 - 1), \qquad G = \mathbb{R}^2_-, \qquad \text{and} \qquad x_0 = (0,0),$$

2. $\mathbb{E} = \mathbb{F} = \mathbb{R}^2$, and c (2 identical constraints) and G are defined by

$$c(x) = (x_2,\ x_2), \qquad G = \mathbb{R}^2_+, \qquad \text{and} \qquad x_0 = (0,0).$$

For these sets X_G and points $x_0 \in X_G$,

(i) determine whether Robinson's constraint qualification condition holds at x_0, without using its equivalence with (CQ-MF),

(ii) find a mudulus of metric regularity of the multifunction $x \mapsto c(x) - G$ at $(x_0, 0)$ if any.

1.3. {1.5} *Robinson's constraint qualification conditions for* (P_{EI}). Viewing problem (P_{EI}) as a particular instance of problem (P_G), show that the Robinson constraint qualification condition is equivalent to the Mangasarian-Fromovitz constraint qualification condition.

1.4. {3} *First order optimality conditions with an additional set-inclusion constraint.* Let $\mathbb{E}$ and $\mathbb{F}$ be two Euclidean vector spaces. Consider the problem

$$\begin{cases} \min\ f(x) \\ x \in Q \\ c(x) \in G, \end{cases} \tag{1}$$

where $f : \mathbb{E} \to \mathbb{R}$ and $c : \mathbb{E} \to \mathbb{F}$ are differentiable functions, Q is a nonempty closed convex set of $\mathbb{E}$, and G is a nonempty closed convex set of $\mathbb{F}$. Show that if x_* is a local minimizer of problem (1) and if

$$0 \in \operatorname{int}\Big(c(x_*) + c'(x_*)(Q - x_*) - G\Big), \tag{2}$$

then, there exists $\lambda_* \in \mathbb{F}$ such that

$$\begin{cases} \nabla f(x_*) + c'(x_*)^* \lambda_* \in (\mathrm{T}_{x_*} Q)^+ \\ \lambda_* \in \mathrm{N}_{c(x_*)}\, G. \end{cases} \tag{3}$$

Second order optimality conditions

1.5. {2} *Example of use of the second order optimality conditions.* Consider the following nonlinear optimization problem in $x \in \mathbb{R}^2$:

$$\begin{cases} \min\ -\frac{1}{2}(x_1^2 + x_2^2) \\ x_2 \geqslant x_1^2 - 1 \\ x_1 \geqslant 0. \end{cases}$$

Using the Lagrangian $\ell : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ defined at (x, λ) by

$$\ell(x, \lambda) = -\tfrac{1}{2}(x_1^2 + x_2^2) + \lambda_1(x_1^2 - x_2 - 1) + \lambda_2(-x_1),$$

it can be shown that the first order optimality conditions are verified by the following primal-dual pairs

$$x = 0 \quad \text{and} \quad \lambda = 0, \tag{4a}$$

$$x = (0, -1) \quad \text{and} \quad \lambda = (1, 0), \tag{4b}$$

$$x = (\sqrt{2}/2, -1/2) \quad \text{and} \quad \lambda = (1/2, 0). \tag{4c}$$

Using the second order optimality conditions, determine *analytically* which of the points in (4) are (strict) local minimum, (strict) local maximum, or undetermined.

1.6. {3} *Sufficient second order optimality conditions for problem* (P_G). Let $\mathbb{E}$ and $\mathbb{F}$ be finite dimensional vector spaces. Consider the general optimization problem

$$(P_G) \qquad \begin{cases} \min f(x) \\ c(x) \in G, \end{cases}$$

where $f : \mathbb{E} \to \mathbb{R}$ and $c : \mathbb{E} \to \mathbb{F}$ are smooth functions, and G is a nonempty *closed convex* set in $\mathbb{F}$ (not necessarily a cone). Its feasible set is denoted by $X_G := \{x \in \mathbb{E} : c(x) \in G\}$. The *Lagrangian* of problem (P_G) is the function $\ell : \mathbb{E} \times \mathbb{F} \to \mathbb{R}$ defined at $(x, \lambda) \in \mathbb{E} \times \mathbb{F}$ by

$$\ell(x, \lambda) = f(x) + \langle \lambda, c(x) \rangle. \tag{5}$$

Let x_* be a local minimum of (P_G) and assume that there exists a multiplier $\lambda_* \in \mathbb{E}$ such that

$$\nabla_x \ell(x_*, \lambda_*) = 0 \qquad \text{and} \qquad \lambda_* \in \mathrm{N}_{c(x_*)} G. \tag{6}$$

We denote by Λ_* the set of optimal multipliers λ_* such that (x_*, λ_*) verifies (6), which is therefore nonempty, and by $L_* := \nabla^2_{xx} \ell(x_*, \lambda_*)$ the Hessian of the Lagrangian with respect to x at a pair (x_*, λ_*) defined in the context where L_* appears. The *critical cone* at a local solution x_* to problem (P_G) is defined by

$$C_* := \{d \in \mathbb{E} : c'(x_*)d \in \mathrm{T}_{c(x_*)} G,\ f'(x_*)d \leqslant 0\}. \tag{7}$$

1) Show that, when x_* satisfies the first order optimality conditions (6), the critical cone also reads

$$C_* := \{d \in \mathbb{E} : c'(x_*)d \in \mathrm{T}_{c(x_*)} G,\ f'(x_*)d = 0\}. \tag{8}$$

2) Show that the following two properties are equivalent ($\|\cdot\|$ is an arbitrary norm):

$$\forall d \in C_* \smallsetminus \{0\},\ \exists\, \lambda_* \in \Lambda_* : \quad \langle L_* d, d \rangle > 0, \tag{9a}$$

$$\exists\, \bar{\gamma} > 0,\ \forall d \in C_*,\ \exists\, \lambda_* \in \Lambda_* : \quad \langle L_* d, d \rangle \geqslant \bar{\gamma} \|d\|^2. \tag{9b}$$

3) Assume that (9b) holds. Show that for all $\gamma \in [0, \bar{\gamma})$, there exists a neighborhood V of x_* such that

$$\forall\, x \in (X_G \cap V) \smallsetminus \{x_*\} : \qquad f(x) > f(x_*) + \frac{\gamma}{2} \|x - x_*\|^2. \tag{10}$$

The *quadratic growth property* (10) implies that x_* is a strict local minimum of (P_G).

2 Linearization methods

The Josephy-Newton algorithm for functional inclusions

2.1. {1} *Explicit variational problem.* Let $\mathbb{E}$ be a Euclidean space, X be a (not necessarily convex) subset of $\mathbb{E}$, and $F : \mathbb{E} \to \mathbb{E}$ be a smooth map. Consider the variational problem

$$(P_{\mathrm{V}}) \quad F(x) + \mathrm{N}_X(x) \ni 0,$$

where $\mathrm{N}_X(x)$ is the normal cone to X at x. Suppose that $\mathbb{F}$ is another Euclidean space and that X has actually the following form

$$X := \{x \in \mathbb{E} : c(x) \in G\},$$

in which $c : \mathbb{E} \to \mathbb{F}$ is a smooth function and G is a closed convex set of $\mathbb{F}$. Show that if x_* is a solution to (P_{V}) satisfying

$$0 \in \operatorname{int}\{c(x_*) + c'(x_*)\mathbb{E} - G\},$$

then, there exists $\lambda_* \in \mathrm{N}_{c(x_*)}\,G$ such that

$$F(x_*) + c'(x_*)^* \lambda_* = 0.$$

2.2. {2} *Josephy-Newton algorithm for a nonlinear complementarity problem.* Let $\mathbb{E}$ and $\mathbb{F}$ be two Euclidean spaces, K be a nonempty closed convex cone of $\mathbb{F}$, K^+ its positive dual, and F and $G : \mathbb{E} \to \mathbb{F}$ be two differentiable functions. Consider the nonlinear complementarity problem

$$K \ni G(x) \perp F(x) \in K^+. \tag{11}$$

Show that the algorithm that computes the next iterate x_{k+1} from the current one x_k by solving the linear complementarity problem in x

$$K \ni \Big(G(x_k) + G'(x_k)(x - x_k)\Big) \perp \Big(F(x_k) + F'(x_k)(x - x_k)\Big) \in K^+, \tag{12}$$

can be viewed as the Josephy-Newton algorithm on a certain variational inequality problem on a cone; which one?

The SQP algorithm in constrained optimization

2.3. *An SQP Algorithm for Solving Problem* (P_G). Consider the optimization problem (P_G) with its Lagrangian $\ell : \mathbb{E} \times \mathbb{F} \to \mathbb{R}$ defined at (x, λ) by

$$\ell(x, \lambda) = f(x) + \langle \lambda, c(x) \rangle.$$

For a given pair $(x_*, \lambda_*) \in \mathbb{E} \times \mathbb{F}$, denote by $L_* := \nabla^2_{xx}\ell(x_*, \lambda_*)$ the Hessian of the Lagrangian at (x_*, λ_*). Let us introduce the set

$$C_* := \{d \in \mathbb{E} : c'(x_*)d \in \mathrm{T}_{c(x_*)}\,G,\ f'(x_*)d \leqslant 0\}.$$

1) Show that the first order optimality conditions at a solution x_* to problem (P_G) can be written

$$F(z_*) + \mathrm{N}_C(z_*) \ni 0, \tag{13}$$

for some variable z_*, some function F, and some closed convex set C to determine.

2) Show in what sense the Josephy-Newton (JN) algorithm on the functional inclusion (13) consists in determining the new iterate $x + d$ from the current one x by computing d as a *stationary point* of

$$\begin{cases} \min_d \langle \nabla f(x), d \rangle + \frac{1}{2} \langle \nabla^2_{xx}\ell(x, \lambda)d, d \rangle \\ c(x) + c'(x) \cdot d \in G. \end{cases} \tag{14}$$

3) In this item, we assume that G is a *convex polyhedron*.

Show that $(x_*, \lambda_*) \in \mathbb{E}\times\mathbb{F}$ is a semi-stable primal-dual solution to (P_G) if and only if $\langle L_* d, d\rangle > 0$ for all pairs $(d,\mu) \in \mathbb{E}\times\mathbb{F}$ satisfying

$$(d, c'(x_*)d, \mu) \neq 0, \tag{15a}$$
$$c(x_*) + c'(x_*)d \in G, \tag{15b}$$
$$\langle \lambda_*, c'(x_*)d\rangle = 0, \tag{15c}$$
$$L_* d + c'(x_*)^* \mu = 0, \tag{15d}$$
$$\lambda_* + \mu \in \mathrm{N}_G(c(x_*)). \tag{15e}$$

2.4. *Nondifferentiable augmented Lagrangian.* Let $\mathbb{E}$ be a Euclidean vector space. Consider the usual optimization problem with equality and inequality constraints

$$(P_{EI}) \qquad \begin{cases} \inf_x \ f(x) \\ c_E(x) = 0 \\ c_I(x) \leqslant 0, \end{cases}$$

where $f : \mathbb{E} \to \mathbb{R}$, E and I form a partition of $[1\!:\!m]$, $m_E := |E|$, $m_I := |I|$, $c_E : \mathbb{E} \to \mathbb{R}^{m_E}$, and $c_I : \mathbb{E} \to \mathbb{R}^{m_I}$. The functions f and c are supposed smooth. The lagrangian of the problem is the function $\ell : \mathbb{E}\times\mathbb{R}^m \to \mathbb{R}$ defined at $(x,\lambda) \in \mathbb{E}\times\mathbb{R}^m$ by

$$\ell(x,\lambda) = f(x) + \lambda^{\mathsf{T}} c(x),$$

where $c(x) := (c_E(x), c_I(x))$.

Let x_* be a local minimum of problem (P_{EI}) at which the KKT conditions hold. We denote by Λ_* the set of optimal multipliers associated with x_*. Suppose that the weak second-order sufficient condition of optimality holds at x_*.

Let $\|\cdot\|_{\mathrm{P}}$ be a norm on $\mathbb{R}^m$ and $\|\cdot\|_{\mathrm{D}}$ be its dual norm with respect to the Euclidean scalar product. Let $\mu \in \mathbb{R}^m$ and $\sigma \in \mathbb{R}_+$ verifying

$$\sigma \geqslant \sup_{\lambda_* \in \Lambda_*} \|\lambda_* - \mu\|_{\mathrm{D}} \qquad \text{and} \qquad \sigma > \|\hat\lambda_* - \mu\|_{\mathrm{D}}, \tag{16}$$

for some $\hat\lambda_* \in \Lambda_*$. We want to show that the function $\Theta_{\mu,\sigma} : \mathbb{E} \to \mathbb{R}$ defined at $x \in \mathbb{E}$ by

$$\Theta_{\mu,\sigma}(x) := f(x) + \mu^{\mathsf{T}} c(x)^{\#} + \sigma \|c(x)^{\#}\|_{\mathrm{P}}$$

has a strict local minimum at x_*. We propose a reasoning by contradiction.

1) Show that if $\Theta_{\mu,\sigma}$ has not a strict local minimum at x_*, one can find a sequence $\{x_k\} \subset \mathbb{E}$, a sequence of positive real numbers $\{t_k\} \downarrow 0$, and a nonzero critical direction d such that $x_k = x_* + t_k d + o(t_k)$.
2) Show that

$$\exists\, \lambda_* \in \Lambda_*, \quad \forall\, k \geqslant 1 : \quad \ell(x_*, \lambda_*) < \ell(x_k, \lambda_*).$$

3) Get a contradiction.

2.5. *Norm assumptions.* For an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^m$, show that the following properties are equivalent (the operators $|\cdot|$ and $(\cdot)^+$ act componentwise):

(i) $0 \leqslant u \leqslant v \;\Rightarrow\; \|u\| \leqslant \|v\|$,

(ii) $u \leqslant v \;\Rightarrow\; \|u^+\| \leqslant \|v^+\|$,

(iii) $v \mapsto \|v^+\|$ is convex.

The semismooth Newton algorithm for equations

2.6. *Upper semi-continuity of the subdifferential.* The subdifferential of a convex function $f : \mathbb{E} \to \mathbb{R}$ at $x \in \mathbb{E}$ is the set

$$\partial f(x) \coloneqq \{s \in \mathbb{E} : f(y) \geqslant f(x) + \langle s, y - x\rangle,\ \forall\, y \in \mathbb{E}\}.$$

Show that the multifunction $\partial f : \mathbb{E} \multimap \mathbb{E} : x \mapsto \partial f(x)$ is upper-semicontinuous at any $x \in \mathbb{E}$.

3 Conic Optimization

Semidefinite Optimization

3.1. *Singular Schur complement.* Let $A \in \mathcal{S}^n$, $C \in \mathcal{S}^m$, and $B \in \mathbb{R}^{n\times m}$. Denote by $A^\dagger$ the pseudo-inverse of A. Show that

$$\begin{pmatrix} A & B \\ B^\mathsf{T} & C \end{pmatrix} \succcurlyeq 0 \qquad \Longleftrightarrow \qquad \begin{cases} A \succcurlyeq 0 \\ C - B^\mathsf{T} A^\dagger B \succcurlyeq 0 \\ \mathcal{R}(B) \subset \mathcal{R}(A). \end{cases} \tag{17}$$

3.2. *Existence of dual solutions in SDO.* We use the notation of the course. Consider the standard primal form of the semidefinite optimization problem

$$\begin{cases} \inf_{X \in \mathcal{S}^n} \langle C, X\rangle \\ \mathcal{A}(X) = b \\ X \succcurlyeq 0, \end{cases}$$

Show that

$$\mathcal{F}_{\mathrm{P}}^s \times \mathcal{F}_{\mathrm{D}} \neq \varnothing \qquad \Longrightarrow \qquad \begin{cases} \mathrm{Sol}(D) \cap (\mathcal{R}(A) \times \mathcal{S}^n) \text{ is nonempty and compact,} \\ \text{there is no duality gap: } \mathrm{val}(D) = \mathrm{val}(P). \end{cases}$$