FIRST ORDER DEPENDENCE ON UNCERTAINTY SETS IN ROBUST OPTIMIZATION

C.H. JEFFREY PANG

Abstract. We show that a first order problem can approximate solutions of a robust optimization problem when the uncertainty set is scaled, and explore further properties of this first order problem.

Contents

1. Introduction

Robust optimization is the methodology of handling optimization problems with uncertain data. In practice, the presence of uncertainties in optimization problems can make nominal solutions meaningless. Such uncertainties can come from data uncertainty in measurement and estimation, or from uncertainty in implementation. We refer to the recent text [3] for more details.

Consider the linear program:

$$\begin{array}{ll} \min_x & \bar{c}^T x + \bar{d} \\ \text{s.t.} & \bar{A}x \leq \bar{b}. \end{array}$$

To account for the uncertainties in the data $(\bar{A}, \bar{b})$, one instead considers a point x to be feasible if it satisfies

$$Ax \leq b \text{ for all } (A, b) \in \mathcal{U}.$$

Here, $\mathcal{U}$ is a set containing the nominal data $(\bar{A}, \bar{b})$. We can consider the translation $\Delta\mathcal{U} = \mathcal{U} - (\bar{A}, \bar{b})$ and ask: What is the behavior of optimal solutions to the robust optimization problem if the set $\Delta\mathcal{U}$ were to be scaled by some factor ϵ? A large value of ϵ corresponds to a more robust solution, and a small value of ϵ places

Date: June 9, 2010.
2000 *Mathematics Subject Classification.* 90C31, 93D09, 49J53.
Key words and phrases. Robust optimization, sensitivity, uncertainty sets, variational analysis.

more importance in the objective function. Understanding the dependence of ϵ allows one to find a balance between optimization and robustness. The first order dependence on ϵ is addressed in Corollary 2.4 for linear programs and Theorem 5.3 for nonlinear programs.

The outline of this paper is as follows. We introduce robust linear programming in Section 2. Before we introduce robust nonlinear programming in Section 4, we recall some topics in variational analysis (or nonsmooth analysis) as presented in the texts [9, 5, 4] in Section 3. In Section 4, we also define the tangential problem, which will be important in Theorem 5.3, our main result. We present first order properties of the tangential problem in Section 6, and study the effects of sums of uncertainty sets in the tangential problem in Section 7.

2. Robust linear programming

We keep our presentation compatible with [3], and begin with the definition of the robust counterpart of a linear program.

Definition 2.1. (Robust counterpart) For $\bar{A} \in \mathbb{R}^{m\times n}$, $\bar{b} \in \mathbb{R}^m$, $\bar{c} \in \mathbb{R}^n$ and $\bar{d} \in \mathbb{R}$, where $m \geq n$, consider the linear program with parameters $(\bar{A}, \bar{b}, \bar{c}, \bar{d})$

$$\begin{aligned} &\min_x \bar{c}^T x + \bar{d} \\ &\text{s.t. } \bar{A}x \leq \bar{b}. \end{aligned} \tag{2.1}$$

The *robust counterpart* (written as RC) of the above linear program is

$$\begin{aligned} &\min_x \left\{ \hat{c}(x) = \sup_{(A,b,c,d)\in\mathcal{U}} [c^T x + d] \mid Ax \leq b \text{ for all } (A,b,c,d) \in \mathcal{U} \right\} \\ = \; &\min_{x,t} \{ t \mid t \geq c^T x + d,\ Ax \leq b \text{ for all } (A,b,c,d) \in \mathcal{U} \}, \end{aligned}$$

where $\mathcal{U}$ is an uncertainty set for the parameters (A, b, c, d), with $(\bar{A}, \bar{b}, \bar{c}, \bar{d}) \in \mathcal{U}$.

In a typical linear program, the variable d does not affect the minimizer, but one has to take perturbations in d into account in a robust optimization problem. The second formulation in the RC shows that we can rewrite the linear program so that c stays constant at $\bar{c}$ and $d = 0$. This is the approach we will take for the rest of this section, and we define $\mathcal{U}$ to be a set containing elements of the form (A, b), where (A, b) are close enough to $(\bar{A}, \bar{b})$. For more details, we refer to [3].

We define ΔA, Δb and the set $\Delta\mathcal{U}$ by the relations

$$\begin{aligned} \Delta A &:= A - \bar{A}, \\ \Delta b &:= b - \bar{b}, \\ \text{and } \Delta\mathcal{U} &:= \mathcal{U} - (\bar{A}, \bar{b}). \end{aligned}$$

The vector x can be chosen so that it stays feasible under these first order perturbations. We write $x = \bar{x} + \Delta x$. The RC is therefore simplified to

$$\begin{aligned} &\min_{\Delta x} c^T(\bar{x} + \Delta x) + d \\ &\text{s.t. } [\bar{A}_i + \Delta A_i](\bar{x} + \Delta x) \leq [\bar{b}_i + \Delta b_i] \text{ for all } (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i, \end{aligned} \tag{2.2}$$

where $\Delta\mathcal{U}_i$ is the uncertainty in the ith row.

When $\Delta\mathcal{U}$ is a small set , we seek to use a first order approximation to determine a robustly feasible x. Letting $x = \bar{x} + \Delta x$, and removing the second order term $(\Delta A_i)(\Delta x)$ in (2.2) gives

$$\begin{aligned} \bar{A}_i(\bar{x}+\Delta x) + (\Delta A_i)\bar{x} - (\bar{b}_i + \Delta b_i) &\leq 0 \text{ for all } (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i, \\ \text{or } [\bar{A}_i\bar{x} - \bar{b}] + \bar{A}_i(\Delta x) + [(\Delta A_i)\bar{x} - \Delta b_i] &\leq 0 \text{ for all } (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i. \end{aligned} \tag{2.3}$$

If $\bar{A}_i\bar{x} - \bar{b}_i < 0$ and if $\Delta\mathcal{U}$ were small enough, this constraint will not be tight in the optimization problem. With these in mind, we define the first order problem of a linear program.

Definition 2.2. (First order problem) Let $\bar{x}$ be an optimal solution to (2.1). The *first order problem* is the problem

$$\begin{aligned} &\min_{\gamma} c^T\gamma \\ &\text{s.t. } \bar{A}_i\gamma + [(\Delta A_i)\bar{x} - \Delta b] \leq 0 \\ &\ \text{for all } (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i \text{ whenever } \bar{A}_i\bar{x} - \bar{b}_i = 0. \end{aligned} \tag{2.4}$$

The first order problem can also be written as

$$\begin{aligned} &\min_{\gamma} c^T\gamma \\ &\text{s.t. } \bar{A}_i\gamma \leq - \max_{(\Delta A, \Delta b)\in\Delta\mathcal{U}} [(\Delta A_i)\bar{x} - \Delta b_i] \\ &\ \text{whenever } \bar{A}_i\bar{x} - \bar{b}_i = 0. \end{aligned}$$

In the case where the optimal solution $\bar{x}$ is nondegenerate, i.e., when $B = \{i \mid \bar{A}_i\bar{x} = \bar{b}_i\}$ is of size n and $\bar{A}_B$ is invertible, the optimal solution $\bar{\gamma}$ of the first order problem is just $\bar{\gamma} = \bar{A}_B^{-1}w$, where w is the vector

$$w = \begin{pmatrix} -\max_{(\Delta A, \Delta b)\in\Delta\mathcal{U}}[(\Delta A_{i_1})\bar{x} - \Delta b_{i_1}] \\ \vdots \\ -\max_{(\Delta A, \Delta b)\in\Delta\mathcal{U}}[(\Delta A_{i_n})\bar{x} - \Delta b_{i_n}] \end{pmatrix}, \tag{2.5}$$

where $i_1, \dots i_n$ are the n elements in B. When $\bar{x}$ is a degenerate solution, the first order problem is still easy to solve. We illustrate with a particular example that the tangential constraints are easily obtained for rectangular uncertainty sets.

Example 2.3. (Rectangular uncertainty) Suppose that the uncertainty set $\Delta\mathcal{U}$ is rectangular, that is

$$\begin{aligned} \Delta\mathcal{U} \ := \ &\big\{(\Delta A, \Delta b) : |\Delta A_{j,k}| \leq \epsilon_{j,k},\ |\Delta b_j| \leq \delta_j \\ &\quad \text{for all } j \in \{1,\dots,m\},\ k \in \{1,\dots,n\}\big\}. \end{aligned}$$

Then for each $i \in B$,

$$\max\{(\Delta A_i)\bar{x} - \Delta b_i \mid (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i\} = \delta_i + \sum_{k=1}^{n} \epsilon_{i,k}|\bar{x}_k|.$$

In Theorem 5.3, we will discuss how an adapted first order problem gives a first order approximation of the solution to a robust optimization problem in a general setting of nonlinear programs. For now, we shall present the corollary in the simpler setting of linear programming.

Corollary 2.4. *(to Theorem 5.3) (First order approximation in linear programming) Consider the robust optimization problem*

$$\begin{aligned} &\min_x c^T x + d \\ &\text{s.t. } (\bar{A}_i + \Delta A_i)x \leq (\bar{b}_i + \Delta b_i) \text{ for all } (\Delta A_i, \Delta b_i) \in \epsilon \Delta \mathcal{U}_i \text{ for all } i, \end{aligned} \tag{2.6}$$

and the first order problem

$$\begin{aligned} &\min_\gamma c^T \gamma \\ &\text{s.t. } \bar{A}_i \gamma \leq - \max_{(\Delta A_i, \Delta b_i) \in \Delta \mathcal{U}_i} [(\Delta A_i)\bar{x} - \Delta b_i] \\ &\qquad \text{for all } i \text{ s.t. } \bar{A}_i \bar{x} = \bar{b}_i. \end{aligned} \tag{2.7}$$

Let $\bar{\Gamma}$ be the set of optimal solutions to (2.7)*. Suppose*

1. *$\Delta\mathcal{U}_i$ are compact convex sets.*
2. *$\bar{\Gamma}$ is bounded.*
3. *There is some γ' such that $\bar{A}_i \gamma' < 0$ whenever $\bar{A}_i \bar{x} = \bar{b}_i$.*
4. *$\bar{x}$ is the unique minimizer of the nominal problem $\min\{c^T x \mid \bar{A}x \leq \bar{b}\}$.*

Then the set of cluster points of any sequence $\{\frac{1}{\epsilon}(\bar{x}_\epsilon - \bar{x})\}$, where $\bar{x}_\epsilon$ is an optimal solution to (2.6) *and $\epsilon \to 0$, is a subset of $\bar{\Gamma}$. The objective value of* (2.6) *, say $\bar{v}_\epsilon$, has an approximation $\bar{v}_\epsilon = \bar{v} + \epsilon \tilde{v} + o(\epsilon)$, where $\tilde{v}$ is the objective value of* (2.7)*.*

In particular, if $\bar{\Gamma}$ contains only one element, say $\bar{\gamma}$, then $\lim_{\epsilon \to 0} \frac{1}{\epsilon}(\bar{x}_\epsilon - \bar{x}) = \bar{\gamma}$, or $\bar{x}_\epsilon \in \bar{x} + \epsilon\bar{\gamma} + o(\epsilon)$.

Proof. The condition that Q is Clarke regular at $\bar{A}\bar{x} - \bar{b}$ holds in this case because $\mathbb{R}^m_-$ is Clarke regular everywhere. The condition that $\bar{x}$ is the unique minimizer in (4) suffices because the domain is convex. The affine function $x \mapsto c^T x + d$ is locally Lipschitz and subdifferentially regular everywhere. □

3. Preliminaries in variational analysis

In this section, we recall the definitions of some nonsmooth objects in variational analysis that will be necessary for the rest of the paper. We recall the definition of normal cones and Clarke regularity.

Definition 3.1. (Normal cones and Clarke regularity) Let $C \subset \mathbb{R}^n$. For a point $\bar{x} \in C$, a vector v is *normal to C at $\bar{x}$ in the regular sense*, or a *regular normal*, written $v \in \hat{N}_C(\bar{x})$, if

$$v^T(x - \bar{x}) \leq o(|x - \bar{x}|) \text{ for all } x \in C.$$

It is *normal to C in the general sense*, or simply a *normal vector*, written $v \in N_C(\bar{x})$, if there are sequences $x_i \to \bar{x}$ and $v_i \to v$ with $v_i \in \hat{N}_C(x_i)$. The set C is *Clarke regular* at $\bar{x}$ if $N_C(\bar{x}) = \hat{N}_C(\bar{x})$.

We refer the reader to [9, Corollary 6.29] for equivalent definitions of Clarke regularity. The sets we will encounter in this paper are all Clarke regular, so this does not cause difficulties.

We recall the definition of the tangent cone, which will be important in our main result.

Definition 3.2. (Tangent cones) The *tangent cone* of a set $C \subset \mathbb{R}^m$ at some $\bar{x} \in C$ is defined by

$$T_C(\bar{x}) := \left\{ w \mid \frac{x_i - \bar{x}}{t_i} \to w \text{ for some } x_i \in C,\, t_i \searrow 0 \text{ and } x_i \to \bar{x} \right\}.$$

Next, we recall sublinearity and equivalent definitions of subdifferential regularity that will also be useful for our main result. We take the definitions of subdifferential regularity from [9, Definition 7.25, Exercise 9.15, Corollary 8.19].

Definition 3.3. (positive homogeneity and sublinearity) A function $h : \mathbb{R}^n \to \mathbb{R}$ is *positively homogeneous* if $h(\lambda x) = \lambda h(x)$ for all x and $\lambda > 0$. It is *sublinear* if in addition

$$h(x + x') \leq h(x) + h(x') \text{ for all } x \text{ and } x'.$$

It is clear that sublinear functions are convex.

Definition 3.4. (Subdifferential regularity) Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz at $\bar{x}$.

(a) We say that the function f is *(subdifferentially) regular at* $\bar{x}$ if the epigraph $\text{epi} f := \{(x,t) \mid t \geq f(x)\}$ is Clarke regular at $(x, f(x))$ as a subset of $\mathbb{R}^n \times \mathbb{R}$.

(b) Define the *subderivative* $df(\bar{x}) : \mathbb{R}^n \to \mathbb{R}$ by

$$df(\bar{x})(w) := \liminf_{\tau \searrow 0} \frac{f(\bar{x} + \tau w) - f(\bar{x})}{\tau}. \tag{3.1}$$

and the *regular subderivative* $\hat{d}f(\bar{x}) : \mathbb{R}^n \to \mathbb{R}$ by

$$\hat{d}f(\bar{x})(w) := \limsup_{\substack{\tau \searrow 0 \\ x \to \bar{x}}} \frac{f(x + \tau w) - f(x)}{\tau}.$$

In general, the regular subderivative is sublinear. The function f is (subdifferentially) regular at $\bar{x}$ if and only if $df(\bar{x}) = \hat{d}f(\bar{x})$. Under subdifferential regularity, it is clear that the liminf in (3.1) can be taken to be a full limit. Also, $T_{\text{epi}(f)}(\bar{x}, f(\bar{x})) = \text{epi}(df(\bar{x}))$.

Since the tangent cone will play a major role in our main result, we now recall some calculus rules for tangent cones, highlighting a constraint qualification condition similar to that of condition (4) in Theorem 5.3. The rest of this section will not be essential to the development of the paper, so one may skip to the next section in a first reading. We now recall a formula for tangent cones under intersections.

Proposition 3.5. *(Tangent cones to intersections) Let* $C = C_1 \cap \cdots \cap C_m$ *for closed sets* $C_i \subset \mathbb{R}^n$*, and let* $\bar{x} \in C$*. Suppose* $\bar{x}$ *is Clarke regular at* C_j *for all* j*. Assume either*

$$\begin{aligned} &\sum_{j=1}^{m} \lambda_j v_j = 0,\ v_j \in N_{C_j}(\bar{x}) \text{ and } \lambda_j \geq 0 \text{ for all } j \in \{1, \ldots, m\} \\ &\text{implies } \lambda_j = 0 \text{ for all } j \in J', \end{aligned} \tag{3.2}$$

or equivalently:

(a) *there are no vectors* $\{y_j\}_{j=1}^m$ *such that* $y_j \perp T_{C_j}(\bar{x})$ *and* $y_1 + \cdots + y_m = 0$ *other than* $y_j = 0$ *for all* $j \in \{1, \ldots, m\}$*, and there is a vector* w *such that* $w \in \mathbb{R}^n \backslash \{0\}$ *such that* $w \in \text{rint}(T_{C_j}(\bar{x}))$ *for all* $j \in \{1, \ldots, m\}$*.*

Then one has

$$T_C(\bar{x}) = T_{C_1}(\bar{x}) \cap \cdots \cap T_{C_m}(\bar{x}),$$

and C is Clarke regular at $\bar{x}$.

Proof. Other than the equivalence of (3.2) and (a), this result is stated in a more general case in [9, Theorem 6.42]. This result is obtained by consider the set $D := C_1 \times \cdots \times C_2 \subset (\mathbb{R}^n)^m$ and the mapping $F : x \mapsto (x, \dots, x) \in (\mathbb{R}^n)^m$ with $X = \mathbb{R}^n$ and applying [9, Theorems 6.31 and 6.41]. The constraint qualification condition required is (3.2). By [9, Exercise 6.39(b)], (3.2) is equivalent to the existence of a w' such that $F(w') \in \operatorname{rint}(T_D(\bar{x}, \dots, \bar{x}))$ and having

$$y \perp T_D(\bar{x}, \dots, \bar{x}) \text{ and } F^*(y) = 0 \text{ implies } y = 0.$$

These conditions are equivalent to that in (a). □

We recall the Mangasarian-Fromovitz constraint qualification.

Definition 3.6. (Mangasarian-Fromovitz constraint qualification) For $\mathcal{C}^1$ functions $f_j : \mathbb{R}^n \to \mathbb{R}$ and $j \in \{1, \dots, m\}$, let

$$Q := \big\{x \in \mathbb{R}^n \mid f_j(x) \leq 0 \text{ for all } j \in \{1, \dots, m\}\big\}.$$

For $\bar{x} \in C$, let $J' := \{j \mid f_j(\bar{x}) = 0\}$. The *Mangasarian-Fromovitz constraint qualification* (MFCQ) is satisfied at $\bar{x}$ if there is a vector $w \in \mathbb{R}^n$ such that

$$\nabla f_j(\bar{x})^T w < 0 \text{ for all } j \in J'.$$

Another equivalent definition of the MFCQ is the following "positive linear independence" condition

$$\sum_{j \in J'} \lambda_j \nabla f_j(\bar{x}) = 0 \text{ and } \lambda_j \geq 0 \text{ for all } j \in J' \text{ implies } \lambda_j = 0 \text{ for all } j \in J'.$$

The classical definition of the MFCQ also takes into account equality constraints in the set Q, which we omit since they are not of immediate interest.

To handle sets defined by nonsmooth constraints, we need to recall the subdifferential.

Definition 3.7. (Subdifferentials) Consider a function $f : \mathbb{R}^n \to \mathbb{R}$ such that f is locally Lipschitz at $\bar{x}$. For a vector $v \in \mathbb{R}^n$, one says that

(a) v is a *regular subgradient* (also known as a *Fréchet subgradient*) of f at $\bar{x}$, written $v \in \hat{\partial} f(\bar{x})$, if

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle + o(|x - \bar{x}|);$$

(b) v is a *(general) subgradient* of f at $\bar{x}$, written $v \in \partial f(\bar{x})$, if there are sequences $x_i \to \bar{x}$ and $v_i \to v$ such that $f(x_i) \to f(\bar{x})$ and $v_i \in \hat{\partial} f(x_i)$.

(c) The set $\hat{\partial} f(\bar{x})$ is the *regular subdifferential,* and the set $\partial f(\bar{x})$ is the *(general) subdifferential.*

(d) The function f is (subdifferentially) regular at $\bar{x}$ if and only if $\partial f(\bar{x}) = \hat{\partial} f(\bar{x})$.

This characterization of subdifferentially regular functions is slightly different from the earlier definitions, but is equivalent in the case of locally Lipschitz functions in view of [9, Corollary 8.11, Theorem 9.13 and Theorem 8.6]. We shall only be concerned with subdifferentially regular functions throughout this paper, so there is no need to distinguish between $\partial f(\bar{x})$ and $\hat{\partial} f(\bar{x})$. We conclude with results on the intersections of tangent cones described by constraints.

Proposition 3.8. *(Tangent cone under constraints) Suppose* $C = \{x \mid f_j(x) \leq 0, j \in J\}$*, and* J *is a finite set. At the point* $\bar{x} \in C$*, let* $J' \subset J$ *be the set of all* j*'s such that* $f_j(\bar{x}) = 0$*. If* f_j *are continuous at* $\bar{x}$ *for all* $j \in J$*,* f_j *are continuously differentiable at* $\bar{x}$ *for all* $j \in J'$ *and the MFCQ is satisfied at* $\bar{x} \in C$*, then*

$$T_C(\bar{x}) = \{z \mid \nabla f_j(\bar{x})^T z \leq 0 \text{ for all } j \in J'\}.$$

In the nonsmooth case, if f_j *were locally Lipschitz and subdifferentially regular at* $\bar{x}$ *for all* $j \in J'$ *and*

$$\sum_{j \in J'} \lambda_j v_j = 0,\ v_j \in \partial f_j(\bar{x}) \text{ and } \lambda_j \geq 0 \text{ for all } j \in J' \quad \text{implies } \lambda_j = 0 \text{ for all } j \in J', \tag{3.3}$$

then $\bar{x}$ *is Clarke regular at* C*, and*

$$\begin{aligned} T_C(\bar{x}) &= \bigcap_{j \in J'} \{z \mid v^T z \leq 0 \text{ for all } v \in \partial f_j(\bar{x})\}. \\ &= \{z \mid v^T z \leq 0 \text{ for all } v \in \bigcup_{j \in J'} \partial f_j(\bar{x})\}. \end{aligned} \tag{3.4}$$

Proof. We prove the general nonsmooth case for this theorem, which implies the smooth case. There is a neighborhood U of $\bar{x}$ such that $C \cap U = [\cap_{j \in J'} C_j] \cap U$, where C_j is defined by $C_j = \{x \mid f_j(x) \leq 0\}$. Furthermore, $0 \notin \partial f_j(\bar{x})$ for all $j \in J'$. By [9, Theorem 10.3] (normal cones to level sets) and [9, Corollary 6.29(d)] (tangent-normal relations in regular sets), C_j is Clarke regular at $\bar{x}$, and the tangent cones $T_{C_j}(\bar{x})$ and normal cones $N_{C_j}(\bar{x})$ are given by

$$\begin{aligned} N_{C_j}(\bar{x}) &= \{\lambda v \mid \lambda \geq 0, \text{ and } v \in \partial f_j(\bar{x})\}, \\ \text{and } T_{C_j}(\bar{x}) &= \{w \mid w^T v \leq 0 \text{ for all } v \in \partial f_j(\bar{x})\}. \end{aligned}$$

Therefore, condition (3.3) becomes

$$\sum_{j \in J'} v_j = 0,\ v_j \in N_{C_j}(\bar{x}) \text{ implies } v_j = 0 \text{ for all } j \in J'.$$

By Proposition 3.5, the tangent cone $T_C(\bar{x})$ is

$$T_C(\bar{x}) = \bigcap_{i \in J'} T_{C_j}(\bar{x}),$$

which gives the formula for the tangent cone in the statement. □

It is well known that for the sets

$$\begin{aligned} C_1 &:= \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq x_1^2\} \\ C_2 &:= \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \leq -x_1^2\}, \end{aligned}$$

we have $T_{C_1 \cap C_2}(0) \subsetneq T_{C_1}(0) \cap T_{C_2}(0)$ but the MFCQ is not satisfied.

The constraint qualification condition (3.3) can be checked by another equivalent condition when $T_{C_j}(\bar{x})$ have nonempty interior.

Proposition 3.9. *(Constraint qualification) Assume the conditions of Proposition 3.8. If* $T_{C_j}(\bar{x})$ *have nonempty interior and* $0 \notin \partial f_j(\bar{x})$ *for all* $j \in J'$*, then the condition* (3.3) *is equivalent to the existence of a vector* w *such that* $w \in \operatorname{int}(T_{C_j}(\bar{x}))$ *(or equivalently* $w^T v < 0$ *for all* $v \in \partial f_j(\bar{x})$*) for all* $j \in J'$*.*

Proof. Recall that by [9, Theorem 10.3], if $0 \notin \partial f(\bar{x})$, then the tangent cone $T_{C_j}(\bar{x})$ is equal to $\{z \mid z^T v \leq 0 \text{ for all } v \in \partial f_j(\bar{x})\}$, and the interior $\text{int}(T_{C_j}(\bar{x}))$ is $\{z \mid z^T v < 0 \text{ for all } v \in \partial f_j(\bar{x})\}$, which gives the equivalence on the conditions on w. Next, the equivalence of (3.3) and the condition in this result follow from Proposition 3.5. ☐

4. Robust nonlinear programming

We look at nonlinear programs of the form

$$\min_x \{c^T x + d \mid \bar{A}x - \bar{b} \in Q\}, \tag{4.1}$$

where $Q \subset \mathbb{R}^k$ is a closed set. Specifically, we consider problems of the form

$$\min_x \{c^T x + d \mid \bar{A}_i x - \bar{b}_i \in Q_i,\, 1 \leq i \leq m\}, \tag{4.2}$$

where $Q_i \subset \mathbb{R}^{k_i}$ are nonempty closed sets, $\bar{A}_i \in \mathbb{R}^{k_i \times n}$, and $\bar{b}_i \in \mathbb{R}^{k_i}$. We may write $\bar{A}$ as a concatenation of the matrices $\bar{A}_i$ and $\bar{b}$ as a concatenation of the vectors $\bar{b}_i$, and this would make (4.1) equivalent to (4.2) for $Q = Q_1 \times \cdots \times Q_m$ and $k = k_1 + \cdots + k_m$. One case of interest is the set $Q_i = \{y \mid f_{i,j}(y) \leq 0 \text{ for all } j \in J\}$ for some $f_{i,j} : \mathbb{R}^{k_i} \to \mathbb{R}$ and the set J is finite. Another case of interest is *conic programs*, which arise when all Q_i's are closed convex pointed cones with nonempty interior.

We now recall the definition of robust feasibility from [3].

Definition 4.1. (Robust feasibility) Let an uncertain problem be given and $\Delta\mathcal{U} = \Delta\mathcal{U}_1 \times \cdots \times \Delta\mathcal{U}_m$ be a perturbation set. A candidate solution $x \in \mathbb{R}^n$ is *robustly feasible* if it remains feasible for all realizations of the perturbation vector from the perturbation set, that is

$$[\bar{A}_i + \Delta A_i]x - [\bar{b}_i + \Delta b_i] \in Q_i \,\forall (i, 1 \leq i \leq m, (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i), \tag{4.3}$$

where $\Delta\mathcal{U}_i \subset (\mathbb{R}^{k_i \times n} \times \mathbb{R}^{k_i})$ is the uncertainty set in $(\bar{A}_i, \bar{b}_i)$.

Remark 4.2. (Decomposing uncertainty sets) In the case where $\Delta\mathcal{U}$ is not a direct product of uncertainty sets, the uncertainty sets $\Delta\mathcal{U}_i$ can be defined as

$$\Delta\mathcal{U}_i := \{(\Delta A_i, \Delta b_i) : (\Delta A_i, \Delta b_i) = \Pi_i(\Delta A, \Delta b) \text{ for some } (\Delta A, \Delta b) \in \Delta\mathcal{U}\},$$

where Π_i is the relevant projection from $\mathbb{R}^{k \times n} \times \mathbb{R}^k$ to $\mathbb{R}^{k_i \times n} \times \mathbb{R}^{k_i}$. It is clear that (4.3) is equivalent to

$$[\bar{A} + \Delta A]x - [\bar{b} + \Delta b] \in Q \,\forall (\Delta A, \Delta b) \in \Delta\mathcal{U}.$$

The definition for robust nonlinear programs encompasses nonlinear objective functions.

Example 4.3. (Nonlinear objective) Consider the robust optimization problem

$$\begin{aligned} &\min_x f(x) \\ &\text{s.t. } [\bar{A}_i + \Delta A_i]x - [\bar{b}_i + \Delta b_i] \in Q_i \,\forall (i, 1 \leq i \leq m, (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i). \end{aligned}$$

We can rewrite this robust problem as

$$\begin{aligned} &\min_{x,t} t \\ &\text{s.t. } [\bar{A}_i + \Delta A_i]x - [\bar{b}_i + \Delta b_i] \in Q_i \,\forall (i, 1 \leq i \leq m, (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i) \\ &\text{and } f(x) \leq t. \end{aligned}$$

The function f is convex if and only if the epigraph $\mathrm{epi}(f) = \{(x,t) \mid f(x) \leq t\}$ is convex. Similarly, for a function f locally Lipschitz at $\bar{x}$, the function f is subdifferentially regular at $\bar{x}$ if and only if $\mathrm{epi}(f)$ is Clarke regular at $\bar{x}$. To prove our results for nonlinear functions, we can prove the result for linear objective functions and then appeal to the second formulation to obtain the result we need.

The formula in the robust optimization constraint can be rewritten as

$$\begin{aligned} & [\bar{A}_i + \Delta A_i](\bar{x} + \Delta x) - [\bar{b}_i + \Delta b_i] \in Q_i \\ \iff \quad & [\bar{A}_i\bar{x} - \bar{b}_i] + \bar{A}_i(\Delta x) + [(\Delta A_i)\bar{x} - \Delta b_i] + (\Delta A_i)(\Delta x) \in Q_i. \end{aligned}$$

As in linear programming, we eliminate the second order term $(\Delta A)(\Delta x)$ to obtain a first order approximation. For nonlinear programs, we also need to approximate the set Q_i at $\bar{A}_i\bar{x} - \bar{b}_i$ by the tangential approximation $T_{Q_i}(\bar{A}_i\bar{x} - \bar{b}_i) + [\bar{A}_i\bar{x} - \bar{b}_i]$ at $\bar{A}_i\bar{x} - \bar{b}_i$. This gives our definition of the tangential problem.

Definition 4.4. (Tangential problem) Let $\bar{x}$ be an optimal solution to a nonlinear programming problem with parameters $(\bar{A}, \bar{b})$ so that Q_i is Clarke regular at $\bar{A}_i\bar{x}-\bar{b}_i$ for all i. The *tangential problem* to the robust optimization problem obtained with constraints as explained in Definition 4.1 is

$$\begin{aligned} & \min_{\gamma} c^T\gamma \\ & \text{s.t. } [\bar{A}_i\bar{x} - \bar{b}_i] + \bar{A}_i\gamma + [(\Delta A_i)\bar{x} - \Delta b_i] \in T_{Q_i}(\bar{A}_i\bar{x} - \bar{b}_i) + [\bar{A}_i\bar{x} - \bar{b}_i] \\ & \qquad \text{for all } (i, 1 \leq i \leq m, (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i), \end{aligned}$$

or equivalently

$$\begin{aligned} & \min_{\gamma} c^T\gamma \\ & \text{s.t. } \bar{A}_i\gamma + [(\Delta A_i)\bar{x} - \Delta b_i] \in T_{Q_i}(\bar{A}_i\bar{x} - \bar{b}_i) \\ & \qquad \text{for all } (i, 1 \leq i \leq m, (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i), \end{aligned} \tag{4.4}$$

which is also equivalent to

$$\begin{aligned} & \min_{\gamma} c^T\gamma \\ & \text{s.t. } \bar{A}_i\gamma + L_i(\Delta\mathcal{U}_i) \in T_{Q_i}(\bar{A}_i\bar{x} - \bar{b}_i) \\ & \qquad \text{for all } (i, 1 \leq i \leq m), \end{aligned}$$

where $L_i : \mathbb{R}^{m_i \times n} \times \mathbb{R}^{m_i} \to \mathbb{R}^{m_i}$ is defined by $L_i(\Delta A_i, \Delta b_i) = (\Delta A_i)\bar{x} - \Delta b_i$. We call the corresponding constraints to the tangential problem the *tangential constraints*.

Remark 4.5. (Clarke regularity assumption) The assumption that each Q_i is Clarke regular at $\bar{A}_i\bar{x} - \bar{b}_i$ in Definition 4.4 comes about because the set $Q_1 \times \cdots \times Q_m$ is Clarke regular at $\bar{A}\bar{x} - \bar{b}$ if and only if Q_i is Clarke regular at $\bar{A}_i\bar{x} - \bar{b}_i$ for all i, and in this case,

$$T_{Q_1 \times \cdots \times Q_m}(\bar{A}\bar{x} - \bar{b}) = T_{Q_1}(\bar{A}_1\bar{x} - \bar{b}_1) \times \cdots \times T_{Q_m}(\bar{A}_m\bar{x} - \bar{b}_m).$$

(see [9, Proposition 6.41].) This property makes the tangential problem independent of how we decompose the set Q as a direct product of sets.

We give some examples of tangential constraints.

Example 4.6. (Examples of tangential constraints) (a) When $\bar{A}_i\bar{x} - \bar{b}_i = 0$ and Q_i is a closed convex cone, then $T_{Q_i}(0) = Q_i$. In this case, the corresponding tangential constraint is obtained by just removing the second order term $(\Delta A_i)(\Delta x)$.

(b) When $\bar{A}_i\bar{x} - \bar{b}_i \in \operatorname{int}(Q_i)$, then $T_{Q_i}(\bar{A}_i\bar{x} - \bar{b}_i) = \mathbb{R}^{k_i}$ and the corresponding tangential constraint vanishes.

In view of Example 4.6, we see that for linear programming, the tangential constraints and first order constraints are equivalent. When $\bar{A}_i\bar{x} - \bar{b}_i \in \partial Q_i \backslash \{0\}$, we may still be able to calculate the tangential constraints using the material recalled in Section 3.

We illustrate the tangential problem with the example on second order cone programming (SOCP).

Example 4.7. (Tangential problem in SOCP) Consider the SOCP problem

$$\begin{aligned} &\min_x c^T x \\ &\text{s.t. } \bar{A}_i x - \bar{b}_i \in Q_{k_i} \text{ for all } 1 \le i \le m, \end{aligned}$$

where $Q_d \subset \mathbb{R}^d$ is the *second order cone*

$$Q_d := \{w = (w_0, \dots, w_{d-1}) \in \mathbb{R}^d \mid \|(w_1, \dots, w_{d-1})\|_2 - w_0 \le 0\}.$$

Given an optimal solution $\bar{x}$, we show how to obtain the tangential constraint. If $\bar{A}_i\bar{x} - \bar{b}_i \in \operatorname{int}(Q_{k_i})$, then $T_{Q_{k_i}}(\bar{A}_i\bar{x} - \bar{b}_i) = \mathbb{R}^{k_i}$ by Example 4.6, and so the tangential constraint vanishes. If $\bar{A}_i\bar{x} - \bar{b}_i = 0$, then $T_{Q_{k_i}}(\bar{A}_i\bar{x} - \bar{b}_i) = Q_{k_i}$ by Example 4.6, so the tangential constraint is

$$\bar{A}_i\gamma + L_i(\Delta\mathcal{U}_i) \subset Q_{k_i}.$$

We now consider the case $\bar{A}_i\bar{x} - \bar{b}_i \in \partial Q_{k_i} \backslash \{0\}$. Let $\bar{z} = \bar{A}_i\bar{x} - \bar{b}_i$. In this case, $\bar{z}_0 = \|(\bar{z}_1, \dots, \bar{z}_{k_i-1})\|_2$. The gradient of the map $(w_0, w_1, \dots, w_{k_i-1}) \mapsto \|(w_1, \dots, w_{k_i-1})\|_2 - w_0$ at $\bar{z}$ is $(-1, \frac{(\bar{z}_1, \dots, \bar{z}_{k_i-1})}{\|(\bar{z}_1, \dots, \bar{z}_{k_i-1})\|_2})$. Let $R : \mathbb{R}^{k_i} \to \mathbb{R}^{k_i}$ be the reflection map

$$R(w_0, w_1, \dots, w_{k_i-1}) := (-w_0, w_1, \dots, w_{k_i-1}),$$

i.e., R multiplies the 0th coordinate by -1. The gradient at $\bar{z}$ can also be written as $\frac{1}{\bar{z}_0} R\bar{z}$. Therefore, by Proposition 3.8,

$$T_{Q_{k_i}}(\bar{A}_i\bar{x} - \bar{b}_i) = \{w \in \mathbb{R}^{k_i} \mid w^T R\bar{z} \le 0\}.$$

Therefore, the tangential constraint is

$$\bar{A}_i\gamma + [(\Delta A_i)\bar{x} - \Delta b_i] \in T_{Q_{k_i}}(\bar{A}_i\bar{x} - \bar{b}_i) \text{ for all } (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i.$$

This can be written equivalently as

$$\begin{aligned} &\bar{z}^T R\bar{A}_i\gamma + \bar{z}^T R[(\Delta A_i)\bar{x} - \Delta b_i] \le 0 \text{ for all } (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i, \\ \text{or} \quad &\bar{z}^T R\bar{A}_i\gamma \le - \max_{(\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i} \bar{z}^T R[(\Delta A_i)\bar{x} - \Delta b_i]. \end{aligned}$$

5. Main result: Approximation using the tangential problem

In Theorem 5.3 we prove that if the uncertainty set in a robust optimization problem is dilated or expanded, then the robust optimal solution can be predicted from the exact solution of the nonrobust problem and the tangential problem.

We now prove a lemma needed for the proof of our main result.

Lemma 5.1. *(Compact sets in convex cones) Let $D \subset \mathbb{R}^n$ be Clarke regular at 0 and $C \subset \mathbb{R}^n$ be a compact convex set such that $\mathrm{int}(T_D(0)) \neq \emptyset$ and $C \subset T_D(0)$. Let $v \in \mathrm{int}(T_D(0))$. Then for all sufficiently small $\delta > 0$, there exists $\bar{\epsilon} > 0$ such that $C + \delta v + \delta^2 \mathbb{B} \subset \frac{1}{\epsilon} D$ for all $\epsilon \in (0, \bar{\epsilon}]$.*

Proof. Since $v \in \mathrm{int}(T_D(0))$, $v + \delta \mathbb{B} \subset \mathrm{int}(T_D(0))$ for all sufficiently small $\delta > 0$, and therefore $C + \delta v + \delta^2 \mathbb{B} \subset \mathrm{int}(T_D(0))$. For every point $w \in C + \delta v + \delta^2 \mathbb{B}$, we can find a convex polyhedral set P_w such that $\mathrm{int}(P_w) \neq \emptyset$ and $P_w \subset \mathrm{int}(T_D(0))$. A compactness argument shows that the set $C + \delta v + \delta^2 \mathbb{B}$ is contained in the interior of finitely many of these convex polyhedral sets, so there is a convex polyhedral set P such that $C + \delta v + \delta^2 \mathbb{B} \subset P \subset \mathrm{int}(T_D(0))$.

By the recession properties of tangent cones and the Clarke regularity of D, there is an $\bar{\epsilon} > 0$ such that $\epsilon \mathrm{conv}(\{0\} \cup P) \subset D$ for all $\epsilon \in [0, \bar{\epsilon}]$ (see [9, Exercise 6.34(a)]. The roots of this result on local recession vectors can be traced back to [7].). Therefore $C + \delta v + \delta^2 \mathbb{B} \subset \mathrm{conv}(\{0\} \cup P) \subset \frac{1}{\epsilon} D$ for all $\epsilon \in [0, \bar{\epsilon}]$ as needed. □

We also need material in set-valued analysis as presented in [9, Chapters 4 and 5] for the proof of Theorem 5.3.

Definition 5.2. [9, Definition 5.4] (Set-valued continuity) We say that S is a *set-valued map*, denoted by $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, if $S(x) \subset \mathbb{R}^m$. A set-valued map S is *outer semicontinuous (osc)* at $\bar{x}$ if

$$\limsup_{x \to \bar{x}} S(x) \subset S(\bar{x}),$$

or equivalently $\limsup_{x \to \bar{x}} S(x) = S(\bar{x})$, but *inner semicontinuous (isc)* at $\bar{x}$ if

$$\liminf_{x \to \bar{x}} S(x) \supset S(\bar{x}),$$

or equivalently when S is closed-valued, $\liminf_{x \to \bar{x}} S(x) = S(\bar{x})$. It is called *continuous* at $\bar{x}$ if both conditions hold, i.e., if $S(x) \to S(\bar{x})$ as $x \to \bar{x}$. Here, the *outer limit* $\limsup_{x \to \bar{x}} S(x)$ and the *inner limit* $\liminf_{x \to \bar{x}} S(x)$ are defined by

$$\begin{aligned} \limsup_{x \to \bar{x}} S(x) &:= \{u \mid \exists x_i \to \bar{x}, \exists u_i \to u \text{ with } u_i \in S(x_i)\} \\ \liminf_{x \to \bar{x}} S(x) &:= \big\{u \mid \forall x_i \to \bar{x}, \exists \{u_i\} \text{ with } u_i \in S(x_i) \\ &\qquad \text{s.t. } u \text{ is the limit of a subsequence of } \{u_i\}\big\}. \end{aligned}$$

If S maps to compact sets, continuity as defined by inner and outer limits above is equivalent to continuity in the Pompieu-Hausdorff distance, which is a metric in the subset of compact sets. We refer to [9] for more details. We also need to recall the definition of epi-convergence. A sequence of functions $h_i : \mathbb{R}^n \to \mathbb{R}$ is said to *epi-converge* to a function $h : \mathbb{R}^n \to \mathbb{R}$, written $h_i \xrightarrow{e} h$, if $\mathrm{epi}(h_i) \to \mathrm{epi}(h)$. The history of epi-convergence can be traced back to the 1960's, and the result we need for our proof ([9, Theorem 7.33]) can be traced back to Salinetti (unpublished, but reported in [8]) and [2]. See [9, Chapter 7].

Here is our theorem on the approximation properties of the tangential problem.

Theorem 5.3. *(Approximation properties of Tangential problem) Consider the robust optimization problem*

$$\begin{aligned} &\min_x \; f(x) \\ (5.1) \qquad &\text{s.t. } [\bar{A}_i + \Delta A_i]x - [\bar{b}_i + \Delta b_i] \in Q_i \text{ for all } (\Delta A_i, \Delta b_i) \in \epsilon \Delta \mathcal{U}_i \text{ for all } i, \end{aligned}$$

and the tangential problem

$$\begin{aligned} &\min_{\gamma} \; df(\bar{x})(\gamma) \\ \text{(5.2)} \qquad &\text{s.t. } \bar{A}_i\gamma + [(\Delta A_i)\bar{x} - \Delta b_i] \in T_{Q_i}(\bar{A}_i\bar{x} - \bar{b}_i) \\ &\text{for all } (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i \text{ for all } i. \end{aligned}$$

Let $\bar{\Gamma}$ be the set of optimal solutions to (5.2), $\Phi = \{x \mid \bar{A}x - \bar{b} \in Q\}$, *and $\bar{x}$ be a solution of the nominal problem* $\min\{f(x) \mid \bar{A}x - \bar{b} \in Q\}$. *Suppose*

(1) *Q_i are closed sets that are Clarke regular at $\bar{A}_i\bar{x} - \bar{b}_i$,*
(2) *$\Delta\mathcal{U}_i$ are compact convex sets*
(3) *$\bar{\Gamma}$ is bounded.*
(4) *There is some γ' such that $\bar{A}_i\gamma' \in \operatorname{int}(T_{Q_i}(\bar{A}_i\bar{x} - \bar{b}_i))$ for all i.*
(5) *(Compactness) If $\{x_i\} \subset \Phi$ and $f(x_i) \to f(\bar{x})$, then $x_i \to \bar{x}$.*
(6) *f is locally Lipschitz and subdifferentially regular at $\bar{x}$.*

Then the set of cluster points of any sequence $\{\frac{1}{\epsilon}(\bar{x}_\epsilon - \bar{x})\}$, where $\bar{x}_\epsilon$ is an optimal solution to (5.1) *and $\epsilon \searrow 0$, is a subset of $\bar{\Gamma}$. The objective value of* (5.1) *, say $\bar{v}_\epsilon$, has an approximation $\bar{v}_\epsilon = \bar{v} + \epsilon\tilde{v} + o(\epsilon)$, where $\tilde{v}$ is the objective value of* (5.2).

In particular, if $\bar{\Gamma}$ contains only one element, say $\bar{\gamma}$, then $\lim_{\epsilon\searrow 0} \frac{1}{\epsilon}(\bar{x}_\epsilon - \bar{x}) = \bar{\gamma}$, or $\bar{x}_\epsilon \in \bar{x} + \epsilon\bar{\gamma} + o(\epsilon)$.

Proof. The proof of this result is broken up into four steps. In steps 1 to 3, we prove this result for the affine function $f(x) = c^T x + d$, and $df(\bar{x})(\gamma) = c^T x$. In step 4, we use the observation in Example 4.3 to treat the case where f is locally Lipschitz and subdifferentially regular at $\bar{x}$.

Step 1: Rewriting the robust optimization problem (5.1)**.**

We rewrite the constraint in the robust optimization problem.

$$\begin{aligned} & [\bar{A}_i + \Delta A_i]x - [\bar{b}_i + \Delta b_i] \in Q_i \\ \Leftrightarrow \quad & [\bar{A}_i + \Delta A_i]x - \bar{A}_i\bar{x} - \Delta b_i \in Q_i - [\bar{A}_i\bar{x} - \bar{b}_i] \\ \Leftrightarrow \quad & \bar{A}_i(x - \bar{x}) + [(\Delta A_i)\bar{x} - \Delta b_i] + (\Delta A_i)(x - \bar{x}) \in Q_i - [\bar{A}_i\bar{x} - \bar{b}_i]. \end{aligned}$$

Hence,

$$\begin{aligned} & A_i x - b_i \in Q_i \text{ for all } (\Delta A_i, \Delta b_i) \in \epsilon\Delta\mathcal{U}_i \\ \Leftrightarrow \quad & \bar{A}_i(x - \bar{x}) + [(\Delta A_i)\bar{x} - \Delta b_i] + (\Delta A_i)(x - \bar{x}) \in Q_i - [\bar{A}_i\bar{x} - \bar{b}_i] \\ & \qquad \text{for all } (\Delta A_i, \Delta b_i) \in \epsilon\Delta\mathcal{U}_i. \end{aligned}$$

The next step is to scale the variables ΔA_i and Δb_i so that the ϵ vanishes from the expression $\epsilon\Delta\mathcal{U}_i$. This gives

$$\begin{aligned} & \bar{A}_i(x - \bar{x}) + [(\Delta A_i)\bar{x} - \Delta b_i] + (\Delta A_i)(x - \bar{x}) \in Q_i - [\bar{A}_i\bar{x} - \bar{b}_i] \\ & \qquad \text{for all } (\Delta A_i, \Delta b_i) \in \epsilon\Delta\mathcal{U}_i \\ \Leftrightarrow \quad & \bar{A}_i\Big[\frac{1}{\epsilon}(x - \bar{x})\Big] + [(\Delta A_i)\bar{x} - \Delta b_i] + \epsilon(\Delta A_i)\Big[\frac{1}{\epsilon}(x - \bar{x})\Big] \\ & \qquad \in \frac{1}{\epsilon}\big[Q_i - [\bar{A}_i\bar{x} - \bar{b}_i]\big] \text{ for all } (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i \\ \text{(5.3)} \qquad \Leftrightarrow \quad & \bar{A}_i\gamma_\epsilon + [(\Delta A_i)\bar{x} - \Delta b_i] + \epsilon(\Delta A_i)\gamma_\epsilon \\ & \qquad \in \frac{1}{\epsilon}\big[Q_i - [\bar{A}_i\bar{x} - \bar{b}_i]\big] \text{ for all } (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i, \end{aligned}$$

where $\gamma_\epsilon := \frac{1}{\epsilon}(x-\bar{x})$ in the final expression. We see that as $\epsilon \searrow 0$, the expressions in (5.3) converge to the corresponding expressions for the tangential constraints.

Let Γ_ϵ denote the set of all feasible γ_ϵ for the robust problem with parameter ϵ, and Γ denote the set of all feasible γ for the tangential problem. Similarly, let $\bar{\Gamma}_\epsilon$ and $\bar{\Gamma}$ denote the set of optimal solutions to the corresponding problems. It is elementary to check that the sets Γ_ϵ, Γ, $\bar{\Gamma}_\epsilon$ and $\bar{\Gamma}$ are all closed.

Step 2: $\lim_{\epsilon\searrow 0}\Gamma_\epsilon = \Gamma$.

Suppose that $\{\gamma_j\}$ is a sequence of feasible solutions to the robust problem with parameter ϵ_j, that is $\gamma_j \in \Gamma_{\epsilon_j}$. Then each γ_j satisfies the formula in (5.3) with parameter ϵ_j. It is clear that any limit of $\{\gamma_j\}$ is a feasible solution of the tangential problem (5.2), so $\limsup_{\epsilon\searrow 0}\Gamma_\epsilon \subset \Gamma$.

Next, we show that $\liminf_{\epsilon\searrow 0}\Gamma_\epsilon \supset \Gamma$. Suppose $\tilde{\gamma} \in \Gamma$. We need to show that for any choice of $\epsilon_j \searrow 0$, we can find $\gamma_j \in \Gamma_{\epsilon_j}$ such that $\tilde{\gamma} = \lim_{j\to\infty}\gamma_j$. Recall that $\tilde{\gamma}$ satisfies

$$\bar{A}_i\tilde{\gamma} + L_i(\Delta\mathcal{U}_i) \subset T_{Q_i}(\bar{A}_i\bar{x} - \bar{b}_i),$$

where the linear map $L_i : \mathbb{R}^{m_i\times n}\times\mathbb{R}^{m_i} \to \mathbb{R}^{m_i}$ by $L_i(\Delta A_i, \Delta b_i) = (\Delta A_i)\bar{x} - \Delta b_i$. It follows from the convexity of $\Delta\mathcal{U}_i$ that $L_i(\Delta\mathcal{U}_i)$ is convex. Since $\bar{A}_i\gamma' \in \text{int}(T_{Q_i}(\bar{A}_i\bar{x}-\bar{b}_i))$, we can apply Lemma 5.1 to tell us that for all sufficiently small $\delta > 0$, there is some $\bar{\epsilon} > 0$ such that $[\bar{A}_i(\tilde{\gamma}+\delta\gamma')+L_i(\Delta\mathcal{U}_i)+\delta^2\mathbb{B}] \subset \frac{1}{\epsilon}[Q_i-(\bar{A}_i\bar{x}-\bar{b}_i)]$ for all $\epsilon \in [0,\bar{\epsilon}]$. Therefore

$$\bar{A}_i(\tilde{\gamma}+\delta\gamma') + L_i(\Delta\mathcal{U}_i) + \delta^2\mathbb{B} \subset \frac{1}{\epsilon}\big[Q_i - [\bar{A}_i\bar{x} - \bar{b}_i]\big].$$

If $\epsilon < \bar{\epsilon}$ and $\epsilon\max\{\|\Delta A_i\| \mid (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i\}\|(\tilde{\gamma}+\delta\gamma')\| < \delta^2$, then for all $(\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i$,

$$\begin{aligned}
(5.4)\qquad & \bar{A}_i(\tilde{\gamma}+\delta\gamma') + [(\Delta A_i)\bar{x} - \Delta b_i] + \epsilon(\Delta A_i)(\tilde{\gamma}+\delta\gamma') \\
\in\; & \bar{A}_i(\tilde{\gamma}+\delta\gamma') + L_i(\Delta\mathcal{U}_i) + \epsilon\|\Delta A_i\|\|(\tilde{\gamma}+\delta\gamma')\|\mathbb{B} \\
\subset\; & \bar{A}_i(\tilde{\gamma}+\delta\gamma') + L_i(\Delta\mathcal{U}_i) + \delta^2\mathbb{B} \\
\subset\; & \frac{1}{\epsilon}\big[Q_i - [\bar{A}_i\bar{x} - \bar{b}_i]\big].
\end{aligned}$$

With this observation, we can choose a sequence $\delta_j \searrow 0$ such that $(\tilde{\gamma}+\delta_j\gamma') \in \Gamma_{\epsilon_j}$, which gives $\liminf_{\epsilon\searrow 0}\Gamma_\epsilon \supset \Gamma$ as needed.

Step 3: $\limsup_{\epsilon\searrow 0}\bar{\Gamma}_\epsilon \subset \bar{\Gamma}$.

Recall that for a closed set $D \subset \mathbb{R}^n$, the *indicator function* $\delta_D : \mathbb{R}^n \to \mathbb{R}\cup\{\infty\}$ is defined by

$$\delta_D(x) := \begin{cases} 0 & \text{if } x \in D \\ \infty & \text{otherwise.} \end{cases}$$

Define $h_\epsilon : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}$ by $h_\epsilon(\gamma) := c^T\gamma + \delta_{\Gamma_\epsilon}(\gamma)$ and $h(\gamma) := c^T\gamma + \delta_\Gamma(\gamma)$. Since $\Gamma_\epsilon \to \Gamma$, we have $\text{epi}(\delta_{\Gamma_\epsilon}) = \Gamma_\epsilon \times [0,\infty) \to \Gamma\times[0,\infty) = \text{epi}(\delta_\Gamma)$ (by [9, Exercise 4.29(a)]), so in other words $\delta_{\Gamma_\epsilon} \xrightarrow{e} \delta_\Gamma$. By [9, Exercise 7.8(a)] we have $h_\epsilon \xrightarrow{e} h$. We seek to apply [9, Theorem 7.33], which gives us the result we need. Before we can do so, we have to check that $\{h_\epsilon\}$ is eventually level bounded, that is, for any $\alpha \in \mathbb{R}$, we have $\cup_{[0,\epsilon]}\{\gamma \mid h_\epsilon(\gamma) \leq \alpha\}$ being bounded for some $\epsilon > 0$.

Recall $\Phi = \{x \mid \bar{A}x - \bar{b} \in Q\}$. By Proposition 6.6, the boundedness of $\bar{\Gamma}$ is equivalent to $\{c\}^\perp \cap T_\Phi(\bar{x}) = \{0\}$. Suppose on the contrary that $\{h_\epsilon\}$ is not eventually level bounded. Then there is some α and sequences $\{\epsilon_i\}$ and $\{\gamma_i\}$ such that $\epsilon_i \searrow 0$, $\{\gamma_i\}$ is unbounded, and $h_{\epsilon_i}(\gamma_i) \leq \alpha$ (or equivalently, $\gamma_i \in \Gamma_{\epsilon_i}$ and $c^T\gamma_i \leq \alpha$). Let us write $x_i = \bar{x} + \epsilon_i\gamma_i$. Since $x_i \in \Phi$ and

$$c^T\bar{x} \leq c^T x_i = c^T\bar{x} + \epsilon_i c^T\gamma_i \leq c^T\bar{x} + \epsilon_i\alpha,$$

we have $c^T x_i \to c^T\bar{x}$. Note that this also gives us $\alpha \geq 0$. By our compactness assumption, $x_i \to \bar{x}$, which means that $|\epsilon_i\gamma_i| \to 0$. This means that $\frac{\gamma_i}{|\gamma_i|}$ converges to a vector γ_∞ in $T_\Phi(0)\backslash\{0\}$. Observe that $c^T\gamma_\infty = \lim_{i\to\infty} c^T\frac{\gamma_i}{|\gamma_i|} \leq \lim_{i\to\infty}\frac{\alpha}{|\gamma_i|} = 0$, and that $c^T\gamma_i \geq 0$, so $c^T\gamma_\infty = 0$. This is a contradiction to $\{c\}^\perp \cap T_\Phi(\bar{x}) = \{0\}$. We can thus apply [9, Theorem 7.33] to conclude that $\min\{c^T\gamma \mid \gamma \in \Gamma_\epsilon\}$ converges to $\min\{c^T\gamma \mid \gamma \in \Gamma\}$ and $\limsup_{\epsilon\searrow 0}\bar{\Gamma}_\epsilon \subset \bar{\Gamma}$, ending the proof of the theorem for the linear case.

Step 4: Locally Lipschitz subdifferentially regular f at $\bar{x}$.

Consider the problem

$$\begin{aligned} &\min_{x,t}\ t \\ &\text{s.t. } [\bar{A}_i + \Delta A_i]x - [\bar{b}_i + \Delta b_i] \in Q_i \text{ for all } (\Delta A_i, \Delta b_i) \in \epsilon\Delta\mathcal{U}_i \text{ for all } i, \\ &\text{and } (x,t) \in \operatorname{epi}(f) = \{(x,t) \mid f(x) \leq t\}. \end{aligned} \tag{5.5}$$

and the tangential problem

$$\begin{aligned} &\min_{\gamma,s}\ s \\ &\text{s.t. } \bar{A}_i\gamma + [(\Delta A_i)\bar{x} - \Delta b_i] \in T_{Q_i}(\bar{A}_i\bar{x} - \bar{b}_i) \\ &\qquad \text{for all } (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i \text{ for all } i. \\ &\text{and } (\gamma, s) \in T_{\operatorname{epi}(f)}(\bar{x}, f(\bar{x})). \end{aligned} \tag{5.6}$$

The robust and tangential problems are equivalent to the respective problems (5.1) and (5.2) in the statement of the theorem. We now show that if conditions (1) to (5) in the theorem statement are satisfied for (5.1) and (5.2), then these conditions hold for (5.5) and (5.6) as well.

For condition (1), we need only to check that $\operatorname{epi}(f)$ is Clarke regular at $(\bar{x}, f(\bar{x}))$, which is immediate from the subdifferential regularity of f at $\bar{x}$. Condition (2) is straightforward. For condition (3), we note that the set of minimizers of (5.6) is just $\bar{\Gamma} \times \{df(\bar{x})(\hat{\gamma})\}$, where $\hat{\gamma}$ is any element in $\bar{\gamma}$. The set $\bar{\Gamma} \times \{df(\bar{x})(\hat{\gamma})\}$ is bounded if and only if $\bar{\Gamma}$ is bounded.

We further assume that f is locally Lipschitz at $\bar{x}$ with Lipschitz modulus κ. For condition (4), suppose γ' is a vector such that $\bar{A}_i\gamma' \in \operatorname{int}(T_{Q_i}(\bar{A}_i\bar{x} - \bar{b}_i))$ for all i. Since $(\mathbf{0}, 1) \in \operatorname{int}(T_{\operatorname{epi}(f)}(\bar{x}, f(\bar{x})))$, we have $(\gamma', (\kappa+1)|\gamma'|) \in \operatorname{int}(T_{\operatorname{epi}(f)}(\bar{x}, f(\bar{x})))$, which verifies condition (4).

We also need to check that given $f(x_i) \to f(\bar{x})$ and $\{x_i\} \subset \Phi$ implies $x_i \to \bar{x}$, we have the compactness condition that $t_i \to f(\bar{x})$, (x_i, t_i) satisfies $f(x_i) \leq t_i$ and $\{x_i\} \subset \Phi$ implies $(x_i, t_i) \to (\bar{x}, f(\bar{x}))$. Since $f(\bar{x}) \leq f(x_i) \leq t_i$ and $t_i \to f(\bar{x})$, we have $f(x_i) \to f(\bar{x})$, which gives $x_i \to \bar{x}$, and thus $(x_i, t_i) \to (\bar{x}, f(\bar{x}))$ as needed. □

We now take a closer look at step 4 of the proof of Theorem 5.3. Consider the general case where $\bar{\Gamma}_\epsilon = \arg\min\{c^T\gamma \mid \gamma \in \Gamma_\epsilon\}$, $\bar{\Gamma} = \arg\min\{c^T\gamma \mid \gamma \in \Gamma\}$ and $\Gamma = \lim_{\epsilon\searrow 0}\Gamma_\epsilon$. It may turn out that $\limsup_{\epsilon\searrow 0}\bar{\Gamma}_\epsilon \subsetneq \bar{\Gamma}$, as the example in Figure 5.1 shows. Example 7.5 shows that it is possible for $\bar{\Gamma}$ to be bounded but not be a singleton set. In such cases, it is possible that $\bar{\Gamma}_\epsilon$ is a singleton set, which occurs when the function f is strictly convex for example.

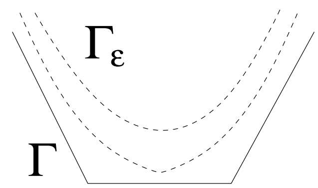

FIGURE 5.1. $\lim_{\epsilon\to 0}\Gamma_\epsilon = \Gamma$, but $\limsup_{\epsilon\to 0}\bar{\Gamma}_\epsilon \subsetneq \bar{\Gamma}$.

We make an observation on the condition $\bar{A}_i\gamma' \in \mathrm{int}(T_{Q_i}(\bar{A}_i\bar{x}-\bar{b}_i))$ for all i in Theorem 5.3.

Example 5.4. (Constraint qualification in tangential problem) The optimization problem

$$\begin{aligned}\min_{x\in\mathbb{R}^2}\ & x_2\\ \text{s.t. }\ & x\in Q_1 := \{x \mid -x_1+x_2^2\le 0\}\\ & x\in Q_2 := \{x \mid x_1+x_2^2\le 0\}\end{aligned}$$

can be written equivalently as

$$\begin{aligned}\min_{x\in\mathbb{R}^2}\ & x_2\\ \text{s.t. }\ & x\in Q_1\cap Q_2 = \{(0,0)\},\end{aligned}$$

and the solution for both problems is $\bar{x}=(0,0)$. The tangential approximation for the first problem is

$$\begin{aligned}\min_{\gamma\in\mathbb{R}^2}\ & \gamma_2\\ \text{s.t. }\ & \left.\begin{aligned}\gamma\in T_{Q_1}(\bar{x}) = \mathbb{R}_+\times\mathbb{R}\\ \gamma\in T_{Q_2}(\bar{x}) = \mathbb{R}_-\times\mathbb{R}\end{aligned}\right\} \implies \gamma\in\{0\}\times\mathbb{R},\end{aligned}$$

while the tangential approximation for the second problem is

$$\begin{aligned}\min_{\gamma\in\mathbb{R}^2}\ & \gamma_2\\ \text{s.t. }\ & \gamma\in T_{Q_1\cap Q_2}(\bar{x}) = \{(0,0)\}.\end{aligned}$$

Clearly, the solutions for the two problems are different. This example shows that depending on how the optimization problem is written, the tangential problems may not be equivalent and may have different solutions. But note that in this case, $T_{Q_1\cap Q_2}(\bar{x}) \subsetneq T_{Q_1}(\bar{x})\cap T_{Q_2}(\bar{x})$, which implies that the MFCQ does not hold, which in turn implies that there is no vector γ' such that $\bar{A}_i\gamma' \in \mathrm{int}(T_{Q_i}(\bar{A}_i\bar{x}-\bar{b}_i))$ for all i. The condition $\bar{A}_i\gamma' \in \mathrm{int}(T_{Q_i}(\bar{A}_i\bar{x}-\bar{b}_i))$ for all i in Theorem 5.3 ensures that the MFCQ holds whenever a group of the $\{(A_i,b_i)\}_{i=1}^m$ contains repetitions, ensuring that the tangent cones of the intersections is the intersections of the corresponding tangent cones through Proposition 3.5.

Theorem 5.3 shows that the decrease in objective function of the robust optimization problem is differentiable in the size of the uncertainty set at $\epsilon=0$. This observation can help give an approximate of the maximum robustness one can afford if the objective is to be above a certain value. The tangential problem also shows that the variables that we should estimate or measure more accurately are those which make the set $L(\Delta\mathcal{U}) = \{(\Delta A)\bar{x}-\Delta b \mid (\Delta A,\Delta b)\in\Delta\mathcal{U}\}$ small. For example, if $\bar{x} = 0$, then more effort should be spent on determining $\bar{b}$ accurately rather than entries in $\bar{A}$. Likewise, $\bar{x}$ determines which variables in $\bar{A}$ should be measured more accurately than others.

Besides these analytical properties, Theorem 5.3 shows that solving the tangential problem can help to obtain a good approximate of the robust solution. The robust optimization problem is known to be more computationally expensive than the original problem, so it will take more effort to obtain a desired level of accuracy. With the tangential problem, we can make use of the previously calculated optimization problem to obtain an approximate of the robust solution. Such an approximation is likely to be simpler than the robust optimization problem (for example, for LP in (2.5) and for SOCP in Example 4.7, though it still may be computationally difficult), and need not be computed to very high accuracy to obtain a good approximate of the robust optimization problem.

Remark 5.5. (Relaxing the constraint qualification in Theorem 5.3) The existence of γ' such that $\bar{A}\gamma' \in \operatorname{int}(T_Q(\bar{A}\bar{x} - \bar{b}))$ can be relaxed slightly if more structure is known about the set Q. All we need is for the chain of inclusions (5.4) to hold. For example, if Q_i is polyhedral and there is some vector γ' such that

$$\bar{A}_i\gamma' \in \operatorname{rint}\big(\operatorname{span}(\Pi_i(\Delta\mathcal{U}_i))\big) \cap T_{Q_i}(\bar{A}_i\bar{x} - \bar{b}_i),$$

where $\Pi_i(\Delta\mathcal{U}_i) = \{\Delta A_i \mid (\Delta A_i, \Delta b_i) \in \Delta\mathcal{U}_i \text{ for some } \Delta b_i\}$, then the chain of inclusions (5.4) would hold as well. A finer analysis on $\Pi_i(\Delta\mathcal{U})$ and local recession vectors can give stronger results.

6. First order optimality conditions of the tangential problem

In this section, we discuss first order optimality conditions of the tangential problem, which can be useful for designing specialized numerical methods for the tangential problem. In view of Theorem 5.3, we also give sufficient conditions for $\bar{\Gamma}$ to be bounded and for $\bar{\Gamma}$ to be a singleton.

As explained in [3, Chapters 5-8], the robust optimization problem is computationally tractable if either Q is polyhedral or $\Delta\mathcal{U}$ is polyhedral, while most other problems encountered in practice are not computationally tractable. Recall that a typical constraint in a robust optimization problem whose nominal solution is $\bar{x}$ is

$$[\bar{A} + \Delta A](\bar{x} + \Delta x) - [\bar{b} + \Delta b] \in Q \text{ for all } (\Delta A, \Delta b) \in \Delta\mathcal{U}.$$

If there were no uncertainty in the matrix $\bar{A}$, then the constraint can be written as

$$\bar{A}(\Delta x) - \Delta b \in Q - [\bar{A}\bar{x} - \bar{b}] \text{ for all } (\Delta A, \Delta b) \in \Delta\mathcal{U}. \tag{6.1}$$

Recall that the tangential constraint is of the form

$$\bar{A}\gamma + [(\Delta A)\bar{x} - \Delta b] \in T_Q(\bar{A}\bar{x} - \bar{b}) \text{ for all } (\Delta A, \Delta b) \in \Delta\mathcal{U}.$$

If the variable Δx in (6.1) were replaced by γ, then we see that (6.1) is similar to the tangential constraint. In other words, the uncertainty in $\bar{A}$ is transferred to $\bar{b}$ in the tangential constraint through $(\Delta A)\bar{x}$. Tangential problems are still often hard to compute efficiently, but the additional structure may be exploited for designing specialized methods. The case where $T_Q(\bar{A}\bar{x} - \bar{b})$ is polyhedral is just robust linear programming and is easy, while the tangential problem for $L(\Delta\mathcal{U})$ being polyhedral reduces to optimizing over the cone $T_Q(\bar{A}\bar{x} - \bar{b})$, as illustrated below.

Example 6.1. (Polyhedral $L(\Delta\mathcal{U})$) Suppose $L(\Delta\mathcal{U})$ is a polyhedral compact set of the form $L(\Delta\mathcal{U}) = \mathrm{conv}(v_1, \ldots, v_J)$. The tangential problem

$$\begin{aligned} &\min_\gamma h(\gamma) \\ &\text{s.t. } \bar{A}\gamma + L(\Delta\mathcal{U}) \subset T_Q(\bar{A}\bar{x} - \bar{b}), \end{aligned}$$

where h is sublinear, is equivalent to

$$\begin{aligned} &\min_\gamma h(\gamma) \\ &\text{s.t. } \bar{A}\gamma + v_j \in T_Q(\bar{A}\bar{x} - \bar{b}) \text{ for all } 1 \le j \le J. \end{aligned}$$

For this section, we define the sets Φ and Ψ by

$$\begin{aligned} \Phi &:= \{x : \bar{A}x - \bar{b} \in Q\}. \\ \Psi &:= \{y : y + L(\Delta\mathcal{U}) \subset T_Q(\bar{A}\bar{x} - \bar{b})\}. \end{aligned} \tag{6.2}$$

Hence Γ can be written similarly as

$$\begin{aligned} \Gamma &= \{\gamma : \bar{A}\gamma + L(\Delta\mathcal{U}) \subset T_Q(\bar{A}\bar{x} - \bar{b})\} \\ &= \{\gamma : \bar{A}\gamma \in \Psi\}. \end{aligned} \tag{6.3}$$

Recall also that

$$\bar{\Gamma} = \arg\min\{df(\bar{x})(\gamma) \mid \gamma \in \Gamma\}. \tag{6.4}$$

Proposition 6.2. *(Tangent space of feasible set) Suppose Q is Clarke regular at $\bar{A}\bar{x} - \bar{b}$. If there is a vector γ' such that $\bar{A}\gamma' \in \mathrm{rint}(T_Q(\bar{A}\bar{x} - \bar{b}))$, then $T_\Phi(\bar{x}) = \{\gamma : \bar{A}\gamma \in T_Q(\bar{A}\bar{x} - \bar{b})\}$, where Φ is defined in* (6.2)*.*

Proof. This follows directly from [9, Theorem 6.31]. The constraint qualification in that Theorem is equivalent to the condition in our statement through [9, Exercise 6.39(b)]. □

The normal cone $N_\Gamma(\gamma)$ can be estimated from the image $N_\Psi(\bar{A}\gamma)$.

Proposition 6.3. *(Normal cone of tangent feasible set) Suppose there is a vector γ' such that $\bar{A}\gamma' \in \mathrm{int}(T_Q(\bar{A}\bar{x} - \bar{b}))$, and Q is Clarke regular at $\bar{A}\bar{x} - \bar{b}$. Then the normal cone $N_\Gamma(\gamma)$ equals $\{\bar{A}^T v \mid v \in N_\Psi(\bar{A}\gamma)\}$, where Γ and Ψ are defined in* (6.2) *and* (6.3)*.*

Proof. Note that Γ can be written in terms of Ψ as $\Gamma = \{\gamma : \bar{A}\gamma \in \Psi\}$. The result follows directly from [9, Theorem 6.14], though we still have to check the constraint qualification condition there. Through [9, Exercise 6.39(b)], the constraint qualification condition required is that there is a vector γ' such that $\bar{A}\gamma' \in \mathrm{int}(T_\Psi(\bar{A}\gamma))$. Note that the recession cone of Ψ is $T_Q(\bar{A}\bar{x} - \bar{b})$. This means that $T_Q(\bar{A}\bar{x} - \bar{b}) \subset T_\Psi(\bar{A}\gamma)$, which shows that $\bar{A}\gamma' \in \mathrm{int}(T_Q(\bar{A}\bar{x} - \bar{b}))$ implies the constraint qualification condition. The conclusion is straightforward. □

One notices that if there is a vector γ' such that $\bar{A}\gamma' \in \mathrm{int}(T_Q(\bar{A}\bar{x} - \bar{b}))$ and $0 \in \mathrm{int}(L(\Delta\mathcal{U}))$ (which holds when $0 \in \mathrm{int}(\Delta\mathcal{U})$), then the tangential problem is feasible.

If a vector $\bar{x}$ is a solution of the problem $\min\{f(x) \mid x \in D\}$, where f is locally Lipschitz and subdifferentially regular at $\bar{x}$ then it is well known that there is a $c \in \partial f(\bar{x})$ such that $-c \in N_D(\bar{x})$ (see [9, Theorem 8.15] for example). We prove the following lemmas, whose proofs do not seem easy to find.

Lemma 6.4. *(Strict minimizers) Let $D \subset \mathbb{R}^n$ be Clarke regular at $\bar{x}$, $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz and subdifferentially regular at $\bar{x}$. Then $-c \in \mathrm{int}(N_D(\bar{x}))$ for some $c \in \partial f(\bar{x})$ implies that $\bar{x}$ is a strict local minimizer of $\min\{f(x) \mid x \in D\}$.*

Proof. Since $-c \in \mathrm{int}(N_D(\bar{x}))$, there is some $\delta > 0$ such that $-c + \delta\mathbb{B} \subset N_D(\bar{x})$, and so there is a neighborhood O of $\bar{x}$ such that for $x \in D \cap O$,

$$\begin{aligned} \Big[-c + \delta \frac{x - \bar{x}}{|x - \bar{x}|} \Big]^T (x - \bar{x}) & \leq \frac{\delta}{2} |x - \bar{x}| \\ \Rightarrow c^T (x - \bar{x}) & \geq \frac{\delta}{2} |x - \bar{x}|. \end{aligned}$$

By restricting O if necessary, for $x \in D \cap O$, we have $f(x) \geq f(\bar{x}) + df(\bar{x})(x - \bar{x}) - \frac{\delta}{4}|x - \bar{x}|$. The well known characterization of the subderivative in terms of support functions (see for example [4, 6, 9]) gives

$$\begin{aligned} f(x) & \geq f(\bar{x}) + df(\bar{x})(x - \bar{x}) - \frac{\delta}{4}|x - \bar{x}| \\ & = f(\bar{x}) + \max_{v \in \partial f(\bar{x})} v^T (x - \bar{x}) - \frac{\delta}{4}|x - \bar{x}| \\ & \geq f(\bar{x}) + c^T (x - \bar{x}) - \frac{\delta}{4}|x - \bar{x}| \\ & \geq f(\bar{x}) + \frac{\delta}{4}|x - \bar{x}|. \end{aligned}$$

Therefore $\bar{x}$ is a strict local minimizer. □

Lemma 6.5. *(Equivalence of strict minimizer condition) Let D be Clarke regular at $\bar{x}$. When f is $\mathcal{C}^1$ at $\bar{x}$, $\nabla f(\bar{x}) \neq 0$ and $-\nabla f(\bar{x}) \in N_D(\bar{x})$ (which holds when $\bar{x}$ is a local minimizer of $\min\{f(x) \mid x \in D\}$), then the conditions $-\nabla f(\bar{x}) \in \mathrm{int}(N_D(\bar{x}))$ and $\{\nabla f(\bar{x})\}^\perp \cap T_D(\bar{x}) = \{0\}$ are equivalent.*

Proof. The fact that $-\nabla f(\bar{x}) \in N_D(\bar{x})$ when $\bar{x}$ is a local minimizer of $\min\{f(x) \mid x \in D\}$ is well known (see [9, Theorem 6.12] for example).

Suppose that $-\nabla f(\bar{x}) \in \mathrm{int}(N_D(\bar{x}))$, and $v \in \{\nabla f(\bar{x})\}^\perp \cap T_D(\bar{x})$. Then $v \in T_D(\bar{x}) = [N_D(\bar{x})]^*$ (the polar cone of $N_D(\bar{x})$). In other words, $v^T s \leq 0$ for all $s \in N_D(\bar{x})$. If $v \neq 0$, since $-\nabla f(\bar{x}) \in \mathrm{int}(N_D(\bar{x}))$, there is some $\epsilon > 0$ such that $\nabla f(\bar{x}) + \epsilon v \in N_D(\bar{x})$, which gives $v^T(\nabla f(\bar{x}) + \epsilon v) \leq 0$, and thus $v^T \nabla f(\bar{x}) < 0$. Since $v \in \{\nabla f(\bar{x})\}^\perp$, this means $v = 0$, so $\{\nabla f(\bar{x})\}^\perp \cap T_D(\bar{x}) = \{0\}$.

Next, suppose $-\nabla f(\bar{x}) \in N_D(\bar{x}) \backslash \mathrm{int}(N_D(\bar{x}))$. Then $N_{N_D(\bar{x})}(-\nabla f(\bar{x})) \supsetneq \{0\}$. (See for example [9, Exercise 6.19].) Let $w \in N_{N_D(\bar{x})}(-\nabla f(\bar{x})) \backslash \{0\}$. Firstly, $-\lambda \nabla f(\bar{x}) \in N_D(\bar{x})$ for all $\lambda \geq 0$, so $w^T(-\lambda \nabla f(\bar{x}) - (-\nabla f(\bar{x}))) \leq 0$ for all $\lambda \geq 0$, that is $(1 - \lambda) w^T \nabla f(\bar{x}) \leq 0$ for all $\lambda \geq 0$. This implies that $w^T \nabla f(\bar{x}) = 0$, or $w \in \{\nabla f(\bar{x})\}^\perp$. Secondly, $w \in N_{N_D(\bar{x})}(-\nabla f(\bar{x}))$ means that $w^T(s + \nabla f(\bar{x})) \leq 0$ for all $s \in N_D(\bar{x})$. Since $w^T \nabla f(\bar{x}) = 0$, this means that $w^T s \leq 0$ for all $s \in N_D(\bar{x})$, which means that $w \in T_D(\bar{x})$. Therefore, $w \in [\{\nabla f(\bar{x})\}^\perp \cap T_D(\bar{x})] \backslash \{0\}$, which gives us the equivalence between the two conditions. □

In Proposition 6.6 below, we show that the conditions in Lemma 6.4 can give us boundedness information on the robust problem. We also give conditions for which $\bar{\Gamma}$ is a singleton.

Proposition 6.6. *(Conditions for $\bar{\Gamma}$ and optimality) Suppose that Q is Clarke regular at $\bar{A}\bar{x}-\bar{b}$, and there is a vector γ' such that $\bar{A}\gamma' \in \operatorname{int}(T_Q(\bar{A}\bar{x}-\bar{b}))$. Let f be locally Lipschitz and subdifferentially regular at $\bar{x}$, and $\bar{x}$ be a minimizer of the tangential problem (5.2). Recall also Φ, Ψ, Γ and $\bar{\Gamma}$ as defined in (6.2), (6.3) and (6.4).*

(a) *The set $\bar{\Gamma}$ is bounded if there is some $c \in \partial f(\bar{x})$ such that $-c \in \operatorname{int}(N_\Phi(\bar{x}))$.*
(b) *If f is $\mathcal{C}^1$ at $\bar{x}$ and $\nabla f(\bar{x}) \neq 0$, then $\bar{\Gamma}$ is bounded if and only if $-\nabla f(\bar{x}) \in \operatorname{int}(N_\Phi(\bar{x}))$, which is also equivalent to $\{\nabla f(\bar{x})\}^\perp \cap T_\Phi(\bar{x}) = \{0\}$.*
(c) *A feasible $\bar{\gamma} \in \Gamma$ is in $\bar{\Gamma}$ if we can find $c \in \partial f(\bar{x})$ such that $-c \in N_{T_\Phi(\bar{x})}(\bar{\gamma})$. The condition $-c \in N_{T_\Phi(\bar{x})}(\bar{\gamma})$ holds when we can find $\{u_i\}_{i=1}^k$ and $\{v_i\}_{i=1}^k$ such that $v_i \in N_{T_Q(\bar{A}\bar{x}-\bar{b})}(\bar{\gamma}+u_i)$, $u_i \in L(\Delta\mathcal{U})$ and $-c = \sum_{i=1}^k \lambda_i \bar{A}^T v_i$ for some $\lambda_i \geq 0$, $1 \leq i \leq k$. Here, L is the linear map $L(\Delta A, \Delta b) = (\Delta A)\bar{x} - \Delta b$.*
(d) *The set $\bar{\Gamma}$ is a singleton if we can find some $c \in \partial f(\bar{x})$ such that $-c \in \operatorname{int}(N_{T_\Phi(\bar{x})}(\bar{\gamma}))$ for some $\bar{\gamma} \in \bar{\Gamma}$. The condition $-c \in \operatorname{int}(N_{T_\Phi(\bar{x})}(\bar{\gamma}))$ holds when we can find $\{u_i\}_{i=1}^k$ and $\{v_i\}_{i=1}^k$ such that $v_i \in N_{T_Q(\bar{A}\bar{x}-\bar{b})}(\bar{A}\bar{\gamma}+u_i)$, $u_i \in L(\Delta\mathcal{U})$, $k = \dim(Q)$, $\{\bar{A}^T v_i\}_{i=1}^k$ is linearly independent, and $-c = \sum_{i=1}^k \lambda_i \bar{A}^T v_i$, where $\lambda_i > 0$ for all $1 \leq i \leq k$.*

Proof. **Part (a):** Seeking a contradiction, suppose that Γ is unbounded, so there is a sequence $\{\gamma_i\}$ of solutions of minimizers of (5.2) such that $|\gamma_i| \to \infty$, with $\frac{\gamma_i}{|\gamma_i|} \to \gamma_\infty$. Note

$$\begin{aligned} \gamma_i &\in \Gamma \\ &= \{\gamma : \bar{A}\gamma + L(\Delta\mathcal{U}) \subset T_Q(\bar{A}\bar{x}-\bar{b})\} \\ &\subset \{\gamma : \bar{A}\gamma \in T_Q(\bar{A}\bar{x}-\bar{b})\} \\ &= T_\Phi(\bar{x}), \end{aligned}$$

so $\gamma_\infty \in T_\Phi(\bar{x})$. On the other hand, since $df(\bar{x})(\cdot)$ is continuous and positively homogeneous, we have

$$\begin{aligned} df(\bar{x})(\gamma_\infty) &= \lim_{i\to\infty} df(\bar{x})\Big(\frac{\gamma_i}{|\gamma_i|}\Big) \\ &= \lim_{i\to\infty} \frac{1}{|\gamma_i|} df(\bar{x})(\gamma_i) \\ &= 0. \end{aligned}$$

The well known characterization of the subderivative in terms of support functions (see for example [4, 6, 9]) gives

$$\begin{aligned} \max_{v\in\partial f(\bar{x})} v^T\gamma_\infty &= df(\bar{x})(\gamma_\infty) = 0 \\ \Rightarrow c^T\gamma_\infty &\leq 0. \end{aligned}$$

But $-c \in \operatorname{int}(N_\Phi(\bar{x}))$, so $c^T\gamma_\infty > 0$. This contradiction tells us that the set $\bar{\Gamma}$ is bounded as needed.

Part (b): In view of part (a) and Lemma 6.5, we just need to prove that if $\bar{\Gamma}$ is bounded, then $\{\nabla f(\bar{x})\}^\perp \cap T_\Phi(\bar{x}) = \{0\}$. We prove the contrapositive. Suppose $w \in [\{\nabla f(\bar{x})\}^\perp \cap T_\Phi(\bar{x})]\backslash\{0\}$. Then $\nabla f(\bar{x})^T w = 0$, and $w \in T_\Phi(\bar{x})$, that is $\bar{A}w \in T_Q(\bar{A}\bar{x} - \bar{b})$ by Proposition 6.2. Let $\bar{\gamma}$ be some element in $\bar{\Gamma}$. Then

$$\begin{aligned} \bar{A}(\lambda w + \bar{\gamma}) + L(\Delta\mathcal{U}) &= \bar{A}\lambda w + [\bar{A}\bar{\gamma} + L(\Delta\mathcal{U})] \\ &\subset T_Q(\bar{A}\bar{x} - \bar{b}) + T_Q(\bar{A}\bar{x} - \bar{b}) \\ &= T_Q(\bar{A}\bar{x} - \bar{b}). \end{aligned}$$

Therefore $\bar{\gamma} + \lambda w \in \bar{\Gamma}$ for all $\lambda \geq 0$, which shows that $\bar{\Gamma}$ is unbounded, concluding our proof.

Parts (c), (d): Note that

$$\begin{aligned} \Psi &= \{\gamma : \gamma + u \in T_Q(\bar{A}\bar{x} - \bar{b}) \text{ for all } u \in L(\Delta\mathcal{U})\} \\ &\subset \{\gamma : \gamma + u_i \in T_Q(\bar{A}\bar{x} - \bar{b}) \text{ for } 1 \leq i \leq k\} \\ &= \bigcap_{i=1}^{k} [T_Q(\bar{A}\bar{x} - \bar{b}) - u_i], \end{aligned}$$

so

$$\begin{aligned} N_\Psi(\bar{A}\bar{\gamma}) &\supset N_{\cap_{i=1}^k [T_Q(\bar{A}\bar{x}-\bar{b})-u_i]}(\bar{A}\bar{\gamma}) \\ &\supset [N_{T_Q(\bar{A}\bar{x}-\bar{b})-u_1}(\bar{A}\bar{\gamma})] + \cdots + [N_{T_Q(\bar{A}\bar{x}-\bar{b})-u_k}(\bar{A}\bar{\gamma})] \\ &\qquad \text{(by [9, Theorem 6.42])} \\ &= [N_{T_Q(\bar{A}\bar{x}-\bar{b})}(\bar{A}\bar{\gamma} + u_1)] + \cdots + [N_{T_Q(\bar{A}\bar{x}-\bar{b})}(\bar{A}\bar{\gamma} + u_k)] \\ &\supset \left\{ \sum_{i=1}^{k} \lambda_i v_i \mid \lambda_i \geq 0 \right\}. \end{aligned}$$

Therefore, by Proposition 6.3,

$$\begin{aligned} N_\Gamma(\bar{\gamma}) &= \{\bar{A}^T y : y \in N_\Psi(\bar{A}\bar{\gamma})\} \\ &\supset \left\{ \sum_{i=1}^{k} \lambda_i \bar{A}^T v_i : \lambda_i \geq 0 \right\}. \end{aligned}$$

For part (c), the condition stated is equivalent to the existence of $c \in \partial f(\bar{x})$ such that $-c \in \{\sum_{i=1}^k \lambda_i v_i \mid \lambda_i \geq 0\}$, which implies $-c \in N_\Gamma(\bar{\gamma})$, which in turn implies $\bar{\gamma} \in \bar{\Gamma}$. Part (d) follows by applying Lemma 6.4. □

The conditions (c) and (d) in Proposition 6.6 can be helpful for designing numerical methods for solving the tangential problem. Due to Clarke regularity, the tangential problem is convex. However, the problem of determining whether a point is feasible is not necessarily easy.

The result corresponding to conditions (c) and (d) in Proposition 6.6 can also be generalized for robust optimization in general. The following result on normal cones in robust optimization combined with results on optimality of nonlinear programs (in Lemma 6.4 for example) gives us the optimality conditions. The proof of the following result is a direct application of [9, Theorems 6.14 and 6.42] similar to the proofs of Propositions 6.3 and 6.6, so we shall only state the result.

Proposition 6.7. *(Normal cones in robust optimization) For $f : \mathbb{R}^n \to \mathbb{R}$, consider the robust optimization problem*

$$\begin{aligned} &\min_x f(x) \\ &\text{s.t. } [\bar{A} + \Delta A]x - [\bar{b} - \Delta b] \in Q \text{ for all } (\Delta A, \Delta b) \in \Delta\mathcal{U}. \end{aligned}$$

Let the sets $\Omega(\Delta A, \Delta b) \subset \mathbb{R}^n$ *and* $\Xi \subset \mathbb{R}^n$ *be defined by*

$$\begin{aligned} \Omega(\Delta A, \Delta b) &:= \{x \mid [\bar{A} + \Delta A]x - [\bar{b} - \Delta b] \in Q\}, \\ \text{and } \Xi &:= \{x \mid [\bar{A} + \Delta A]x - [\bar{b} - \Delta b] \in Q \text{ for all } (\Delta A, \Delta b) \in \Delta\mathcal{U}\} \\ &= \bigcap_{(\Delta A, \Delta b) \in \Delta\mathcal{U}} \Omega(\Delta A, \Delta b). \end{aligned}$$

Let $\bar{x} \in \Xi$.

(1) *If* Q *is Clarke regular at* $[\bar{A} + \Delta A]\bar{x} - [\bar{b} - \Delta b]$ *and the only vector* $y \in N_Q([\bar{A} + \Delta A]\bar{x} - [\bar{b} - \Delta b])$ *for which* $[\bar{A} + \Delta A]^T y = 0$ *is* $y = 0$, *then*

$$N_{\Omega(\Delta A, \Delta b)}(\bar{x}) = \{[\bar{A} + \Delta A]^T y \mid y \in N_Q([\bar{A} + \Delta A]\bar{x} - [\bar{b} - \Delta b])\}.$$

(2) *For any finite set* $\{(\Delta A_i, \Delta b_i)\}_{i \in I} \subset \Delta\mathcal{U}$ *such that* $\bar{x}$ *is Clarke regular at* $\Omega(\Delta A_i, \Delta b_i)$ *for all* $i \in I$, *the normal cone* $N_\Xi(\bar{x})$ *satisfies* $N_\Xi(\bar{x}) \supset \sum_{i \in I} N_{\Omega(\Delta A_i, \Delta b_i)}(\bar{x})$.

7. Addition of uncertainty sets in the tangential problem

For much of this section we focus on the tangential robust problem on addition of uncertainty sets. More specifically, we ask what we can say about the tangential problem with uncertainty set $\lambda_1 \Delta\mathcal{S}_1 + \lambda_2 \Delta\mathcal{S}_2$ given knowledge of the optimal solutions of the tangential problem for the uncertainty sets $\Delta\mathcal{S}_1$ and $\Delta\mathcal{S}_2$. Such a problem can arise from having to considering robust optimization problems with errors which are a sum of two or more unknown sources.

We begin with some elementary properties.

Proposition 7.1. *(Elementary properties of uncertainty sets) Suppose* Q *is Clarke regular at* $\bar{A}\bar{x} - \bar{b}$. *For an uncertainty set* $\Delta\mathcal{U}$, *suppose the solution of the tangential problem is defined by*

$$\begin{aligned} v(\Delta\mathcal{U}) &:= \min_\gamma h(\gamma) \\ &\qquad \text{s.t. } \bar{A}\gamma + [(\Delta A)\bar{x} - \Delta b] \in T_Q(\bar{A}\bar{x} - \bar{b}) \text{ for all } (\Delta A, \Delta b) \in \Delta\mathcal{U}, \end{aligned}$$

where h *is sublinear. Then*

- *If* $\Delta\mathcal{S}' = \lambda\Delta\mathcal{S}$, *then* $v(\Delta\mathcal{S}') = \lambda v(\Delta\mathcal{S})$.
- *If* $\Delta\mathcal{S}' \supset \Delta\mathcal{S}$, *then* $v(\Delta\mathcal{S}) \leq v(\Delta\mathcal{S}')$.
- *If* $h(\gamma) = c^T\gamma$ *and* $\Delta\mathcal{S}' = \Delta\mathcal{S} + \{\bar{s}\}$, *then* $v(\Delta\mathcal{S}') = v(\Delta\mathcal{S}) + c^T\bar{s}$.

We have the following result to study how set addition affects the solution to the tangential problem.

Proposition 7.2. *(Set addition) Recall the definition of* $v(\Delta\mathcal{U})$ *in Proposition 7.1. Suppose* $\Delta\mathcal{S} = \lambda_1\Delta\mathcal{S}_1 + \lambda_2\Delta\mathcal{S}_2$, *where* $\lambda_1, \lambda_2 \geq 0$. *Then* $v(\Delta\mathcal{S}) \leq \lambda_1 v(\Delta\mathcal{S}_1) + \lambda_2 v(\Delta\mathcal{S}_2)$.

Proof. In view of Proposition 7.1, we only need to prove the case for $\lambda_1 = \lambda_2 = 1$. Suppose $\bar{\gamma}_1$ and $\bar{\gamma}_2$ are solutions to

$$\begin{aligned} &\min_\gamma h(\gamma) \\ &\text{s.t. } \bar{A}\gamma + [(\Delta A)\bar{x} - \Delta b] \in T_Q(\bar{A}\bar{x} - \bar{b}) \text{ for all } (\Delta A, \Delta b) \in \Delta\mathcal{S}_i \end{aligned}$$

for $i = 1, 2$. If $(\Delta A, \Delta b) \in \Delta\mathcal{S}_1 + \Delta\mathcal{S}_2$, then

$$[(\Delta A)\bar{x} - \Delta b] = [(\Delta A_1)\bar{x} - \Delta b_1] + [(\Delta A_1)\bar{x} - \Delta b_2],$$

for some $[(\Delta A_i)\bar{x} - \Delta b_i] \in \Delta\mathcal{S}_i$ for $i = 1, 2$. This means that

$$\begin{aligned} & \bar{A}(\bar{\gamma}_1 + \bar{\gamma}_2) - [(\Delta A)\bar{x} - \Delta b] \\ = & \big[\bar{A}\bar{\gamma}_1 - [(\Delta A_1)\bar{x} - \Delta b_1]\big] + \big[\bar{A}\bar{\gamma}_2 - [(\Delta A_2)\bar{x} - \Delta b_2]\big] \\ \in & T_Q(\bar{A}\bar{x} - \bar{b}) + T_Q(\bar{A}\bar{x} - \bar{b}) \\ = & T_Q(\bar{A}\bar{x} - \bar{b}). \end{aligned}$$

So $\bar{\gamma}_1 + \bar{\gamma}_2$ is a feasible, though not necessarily optimal, solution to the tangential problem where the uncertainty set is $\Delta\mathcal{S}_1 + \Delta\mathcal{S}_2$, which shows that $v(\Delta\mathcal{S}) \leq h(\bar{\gamma}_1 + \bar{\gamma}_2) \leq h(\bar{\gamma}_1) + h(\bar{\gamma}_2) = v(\Delta\mathcal{S}_1) + v(\Delta\mathcal{S}_2)$. □

In a nondegenerate linear programming problem, we do have equality in Proposition 7.2. The assumption that $\bar{A}$ is an invertible square matrix below is not restrictive, because this is exactly what happens in a nondegenerate linear program.

Proposition 7.3. *(Set addition in nondegenerate linear programming) Consider the tangential problem having a linear programming structure*

$$\begin{aligned} v(\Delta\mathcal{U}) \quad := \quad & \min_{\gamma} c^T\gamma \\ & \text{s.t. } \bar{A}\gamma + [(\Delta A)\bar{x} - \Delta b] \leq 0 \text{ for all } (\Delta A, \Delta b) \in \Delta\mathcal{U}. \end{aligned}$$

Assume that $\bar{A}$ is square and invertible, and $-c = \sum_{i=1}^{k} \lambda_i \bar{A}_i$ for some $\lambda_i > 0$, $1 \leq i \leq k$. Then $v(\Delta\mathcal{S}_1 + \Delta\mathcal{S}_2) = v(\Delta\mathcal{S}_1) + v(\Delta\mathcal{S}_2)$.

Proof. Recall that feasibility can be rewritten as

$$\bar{A}_i\gamma + \max_{(\Delta A,\Delta b)\in\Delta\mathcal{U}} [(\Delta A_i)\bar{x} - \Delta b_i] \leq 0 \text{ for all } i.$$

Form the vector $w(\Delta\mathcal{U}) \in \mathbb{R}^m$ by

$$[w(\Delta\mathcal{U})]_i := -\max_{(\Delta A,\Delta b)\in\Delta\mathcal{U}} [(\Delta A_i)\bar{x} - \Delta b_i].$$

The value $v(\Delta\mathcal{U})$ is equal to $c^T\bar{A}^{-1}w(\Delta\mathcal{U})$. The condition on $-c$ in the statement assures that this minimizer is unique through convexity and Lemma 6.4. It is clear that for each i, we have

$$\begin{aligned} & \max_{(\Delta A,\Delta b)\in\Delta\mathcal{S}_1+\Delta\mathcal{S}_2} [(\Delta A_i)\bar{x} - \Delta b_i] \\ = & \max_{(\Delta A,\Delta b)\in\Delta\mathcal{S}_1} [(\Delta A_i)\bar{x} - \Delta b_i] + \max_{(\Delta A,\Delta b)\in\Delta\mathcal{S}_2} [(\Delta A_i)\bar{x} - \Delta b_i], \end{aligned}$$

which implies that $w(\Delta\mathcal{S}_1 + \Delta\mathcal{S}_2) = w(\Delta\mathcal{S}_1) + w(\Delta\mathcal{S}_2)$. This immediately gives $v(\Delta\mathcal{S}_1 + \Delta\mathcal{S}_2) = v(\Delta\mathcal{S}_1) + v(\Delta\mathcal{S}_2)$ as needed. □

The next step is to ask whether the property in Proposition 7.3 is satisfied for problems of the form

$$\begin{aligned} & \min_{\gamma} c^T\gamma \\ & \text{s.t. } \bar{A}\gamma + [(\Delta A)\bar{x} - \Delta b] \in Q \text{ for all } (\Delta A, \Delta b) \in \Delta\mathcal{U}, \end{aligned}$$

where Q is a convex cone. If $Q = \{y \mid \tilde{A}y \leq 0\}$, where $\tilde{A}$ is an invertible square matrix, then the constraint above can be transformed into

$$\tilde{A}\big[\bar{A}\gamma + [(\Delta A)\bar{x} - \Delta b]\big] \leq 0 \text{ for all } (\Delta A, \Delta b) \in \Delta\mathcal{U},$$

and we can apply Proposition 7.3 on $\tilde{A}\bar{A}$.

In the general case, the equality $v(\Delta\mathcal{S}_1+\Delta\mathcal{S}_2)=v(\Delta\mathcal{S}_1)+v(\Delta\mathcal{S}_2)$ may not hold. We give two examples to illustrate this. In the first example, we have a degenerate linear program, while in the second example, we have different cones for a conic programming problem.

Example 7.4. (Inequality in sums of uncertainty sets 1) Consider the problem

$$\begin{array}{rcl}\bar{v}(\Delta\mathcal{U}) & := & \min_{x\in\mathbb{R}^2} x_2 \\ & & \text{s.t. } (\bar{A}+\Delta A)x-(\mathbf{0}+\Delta b)\leq 0 \text{ for all } (\Delta A,\Delta b)\in\Delta\mathcal{U}, \\ & & \text{,where } \bar{A}=\begin{pmatrix}-1 & -1\\ 1 & -1\\ 0.5 & -1\end{pmatrix}, \bar{b}=\mathbf{0}.\end{array}$$

The optimal solution to the nonrobust problem is $\bar{x}=0$, and the tangential problem is

$$\begin{array}{rcl}v(\Delta\mathcal{U}) & := & \min_{\gamma\in\mathbb{R}^2} \gamma_2 \\ & & \text{s.t. } \bar{A}\gamma+[(\Delta A)\mathbf{0}-\Delta b]\leq 0 \text{ for all } (\Delta A,\Delta b)\in\Delta\mathcal{U}.\end{array}$$

We illustrate this example in Figure 7.1. Let the uncertainty sets $\Delta\mathcal{S}_1$ and $\Delta\mathcal{S}_2$ be defined by

$$\begin{array}{rcl}\Delta\mathcal{S}_1 & := & \{\mathbf{0}\}\times\{\Delta b:\Delta b_1=0,|\Delta b_2|\leq 4\epsilon,\Delta b_3=0\}\\ \Delta\mathcal{S}_2 & := & \{\mathbf{0}\}\times\{\Delta b:\Delta b_1=0,\Delta b_2=0,|\Delta b_3|\leq 3\epsilon\}\\ \Delta\mathcal{S}_1+\Delta\mathcal{S}_2 & = & \{\mathbf{0}\}\times\{\Delta b:\Delta b_1=0,|\Delta b_2|\leq 4\epsilon,|\Delta b_3|\leq 3\epsilon\}.\end{array}$$

To find $v(\Delta\mathcal{S}_1)$, we note that tangential problem becomes

$$\begin{pmatrix}-1 & -1\\ 1 & -1\\ 0.5 & -1\end{pmatrix}\gamma\leq\begin{pmatrix}0\\ -4\epsilon\\ 0\end{pmatrix}.$$

The first two rows are the active constraints, which gives a solution of $\bar{\gamma}(\Delta\mathcal{S}_1)=(-2\epsilon,2\epsilon)$ and $v(\Delta\mathcal{S}_1)=2\epsilon$. To find $v(\Delta\mathcal{S}_2)$, a similar set of calculations shows that the first and third constraints are the active constraints, which gives $\bar{\gamma}(\Delta\mathcal{S}_2)=(-2\epsilon,2\epsilon)$ and $v(\Delta\mathcal{S}_2)=2\epsilon$. Similarly, $v(\Delta\mathcal{S}_1+\Delta\mathcal{S}_2)=2\epsilon$, and all constraints are active. We have $v(\Delta\mathcal{S}_1+\Delta\mathcal{S}_2)<v(\Delta\mathcal{S}_1)+v(\Delta\mathcal{S}_2)$ as needed.

Here is a second example for the case when the cone Q is slightly more complicated.

Example 7.5. (Inequality in sums of uncertainty sets 2) Consider the problem

$$\begin{array}{rcl}\bar{v}(\Delta\mathcal{U}) & := & \min_{\gamma\in\mathbb{R}^3} x_3 \\ & & \text{s.t.}(I+\Delta A)x-(\mathbf{0}+\Delta b)\in Q \text{ for all } (\Delta A,\Delta b)\in\Delta\mathcal{U}.\end{array}$$

Here, the matrix $\bar{A}=I$ is the 3×3 identity matrix, $\bar{b}\in\mathbb{R}^3$ is the zero vector, and the convex cone Q is defined by

$$Q:=\{x\in\mathbb{R}^3: x_3\geq\max(|x_1|,|x_2|)\}.$$

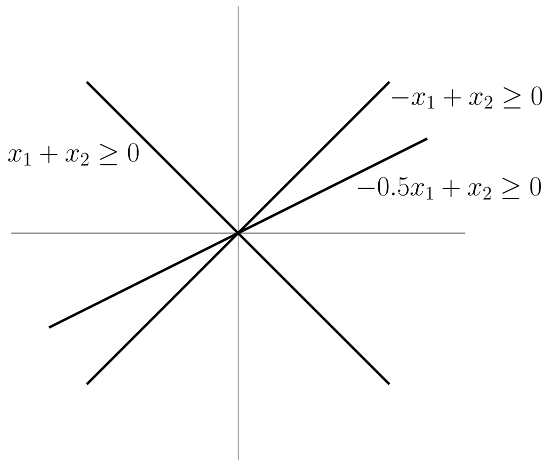

FIGURE 7.1. Illustration of Example 7.4.

The optimal solution to the nonrobust problem is $\bar{x} = \mathbf{0}$, and the tangential problem is

$$\begin{aligned} v(\Delta\mathcal{U}) \quad := \quad & \min_{\gamma \in \mathbb{R}^3} \gamma_3 \\ & \text{s.t.} \gamma + [(\Delta A)\mathbf{0} - \Delta b] \in T_Q(\mathbf{0}) \text{ for all } (\Delta A, \Delta b) \in \Delta\mathcal{U}. \end{aligned}$$

Let $a < b$ and the sets $\mathcal{S}_1$ and $\mathcal{S}_2$ be defined by

$$\begin{aligned} \Delta\mathcal{S}_1 \quad &:= \quad \{\mathbf{0}\} \times \left\{ \Delta b \in \mathbb{R}^3 : |\Delta b_1| \leq \frac{a}{2}, |\Delta b_2| \leq \frac{b}{2}, |\Delta b_3| \leq \frac{\delta}{2} \right\} \\ \Delta\mathcal{S}_2 \quad &:= \quad \{\mathbf{0}\} \times \left\{ \Delta b \in \mathbb{R}^3 : |\Delta b_1| \leq \frac{b}{2}, |\Delta b_2| \leq \frac{a}{2}, |\Delta b_3| \leq \frac{\delta}{2} \right\} \\ \Delta\mathcal{S}_1 + \Delta\mathcal{S}_2 \quad &= \quad \{\mathbf{0}\} \times \left\{ \Delta b \in \mathbb{R}^3 : |\Delta b_1| \leq \frac{a+b}{2}, |\Delta b_2| \leq \frac{a+b}{2}, |\Delta b_3| \leq \delta \right\}. \end{aligned}$$

See Figure 7.2 for an illustration of the convex cone Q, and the projection of $\Delta\mathcal{S}_1$, $\Delta\mathcal{S}_2$ and $\Delta\mathcal{S}_1 + \Delta\mathcal{S}_2$ onto the 2-dimensional space corresponding to the first 2 coordinates in $\mathbb{R}^3$. It is elementary that $v(\Delta\mathcal{S}_1) = v(\Delta\mathcal{S}_2) = \frac{b}{2} + \frac{\delta}{2}$, and $v(\Delta\mathcal{S}_1 + \Delta\mathcal{S}_2) = \frac{a+b}{2} + \delta$. This gives $v(\Delta\mathcal{S}_1 + \Delta\mathcal{S}_2) < v(\Delta\mathcal{S}_1) + v(\Delta\mathcal{S}_2)$. If the calculations had been performed with the second order cone Q' defined by

$$Q' := \left\{ x \in \mathbb{R}^3 : x_3 \geq \sqrt{x_1^2 + x_2^2} \right\}$$

instead, then we also get the conclusion $v(\Delta\mathcal{S}_1 + \Delta\mathcal{S}_2) < v(\Delta\mathcal{S}_1) + v(\Delta\mathcal{S}_2)$.

ACKNOWLEDGMENTS

I am grateful to Henry Wolkowicz for conversations and his initial ideas in [10] that led to me studying this problem, and to the Natural Sciences Engineering Research Council (Canada) for supporting the research. The research for this paper was carried out in the Department of Combinatorics and Optimization, University of Waterloo, when I was on a postdoctoral position there, and I gratefully acknowledge them for providing a splendid working environment.

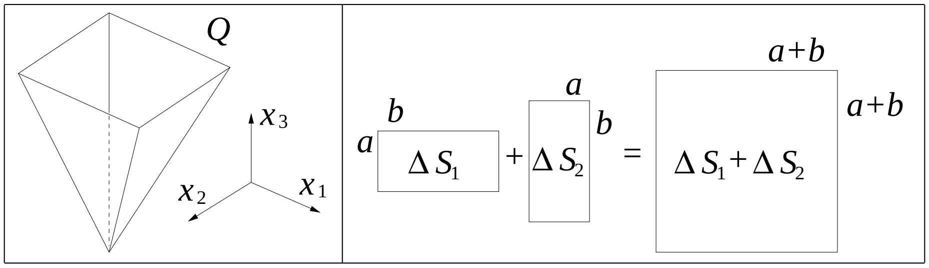

FIGURE 7.2. These diagrams illustrate Example 7.5. The figure on the left illustrates the convex cone Q in $\mathbb{R}^3$, while the figure on the right illustrates the first 2 coordinates for the shapes $\Delta\mathcal{S}_1$ and $\Delta\mathcal{S}_2$.

REFERENCES

[1] F. Alizadeh and D. Goldfarb, *Second-order cone programming,* Mathematical Programming, Ser B 95:3–51 (2003).
[2] H. Attouch and R. J.-B. Wets, *Approximation and convergence in nonlinear optimization*, in Nonlinear Programming 4, edited by O. Mangasarian, R. Meyer and S. Robinson, pp. 367–395, Academic Press, New York, 1981.
[3] A. Ben-Tal, L. El Ghaoui and A. Nemirovski. *Robust Optimization*, Princeton Series in Applied Mathematics, Princeton, 2009.
[4] F.H. Clarke. *Optimization and Nonsmooth Analysis.* Wiley, New York, 1983. Republished as Vol. 5, Classics in Applied Mathematics, SIAM, 1990.
[5] B.S. Mordukhovich. *Variational Analysis and Generalized Differentiation I and II.*, Grundlehren der mathematischen Wissenschaften, Vols 330 & 331, Springer, Berlin, 2006.
[6] R. T. Rockafellar, *Convex Analysis*, Princeton, 1970.
[7] R. T. Rockafellar, *Clarke's tangent cones and the boundaries of closed sets in* $\mathbb{R}^n$, Nonlinear Analysis: Theory, Methods and Applications, 3, 145–154, 1979.
[8] R. T. Rockafellar and R. J.-B. Wets, *Variational Systems, an introduction*, in Multifunctions and Integrands: Stochastic Analysis, Approximation and Optimization, edited by G. Salinetti, Lecture Notes in Mathematics, 1091, pp. 1–54, Springer Verlag, Berlin, 1984.
[9] R.T. Rockafellar and R.J.-B. Wets. *Variational Analysis*, Grundlehren der mathematischen Wissenschaften, Vol 317, Springer, Berlin, 1998.
[10] H. Wolkowicz and W.L.N. Yeung, *An alternative approach to sensitivity analysis in linear programming*, 2009 (unpublished).

Current address: Department of Combinatorics and Optimization, University of Waterloo, 200 University Ave West, Waterloo, ON, Canada N2L 3G1.

E-mail address: chj2pang@math.uwaterloo.ca