Journal of Convex Analysis
Volume 26 (2019), No. 2, 605–616

Uncertainty Relations in Optimization

A. I. Propoi
Federal Research Center for Computer Science and Control,
Russian Academy of Sciences, Moscow 119333, Russia
propoi@gmail.com

Received: November 08, 2017
Revised manuscript received: June 29, 2018
Accepted: July 11, 2018

Uncertainty relations for a pair of dual (conjugate) cones are established. For this purpose a measure of dispersion for a cone as the norm (in some metric) of the deviation of a ray from the axis of a cone is introduced. An application of the uncertainty relations to the epsilon-subdifferential is considered.

Keywords: Convex analysis, uncertainty relations, dual conjugate cones, epsilon-subdifferential.
2010 Mathematics Subject Classification: 68Q25, 68R10, 68U05.

1. Introduction

A basis of optimization methods is the duality relation between a feasible direction and a linear functional that approximates an objective function [3], [4], [5], [7].

If the set of feasible directions and the objective function are smooth and strictly convex, then the duality relation is a bijection: each linear functional uniquely defines the optimal (dual) direction and each direction is related to the dual linear functional. However, in many optimization problems it is impossible to specify a direction or a linear functional. In this case we have a pair of dual (conjugate) cones.

Geometrical reasoning shows that the "narrower" a cone, the "wider" the conjugate cone. Thus, to formalize this relation it is necessary to define the parameters that describe the dispersion of a cone.

The paper is organized as follows. In Section 2 we introduce a cutting hyperplane instead of the supporting hyperplane. In this case the supporting vector is replaced by a cone.

In Section 3 we study properties of dual cones. First, we consider a generalization of the Riesz representation theorem and show how a linear functional can be represented by a vector dual to this functional in some metric (inner product). Then a pair of dual cones in the dual spaces can be replaced by a pair of dual cones in one space. In conclusion of Section 3 a measure of dispersion for a cone is introduced. For this we fix an axis of the cone (it can be any interior ray of the cone in some metric) and the deviation $\|w\|$ from the axis measured at some

ISSN 0944-6532 / $ 2.50

fixed distance s from the vertex of the cone. Thus the dispersion of a cone can be characterized by these (interrelated) indices: either by $\|w\|$ or by s.

After establishing the measure of dispersion for a cone the derivation of the uncertainty relations is straightforward (Section 4). We derive the uncertainty relations in two cases: for the pair of dual cones and for a cone interval. In the latter case the uncertainty relation has a clear interpretation in optimization methods: the limit of the uncertainty relations is defined by the chosen distance from the initial (current) point to the final point.

In Section 5 we consider a different approach to the uncertainty relations where the uncertainty is described by a ball at the end (or at the beginning) of a vector. We establish the uncertainty relations for this case and show that they are a consequence of the geometric structure of the cone interval defined by the dual cones.

In Section 6 we analyze the connection of the uncertainty relations to the notion of ε-subdifferential [12].

2. Statement of the problem

Before the statement of the problem we consider two examples.

Example 2.1. Let the state space of a control system (CS) be a smooth manifold M; a point $x \in M$ be a state of CS. CS is defined by the set of feasible velocities $B(x)$ on the tangent bundle TM. If M is the Riemann manifold, $B(x)$ is an ellipsoid in the tangent space T_xM.

A *local approximation* (at the state x or at the time t) of an objective function is defined by a linear functional p^*. Then the optimization problem (in T_xM)

$$\sup_{u \in B(x)} p^*(u) \tag{1}$$

determines the optimal velocity from the point x for the given p^*.

In static optimization problems, $p^*(x) = d\Phi(x)$, $d\Phi(x)$ is the first differential of the objective function $\Phi(x)$ (if it exists) [3], [5], [7], [8]. In dynamic optimization problems the evolution of u and p^* is described by the conjugate equations; Equation (1) is the Pontryagin's maximum principle, p^* is the generalized momentum [10]. □

Example 2.2. Let the vector ψ in a real Hilbert space define the trajectory of CS. For a Hilbert space a natural assumption is to define a feasible set by a quadratic form in H, and an objective function – by a linear functional in H^*. Then the choice of the optimal trajectory is reduced to the problem of the type (1) in the space H. Such an approach is used in model predictive control [11]. □

Taking into account these examples we state the problem.

We consider a real finite-dimensional Hilbert space X. In this space the feasible set $B = \{u \,|\, (u, Gu) \leq 1\}$ is given, where G is a (bounded) positive self-adjoint operator in X. This operator defines the metric in X: $\|u\|_G^2 = (u, Gu)$. We will use G as the label of this metric, g is the label of original metric (u, v) of the space X. The Hilbert space with the metrics g and G will be denoted as X and $X(G)$. The set B is the ellipsoid in the metric g and the ball in the metric G.

We fix a linear functional $p^* \in X^*$ on X and, instead of the optimization problem (1), consider the following system of equations

$$(u, Gu) \leq 1, \tag{2}$$

$$p^*(u) = s, \tag{3}$$

where the value $s > 0$ is given. For a normalized linear functional: $0 < s \leq 1$.

The system (2) and (3) defines the cross-section of the ellipsoid (2) by the hyperplane (3). The set of all vectors λu, where u satisfies (2), (3) and $\lambda \geq 0$, defines the proper cone [3] denoted as $C_s(p^*)$.

We consider the value of $p^*(u)$ as the goal for the choice of the (optimal) direction; correspondingly, $C_s(p^*)$ is the cone of improving directions, because $p^*(u) \geq 0$ for all $u \in C_s(p^*)$.

If $s = 1$, then (3) defines the hyperplane that supports the ellipsoid (2) at the vector v_p: $p^*(v_p) = (v_p, Gv_p) = 1$. In this case the cone $C_s(p^*)$ is reduced to the ray λv_p, $\lambda \geq 0$, which determines the optimal (dual) direction for the given p^*. Thus, the specification of s, $0 < s \leq 1$, produces the uncertainty in the choice of a direction (for the given goal p^*).

These constructions can be repeated in the dual space X^*:

$$(p^*, Gp^*) \leq 1, \tag{4}$$

$$p^*(v) = s_*, \tag{5}$$

where $v \in X$ and $s_* > 0$ are given.

The corresponding cone $C^*_{s_*}(v)$ in X^* characterizes the uncertainty in the choice of the goal p^* (for the given direction v).

Problem (1) or the system (2), (3) for $s = 1$, define the duality relation between the given goal p^* and the optimal direction v_p. This dual relation is generalized if $0 < s \leq 1$ and leads to the dual (conjugate) cones.

Geometrical reasoning shows that the "narrower" the cone C, the "wider" the conjugate cone C^*. In the next sections we will formalize this relation.

3. Dual cones

Let p^* be a linear functional on X. The linear functional p^* defines the decomposition of the space X into the direct sum: $X = X_v \oplus N(p^*)$, where X_v is the one-dimensional subspace of X spanned by the vector v, $N(p^*)$ is the null space of p^*: $p^*(w) = 0$, $w \in N(p^*)$.

Any vector $u \in X$ can be represented as

$$u = sv + w,\ w \in N(p^*). \tag{6}$$

We will choose the vector $v = p$ orthogonal to the subspace $N(p^*)$: $(p, w) = 0$. In this case

$$s = \frac{p^*(u)}{p^*(p)} = \frac{(p, u)}{(p, p)}. \tag{7}$$

In particular, $p^*(u) = (p, u)$, if $p^*(p) = (p, p)$ (e.g., $p^*(p) = 1$ and $(p, p) = 1$).

Equation (7) is the Riesz representation theorem: a given linear functional p^* is related to the vector $v = v_p$ in such a way that (7) is true for all $u \in X$. Conversely, a given vector $v \in X$ is related to p^*_v for which (7) is true.

Equation (7) can be generalized for a metric G:

$$\frac{p^*(u)}{p^*(v)} = \frac{(Gv, u)}{(Gv, v)}. \tag{8}$$

In this case the vector v is orthogonal to the subspace $N(p^*)$ in the metric G: $(Gv, w) = 0$.

The vectors in X and X^* connected by (7) or (8) are dual; correspondingly, a mapping $X \to X^*$ is a duality mapping.

For the metric G the duality mapping is defined by

$$p = Gv. \tag{9}$$

Equation (9) also follows from the solution of the optimization problem (1) with the feasible set defined in (2). Thus, $p^*(u) = (p, u) = (Gv, u)$ for all $u \in X$ (in the normalized case).

The vectors p and v in (9) can be considered as two representations of the linear functional p^* (correspondingly, in metrics g and G).

Remark 3.1. For a Banach space the duality mapping $X \to X^*$ is defined as

$$D(v) = \{p^* \in X^* \mid p^*(v) = \|p^*\|\|v\|,\ \|p^*\| = \|v\|\}.$$

and in general it is multivalued [4]. For example, $\|v\| = \|v\|_B$, where $\|v\|_B$ is the Minkowski functional for the set B, which is assumed to be convex and central-symmetric [12]. □

We will now generalize the notion of duality for cones [3]. Let C and C^* be the closed convex cones in X and X^*.

Definition 3.2. The cone C^* is *dual* to the cone C if

$$C^* = \{p^* \in X^* \mid p^*(u) \geq 0,\ u \in C\}. \tag{10}$$

Hence,
$$C = \{x \in X \mid p^*(u) \geq 0,\ p^* \in C^*\}. \tag{11}$$

Thus, the cone C contains all directions u that improve values $p^*(u)$ for all $p^* \in C^*$.

Above we have shown that a bilinear form on $X^* \times X$ is related to a bilinear form on $X \times X$. Therefore the pair of dual cones in $X^* \times X$ can be related to the pair of dual cones in the space X:

$$\bar{C} = \{p \,|\, (p, u) \geq 0,\ u \in C\}, \tag{12}$$

$$C = \{u \,|\, (p, u) \geq 0,\ p \in \bar{C}\}. \tag{13}$$

In (12) and (13) we do not specify the metric of the space. Because p in (12) and (13) is a representation of the linear functional p^* (in some metric), the cone (13) can also be interpreted as the cone of improving directions.

We fix a number α, $0 < \alpha \leq 1$, and define the elliptic cone

$$C_\alpha(p) = \{u \in X \,|\, (p, u) \geq \alpha \|p\| \|u\|\}, \tag{14}$$

where α is the cosine of the angle between the axis λp, $\lambda \geq 0$, of the cone and its extremal rays λu_e. Any vector $u \in X$ can be represented as $u = sp + w$, $(p, p) = 1$, $(p, w) = 0$. Then $(p, u) = s$. This equation defines the hyperplane $H_s(p)$ with the normal p. The cross-section of the unit ball by this hyperplane defines the cone $C_s(p) = C_s(p^*)$:

$$C_s(p) = \{\lambda u \in X, \lambda \geq 0 \,|\, u = sp + w,\ (p, w) = 0,\ \|w\| \leq \|w_e\|\}, \tag{15}$$

according to (2), (3) and (7).

The *dispersion* of a cone can be characterized either by the (maximal) value of deviation $\|w\|$ from the axis λp, $\lambda \geq 0$, at the fixed distance s from the vertex of the cone or by the distance s for the fixed $\|w\|$. Hence, $C_s(p^*) = C_s(p) = C(w)$.

According to the Pythagorean theorem these characteristics are connected with each other: $\|u\|^2 = s^2 + \|w\|^2$. For $\|u\|^2 = \|p\|^2 = 1$: $s = \alpha$ in (14) and (15). For $w = \beta w_0$, $\|w_0\| = 1$, we have $\alpha^2 + \beta^2 = 1$, where β is the maximal value of the sine of the angle between the axis λp, $\lambda \geq 0$, of the cone and its extremal rays.

Clearly, $C_{s'}(p) \subset C_s(p)$, if $1 \geq s > s' > 0$, $\|w\|$ is fixed, and $C(w) \subset C(w')$, if $\|w'\| > \|w\|$, s is fixed.

Remark 3.3. As $sp = P_p u$ and $w = P_w u$, where $P_p u$ and $P_w u$ are the orthogonal projections onto the subspaces X_p and X_w, we obtain that $s^2 = (u, P_p u)$ and $\|w\|^2 = (u, P_w u)$. Thus, the dispersion of a cone can be characterized either by the value of $(u, P_p u)$ or $(u, P_w u)$.

Remark 3.4. For the fixed metric G the ellipsoid $E_G = \{w \,|\, (w, Gw) \leq 1 - s^2\}$, $w = u - sv_p$, $(Gv_p, w) = 0$, $\|u\|_G = \|v_p\|_G = 1$, is the cross-section of the ellipsoid (2) and the cone $C_s(p^*)$ by the hyperplane (3); in this case the axis λv_p, $\lambda \geq 0$, is orthogonal to the hyperplane (3) in the metric G.

If a direction is considered as a random vector with the mean vector $\bar{v}$ and the covariance matrix Σ, then we can define the dispersion (concentration) ellipsoid [6]: $E_\Sigma = \{u \,|\, ((u - \bar{v}), \Sigma^{-1}(u - \bar{v}))) \leq 1\}$. In the latter case the cone (15) is the cone of dispersion with the axis $\lambda \bar{v}$, $\lambda \geq 0$.

In this context the parameter s can be considered as a random variable that defines the accuracy in the constructing the supporting hyperplane.

Remark 3.5. An elliptic cone can be defined in a linear space: the cone C is elliptic, if any interior ray λv, $\lambda \geq 0$, of C determines the center of symmetry for a hyperplane section of C [1].

For a Hilbert space it means that for any interior ray λv, $\lambda \geq 0$ of an elliptic cone there is the metric $G = G_v$ and the hyperplane: $(p, u) = s$, $p = Gv$, that define the ellipsoid E_G, where $w = u - sv$, $u \in C$.

4. Uncertainty relations: I

Let $C = C_s(v) = C(w)$ and $\bar{C} = \bar{C}_s(v) = \bar{C}(\bar{w})$ be the dual cones.

Here $u = sv \pm w$, $(v, w) = 0$, $u \in C$; $p = sv \pm \bar{w}$, $(v, \bar{w}) = 0$, $p \in \bar{C}$; i.e., the cones C and $\bar{C}$ differ only by the (maximal) deviations $\|w\|$ and $\|\bar{w}\|$ from the same axis and at the same distance from the vertex of the cone.

By the definition of dual cones $(p, u) \geq 0$ for all $u \in C$ and $p \in \bar{C}$. In particular, this inequality holds for $u^- = sv - w$. Then $(p, u^-) = s^2\|v\|^2 - \|\bar{w}\|\|w\| \geq 0$.

We denote by $\sigma(u) = \|w\|$ and $\sigma(p) = \|\bar{w}\|$ the dispersions of the cones $C = C_s(v) = C(w)$ and $\bar{C} = \bar{C}_s(v) = \bar{C}(\bar{w})$. Finally,

$$\sigma(u)\sigma(p) \leq s^2\|v\|^2. \tag{16}$$

As v is the same vector in the cones C and $\bar{C}$ we can set $\|v\| = 1$. Then (16) becomes $\sigma(u)\sigma(p) \leq s^2$.

Thus, there is a maximum for the product of the dispersions, and increasing the dispersion $\sigma(u)$ of the cone C leads to decreasing the dispersion $\sigma(p)$ of the dual cone $\bar{C}$ and vice versa.

Remark 4.1. For the maximal deviations: $\|\bar{w}_e\|\|w_e\| = s^2\|v\|^2$. This equality also follows from the similarity of the rectangular triangles defined by the vectors sums $u_e^- = sv - w_e$ and $p_e = sv + \bar{w}_e$:

$$\frac{\|w_e\|}{s\|v\|} = \frac{s\|v\|}{\|\bar{w}_e\|}.$$

If the cone $C(w)$ is given, then the dual cone $\bar{C}(\bar{w})$, $\|\bar{w}\| = s^2\|v\|^2/\|w\|$, is defined. In particular, if $\|\bar{w}\| = \|w\| = s\|v\|$, then $C = \bar{C}$. If $\|w\| < s\|v\|$, then $C \subset \bar{C}$ and $\bar{C} \subset C$ if $\|w\| > s\|v\|$.

Remark 4.2. If we introduce a measure of accuracy for specifying a direction (instead of the dispersion), then (16) becomes similar to Heisenberg's uncertainty relation: the more the accuracy of specifying the direction u, the less precisely the goal of the direction (the generalized momentum) p is determined [9].

We will now consider the case which has a clear interpretation in optimization theory.

Let x_1 be the current point in an optimization process, $B_{s_1}(x_1) = \{z \mid \|z - x_1\| \leq s_1\}$ be the feasible set, and $q(z) = \|z - x_2\|$ be the (current) objective function. As above, we do not specify the metric of the space, but assume that norms that define the feasible set and the objective function are the same. In this case there exists a separating hyperplane [3]:

$$p_v^*(u) = (v, u) = s_1, \tag{17}$$

where the vector hv, $(v, v) = 1$, connects the points x_1 and x_2 and is orthogonal to this hyperplane which supports the balls $B_{s_1}(x_1)$ and $B_{s_2}(x_2)$ at $s_1 v$.

Let u be the vector from the point x_1: $z = x_1 + u$. We represent this vector as $u = s_1 v + w$, $(v, w) = 0$, $(v, v) = 1$. Similarly, let $-\bar{u}$ be the vector from the point x_2: $z = x_2 - \bar{u}$, $-\bar{u} = -s_2 v + w$, $(v, w) = 0$, $(v, v) = 1$.

In this case the vector $u + \bar{u}$ defines the path from the point x_1 to the point x_2: $x_2 = x_1 + u + \bar{u} = z + \bar{u}$, where $(v, u) = s_1$, $(v, \bar{u}) = s_2$, $s_1 + s_2 = 1$.

We consider the balls $B_{r_1}(x_1)$ and $B_{r_2}(x_2)$ with the radii $r_1^2 = s_1^2 + \|w\|^2$ and $r_2^2 = s_2^2 + \|w\|^2$. The intersections of the corresponding spheres, where $\|w\|$ is a parameter, define the hyperplane (17) [1].

If $\|w\| \leq \|w_e\|$, $\|w_e\|$ is given, then the intersections of these spheres the cone interval define $C_{s_1 s_2}(x_1; x_2) = C_{s_1}(x_1) \cap C_{s_2}^-(x_2)$, where $\lambda u \in C_{s_1}(x_1)$, $\lambda \geq 0$, and $\lambda \bar{u} \in C_{s_2}^-(x_2)$, $\lambda \leq 0$.

The cone interval $C_{s_1 s_2}(x_1; x_2)$ consists of feasible paths $u + \bar{u}$ from the point x_1 to the point x_2. All these paths change their direction at the points of the hyperplane (17), except the path $x_2 = x_1 + hv$, $h = s_1 + s_2$.

Remark 4.3. A proper cone C defines a partial ordering on X [3]:

$$x \leq y \ \Leftrightarrow y - x \in C \Leftrightarrow y \in C(x).$$

The *order cone* or the *cone interval* is defined as

$$C(x, y) = \{z \mid x \leq z \leq y\} = C(x) \cap C^-(y).$$

We will choose the value $\|w_e\|$ in such a way that the cone $C_{s_2}^-$ will be dual to the cone C_{s_1} (denoted as $\bar{C}_{s_2}^-(x_2)$. In this case for all feasible paths the direction $\bar{u}$ is improving with respect to u: $(u, \bar{u}) \geq 0$. In particular, for the extremal rays $(u_e, \bar{u}_e) = 0$, i.e., u_e is the conjugate direction to $\bar{u}_e$ (in the metric which defines the cone interval $C_{s_1 s_2}(x_1; x_2)$).

From the inequality $(u, \bar{u}) \geq 0$ and the representations of the vectors u and $\bar{u}$ it follows that $s_1 s_2 - \|w\|^2 \geq 0$ for $(v, v) = 1$. Thus

$$s_1 s_2 \geq \|w\|^2. \tag{18}$$

As was shown the value s_1 defines the maximal length of the path along v with the dispersion less than $\|w\|$. The similar meaning has the value s_2. The relation (18) shows that the product of these values is limited from below by the given dispersion $\|w\|$.

To compare the dispersions of the cones $C_{s_1}(x_1)$ and $\bar{C}^-_{s_2}(x_2)$ it is necessary that the vectors w and $\bar{w}$ must be defined at the same distance from the vertices of the cones: $s_1 = s_2 = s$. In this case $2s = h$.

From $(\bar{u}, u) = 0$ it follows that $\|\bar{w}\|\|w\| = s^2$. This equality is the known equality from the school geometry which relates the height of a rectangular triangle to the projections of its legs on the hypotenuse.

For interior paths $u + \bar{u}$ of the cone interval $C_{s_1 s_2}(x_1; x_2)$ we have $h' \leq h$ and $(u, \bar{u}) \geq 0$. Finally,

$$\|\bar{w}\|\|w\| \leq \frac{1}{4}\, h^2. \tag{19}$$

This inequality has a clear interpretation: if the distance from the initial point x_1 to the final point x_2 is limited from below by the value $h > 0$, then the more the dispersion of feasible directions, the less the dispersion of conjugate directions and vice versa.

5. Uncertainty relations: II

We will now consider a different approach describing uncertainty.

Let the ray λu, $\lambda \geq 0$, be the direction from the point x_1: $z = x_1 + \lambda u$. We consider the set of vectors u, $\|u - v\| \leq \rho$, v and ρ are given. This set defines the cone from the point x_1 with the axis v:

$$C_\rho(x_1) = \{\lambda u,\, \lambda \geq 0 \,|\, \|u - v\| \leq \rho\}. \tag{20}$$

We note that the definition of the cone is symmetric w.r.t. the vectors v and u. Thus, in the case considered the uncertainty in the realization of a direction from the point x_1 is represented by the cone (20).

We denote by $B_\rho(x_2)$ the ball of radius ρ with center at $x_2 = x_1 + v$. Then

$$C_\rho(x_1) = \{\lambda u,\, \lambda \geq 0 \,|\, u = v + y,\, y \in B_\rho(x_2)\}.$$

Let λu_e, $\lambda \geq 0$, be the tangent ray from the point x_1 to the ball $B_\rho(x_2)$, $x_{12} = x_1 + u_e$ be the tangent point (λu_e, $\lambda \geq 0$, is the extremal ray of the cone (20)). By the projection theorem $u_e = P_{u_e} v$, where $P_{u_e} v$ is the orthogonal projection onto the one-dimensional subspace defined by the vector u_e. Thus

$$u_e = v + y_e,\; u_e = P_{u_e} v,\; y_e = P_{y_e} v,\; (u_e, y_e) = 0. \tag{21}$$

We shall also use the representation of the vector u_e considered in Section 4:

$$u_e = sv + w_e,\; sv = P_v u_e,\; w_e = P_w u_e,\; (v, w_e) = 0. \tag{22}$$

We will analyze relations between parameters that specify the cone (20).

The representations (21) and (22) of the vector u_e give $\|u_e\|^2 = (u_e, v + y_e) = (v, u_e) = (v, sv + w_e) = s\|v\|^2$, i.e., $\|u_e\|^2 = s\|v\|^2$. In a similar way by using the representations of the vector y_e: $y_e = -v + u_e = -v + sv + w_e = -s^+ v + w_e$, $s^+ = 1 - s$, we obtain that $\|y_e\|^2 = -(v, y_e) = s^+\|v\|^2$.

Thus $\|u_e\|^2 + \|y_e\|^2 = s\|v\|^2 + s^+\|v\|^2 = \|v\|^2$. Finally, $(u_e, y_e) = 0$ implies that $s^+ s\|v\|^2 = \|w_e\|^2$ (cf. (18)).

We set: $\|y_e\| = \rho$, $\|u_e\| = \bar{\rho}$ and $\|w_e\| = \sigma$. From the above equations we see that the cone (20) for the given vector v is characterized by the interrelated parameters ρ, $\bar{\rho}$, s, s^+, and σ. Moreover, these parameters define the cone dual to the cone (20) if we assume that $\bar{u}_e = y_e$ and $\bar{y}_e = u_e$:

$$\bar{C}_{\bar{\rho}}(x_2) = \{\lambda\bar{u},\ \lambda \geq 0 \,|\, \bar{u} = -v + \bar{y},\ \bar{y} \in B_{\bar{\rho}}(x_1)\}. \tag{23}$$

The extremal rays of the cone (23) are the tangent rays to the ball $B_{\bar{\rho}}(x_1)$; the point x_{12} is the intersection of the tangent rays from the points x_1 and x_2 to the balls $B_\rho(x_2)$ and $B_{\bar{\rho}}(x_1)$. Thus $\bar{u}_e = -v+\bar{y}_e = -s^+v+w_e$, $(\bar{u}_e, \bar{y}_e) = 0$, $(v, w_e) = 0$, where $\bar{u}_e = -P_{\bar{u}_e}v$, $\bar{y}_e = P_{\bar{y}_e}v$.

The intersection of these cones define the cone interval $C_\rho(x_1) \cap \bar{C}_{\bar{\rho}}(x_2)$. If $w_e = P_w u_e = P_w \bar{u}_e$, then $C_\rho(x_1) \cap \bar{C}_{\bar{\rho}}(x_2) = C_s(x_1) \cap \bar{C}_{s^+}(x_2)$.

In the general case the parameters that specify the dual cone are defined independently of the (primal) cone (20). But in all cases the cone intervals defined by the dual cones will be similar:

$$kv = \bar{v},\ ky_e = \bar{u}_e,\ ku_e = \bar{y}_e,\ kw_e = \bar{w}_e,\ k > 0. \tag{24}$$

Further we will omit the lower index e for the notation of the extremal rays.

As was shown in Section 4 to establish uncertainty relations it is necessary to consider the product of uncertainties.

Proposition 5.1. *Let the extremal rays of the dual cones $C_\rho(x_1)$ and $C_{\bar{\rho}}(x_2)$ be represented as* (21), (22) *and*

$$\bar{u} = -\bar{v} + \bar{y} = -\bar{s}\bar{v} + \bar{w},\ \bar{s} = \bar{s}(\bar{v}),\ (\bar{u}, \bar{y}) = 0,\ (\bar{v}, \bar{w}) = 0. \tag{25}$$

Then
$$\|\bar{u}\|^2\|u\|^2 = \bar{s}s\|\bar{v}\|^2\|v\|^2 = \|\bar{w}\|\|w\|\|\bar{v}\|\|v\| = \|y\|^2\|\bar{y}\|^2, \tag{26}$$

where $\|y\| = \rho$, $\|\bar{y}\| = \bar{\rho}$, $s + \bar{s} = 1$.

Proof. As was shown above the representations (21) and (22) of the vector u lead to $\|u\|^2 = s\|v\|^2$. In a similar way the representation (25) of the vector $\bar{u}$ gives $\|\bar{u}\|^2 = \bar{s}\|\bar{v}\|^2$. Then the first equality in (26) is the direct consequence of these representations of the vectors $\bar{u}$ and u.

The second equality in (26) follows from the duality relation: $(\bar{u}, u) = 0$.

From the representation: $y = -v + u = -s^+v + w$, $s^+ = 1 - s$, it follows that $\|y\|^2 = s^+\|v\|^2$. Similarly we get $\|\bar{y}\|^2 = \bar{s}^+\|v\|^2$, $\bar{s}^+ = 1 - \bar{s}$. The duality relation $(\bar{y}, y) = 0$ gives $\bar{s}^+s^+\|\bar{v}\|\|v\| = \|\bar{w}\|\|w\|$. Then $\|\bar{y}\|^2\|y\|^2 = \|\bar{w}\|\|w\|\|\bar{v}\|\|v\| = \|\bar{u}\|^2\|u\|^2$. Hence $\bar{s}s = \bar{s}^+s^+$, or $\bar{s} + s = 1$.

Note that the equality $\|\bar{u}\|\|u\| = \|y\|\|\bar{y}\|$ also follows from the similarity relations (24): $\|\bar{u}\|\|u\| = \|ky\|\|\frac{1}{k}\bar{y}\| = \|y\|\|\bar{y}\|$. □

Remark 5.2. The relations (26) can also be obtained from the similarity of the rectangular triangles defined by the vector sums in (21), (22) and (25). This remark and the Remarks 3.1 and 4.1 show that the uncertainty relations also hold in a Banach space.

Remark 5.3. The points $x_2 = x_1 + v$ and $z_{12} = x_1 + sv$ for which $\bar{\rho}^2 = \|v\|\|sv\|$ holds are called conjugate w.r.t. the sphere of the radius $\bar{\rho}$. Similarly, the points $x_1 = x_2 - v$ and $z_{21} = x_2 - s^+ v$ are conjugate w.r.t. the sphere of the radius ρ if $\rho^2 = \|v\|\|s^+ v\|$. The conjugate points are used for the construction of the Green function of a sphere.

We consider particular cases of (26).

Let $v = k\bar{v}$. Then (25) becomes $\bar{u}_1 = -v + \bar{y}_1 = -\bar{s}v + \bar{w}_1$, where $\bar{u}_1 = k\bar{u}$, $\bar{y}_1 = k\bar{y}$, $\bar{w}_1 = k\bar{w}$. In this case (26) gives $\bar{\rho}_1^2\rho^2 = \bar{s}s\|v\|^4 = \|\bar{w}_1\|\|w\|\|v\|^2$ for $\|u - v\| \le \rho$ and $\|\bar{u}_1 - v\| \le \bar{\rho}_1$. In particular, $\bar{\rho}_1^2\rho^2 = \bar{s}s = \|\bar{w}_1\|\|w\|$ for $\|v\| = 1$.

Let $s\|\bar{v}\|^2 = \bar{s}\|v\|^2$, then $k = \sqrt{s/\bar{s}}$. In this case (26) gives $\bar{\rho}^2\rho^2 = \bar{s}^2\|v\|^4 = s^2\|\bar{v}\|^4$, or $\bar{\rho}\rho = s$ for $\|\bar{v}\| = 1$.

Let $u = v + y = sv + w$ and $\bar{u} = -\bar{v} + \bar{y} = -\bar{s}v + w$, $\bar{s} + s = 1$, i.e., the dual cones are defined by the same deviation vector w (for the fixed vector v: $x_2 = x_1 + v$).

By using (24) we obtain from (26) that

$$\bar{\rho}^2\rho^2 = \bar{s}s\|v\|^4 = \|w\|^2\|v\|^2. \tag{27}$$

In particular, $\bar{\rho}\rho = \|w\|$ for $\|v\| = 1$. Note that the equalities in (27) hold for the extremal vectors of the cone interval y and $\bar{y}$: $(\bar{y}, y) = 0$, $\rho = \|y\|$ and $\bar{\rho} = \|\bar{y}\|$. For any vectors of the cone interval

$$\bar{\rho}\rho \ge \sigma, \tag{28}$$

where $\sigma = \|w\|$ and $\|v\| = 1$.

In conclusion of this section we state the uncertainty relations in the following form.

Proposition 5.4. *If the uncertainty in the realization of the direction is limited by the value ρ: $\|u - v\| \le \rho$, then the uncertainty in the realization of the dual direction is limited by the value ρ/σ, $\sigma > 0$: $\|\bar{u} - v\| \le \rho/\sigma$.*

Note that, according to Section 3, the last inequality can also be written in the form $\|u^* - v^*\| \le \rho/\sigma$, which is similar to the relations considered in [5].

6. The ε-subdifferential

In the above sections a basic notion was the *feasible set*. In this section the basic notion will be the *objective function*.

We will consider the objective function in a Hilbert space X in two forms:

$$q(x) = \frac{1}{2}((x - x_c), G(x - x_c)), \tag{29}$$

where G is the positive (bounded) self-adjoint operator, x_c is the null of the function q, and

$$q_1(x) = \frac{1}{2}(x, Ax) + (b, x) + c. \tag{30}$$

Clearly $q(x) = q_1(x)$, if $G = A$, $Gx_c = b$, $\frac{1}{2}(x_c, Gx_c) = c$.

In the space $X(G)$ the objective function (29) can be written as $q(u) = \frac{1}{2}\|u\|_G^2$, $\|u\|_G^2 = (u, Gu)$, $u = x - x_c$. The function (29) defines the level sets $E_\rho(x_c) = \{u \,|\, (u, Gu) \le \rho\}$, $\rho > 0$. These sets are ellipsoids in the space X and balls in the space $X(G)$.

Let $v^* = dq(x)$ be the first differential of the function q at the point $x = x_c + v$. The representation of the linear functional $v^* \in X^*$ in the space X is called gradient. Hence, the vector Gv is the gradient of the function q at the point $x = x_c + v$ in the space X and v is the gradient of the function q at this point in the space $X(G)$.

Further we will not specify the metric of the space. Besides, we will omit the multiplier 1/2 in the definition of the objective function.

We fix the point $x_0 = x_c + v_0$ and define the ball $B_\rho(x_c) = \{u \,|\, \|u\| \le \rho\}$, $\rho = \|v_0\|$.

The first differential of the function q at the point x_0, the linear functional $v_0^* = dq(x_0) \in X^*$, defines the hyperplane: $v_0^*(u) = (v, u) = \rho$ with the normal v_n: $v_0^*(v_n) = (v_n, v_n) = 1$, that supports the ball $B_\rho(x_c)$ at the point x_0: $q(x) \ge q(x_0) + v_0^*(x - x_0)$ for all $x \in X$.

We also define an *ε-subdifferential* $v_\varepsilon^* \in X^*$ of the function q at the point x_0 [12]:

$$q(x) \ge q(x_0) - \varepsilon + v_\varepsilon^*(x - x_0), \tag{31}$$

or $\|u\|^2 \ge \|v_0\|^2 - \varepsilon + v_\varepsilon^*(u)$. The linear functional v_ε^* defines the hyperplane: $v_\varepsilon^*(u) = (v_\varepsilon, u) = \rho_\varepsilon$ with the normal v_ε: $v_\varepsilon^*(v_\varepsilon) = (v_\varepsilon, v_\varepsilon) = 1$. According to (31) this hyperplane supports the ball $B_{\rho_\varepsilon}(x_c)$ at the point $x_{0\varepsilon} = x_c + y_\varepsilon$ and contains the point $x_0 = x_c + v_0$: $(v_\varepsilon, v_0) = (v_\varepsilon, y_\varepsilon)$.

We calculate ρ_ε. Let λu_0, $\lambda \ge 0$, be the tangent ray from the point x_0 to the ball $B_{\rho_\varepsilon}(x_c)$. Then $y_\varepsilon = v_0 + u_0$, $(u_0, y_\varepsilon) = 0$, $x_{0\varepsilon} = x_c + y_\varepsilon = x_0 + u_0$. From (31) we get $\|y_\varepsilon\|^2 = \|v_0\|^2 - \|u_0\|^2$ or $\rho_\varepsilon^2 = \rho_0^2 - \varepsilon$, where $\rho_\varepsilon = \|y_\varepsilon\|$, $\rho_0 = \|v_0\|$, $\varepsilon = \|u_0\|^2$.

Note that these calculations were made in the space $X(G)$. Hence $\rho_\varepsilon^2 = (y_\varepsilon, Gy_\varepsilon)$; the vector y_ε is dual to v_ε^* (or to v_ε) in the space $X(G)$, and $(u_0, Gy_\varepsilon) = 0$, $Gy_\varepsilon = v_\varepsilon$.

We will call the ball $B_{\rho_\varepsilon}(x_c)$ the *target set*. The value ρ_ε estimates the uncertainty in the determination of the target x_c: if $\rho_\varepsilon \to 0$, then $B_{\rho_\varepsilon}(x_c) \to x_c$.

The ball $B_{\rho_\varepsilon}(x_c)$ defines the cone:

$$C_\varepsilon(x_0) = \{\lambda u, \lambda \ge 0 \,|\, u = -v_0 + y,\ y \in B_{\rho_\varepsilon}(x_c)\}. \tag{32}$$

All hyperplanes that support the ball $B_{\rho_\varepsilon}(x_c)$ and have the common point x_0 are tangent to the cone (32).

We will call this cone the *cone of feasible directions* u from the point x_0: $x = x_0+u$. In a similar way we define the ball $B_{\sqrt{\varepsilon}}(x_0) = \{u \mid \|u\| \le \sqrt{\varepsilon}\}$ and call it the *feasible set*. This set estimates the uncertainty in the determination of the point x_0.

The ball $B_{\sqrt{\varepsilon}}(x_0)$ defines the cone:

$$\bar{C}_\varepsilon(x_c) = \{\lambda\bar{u}, \lambda \ge 0 \mid \bar{u} = v_0 + \bar{y},\ \bar{y} \in B_{\sqrt{\varepsilon}}(x_0)\}. \tag{33}$$

All hyperplanes that supports the ball $B_{\sqrt{\varepsilon}}(x_0)$ and have the common point x_c are tangent to the cone (33). We call the negative cone $\bar{C}_\varepsilon^-(x_c)$ the *cone of target directions* to the point x_c: $x_c = x + \bar{u}$.

Thus the ε-subdifferential of the (smooth) function $q(x)$ defines the pair of dual cones and the corresponding uncertainty relations considered in Section 5.

References

[1] A. D. Alexandrov: A contribution to chronogeometry, Canadian J. Math. 19(4) (1967) 1119–1128.

[2] D. P. Bertsekas: Dynamic Programming and Optimal Control, vol. I and II, Athena Scientific, London (1995).

[3] S. Boyd, L. Vandenberghe: Convex Optimization, Cambridge University Press, Cambridge (2004).

[4] F. E. Browder: Multi-valued monotone nonlinear mappings and duality mappings in Banach spaces, Trans. Amer. Math. Soc. 118 (1965) 338–351.

[5] I. Ekeland, R. Temam: Convex Analysis and Variational Problems, North-Holland, Amsterdam et al. (1976).

[6] M. Friendly, G. Monette, J. Fox: Elliptical insights: understanding statistical methods through elliptical geometry, Statistical Science 28(1) (2013) 1–39.

[7] A. D. Ioffe, V. M. Tihomirov: Theory of Extremal Problems, North-Holland, Amsterdam et al. (1979).

[8] B. S. Mordukhovich, N. M. Nam: An Easy Path to Convex Analysis and Applications, Morgan & Claypool, San Rafael (2013).

[9] A. Peres: Quantum Theory: Concepts and Methods, Kluwer, Dordrecht (2002).

[10] L. Pontryagin, V. Boltayanskii, R. Gamkrelidze, E. Mishchenko: The Mathematical Theory of Optimal Processes, Wiley, New York et al. (1962).

[11] A. I. Propoi: Feasible set in model predictive control, in: Model Predictive Control: Theory, Practices and Future Challenges, C. Wade (ed.), Nova Science, Hauppauge (2015).

[12] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton (1997).