1 Introduction

In the classical calculus a continuously Fréchet differentiable mapping $F$ from a Banach space $X$ into another Banach space $Y$ is called regular at $x$ if $F'(x)$, the derivative of $F$ at $x$, maps $X$ onto $Y$. According to the Lusternik–Graves theorem this implies that the images of small balls around $x$ cover balls centered at $x$ whose radii are proportional to the radii of the corresponding balls in $X$. Moreover, it turns out that the same is true for all $x$ near $x$ and a common coefficient of proportionality can be chosen for all $x$ of a neighborhood of $x$. These are among the most fundamental facts of the classical calculus which are behind such results as the inverse and the implicit function theorem.

Furthermore, it was discovered in the 70s [10, 16, 17] that the regularity property implies distance estimates to solutions of $F(x) = y$, namely $d(x, F^{-1}(y)) \leq K\|y - F(x)\|$ with $K$ closely connected with the above mentioned coefficient proportionality (in fact reciprocal to it). This discovery triggered extensive studies of stability properties of solution to all kinds of equations.

Much of that can be extended as far as to set-valued mappings between complete metric spaces [9]. Namely, if $F : X \rightrightarrows Y$ is such a mapping, then for any $(\bar{x}, \bar{y}) \in \text{Graph } F$ we define the rate of surjection as follows:

$$\text{sur} F(\bar{y}|\bar{y}) = \liminf \left( \frac{1}{\lambda} \text{Sur} F(x|y)(\lambda), \lambda \to 0^+ \right)$$

where $\text{Sur} F(x|y)(\lambda) = \sup\{r \geq 0 : y + rB_Y \subset F(x + \lambda B_X)\}$ is the modulus of surjection at $(x, y)$ (see e.g. [8]).

It is clear that $\text{sur} F(\bar{y}|\bar{y})$ is the upper bound of $r \geq 0$ such that the inclusion

$$y + rtB_Y \cap V \subset F(x + tB_X)$$
holds for all sufficiently small $t$ if $(x, y) \in \text{Graph } F$ is sufficiently close to $(\bar{x}, \bar{y})$. It turns out furthermore that whenever $r > 0$ has this property, the reciprocal of it $K = r^{-1}$ is such that
\[
d(x, F^{-1}(y)) \leq Kd(y, F(x))
\]
for all $(x, y)$ of a neighborhood of $x$. The latter property is known as metric regularity of $F$ near $(\bar{x}, \bar{y})$ and the lower bound of such $K$ is called the rate of metric regularity of $F$ at the point and is denoted $\text{reg } F(\bar{x}|\bar{y})$. Thus we always have
\[
\text{sur } F(\bar{x}|\bar{y}) = \text{reg } F(\bar{x}|\bar{y})^{-1}.
\]

It turns out that the rates of surjection and metric regularity are closely connected with a number a quantitative characteristics playing important role in analysis and optimization, such as condition numbers of solutions of systems of equations and inequalities of various types or error bounds for all kind of inconsistent systems. Both carry valuable information about quality of convergence of associated numerical procedures. Very recently [4] it was understood that the rate of surjection is the exact boundary of the size of perturbations of the mapping which still do not destroy regularity. These and other developments underscore the necessity of knowing the exact values of the rates.

The modern calculus of subdifferentials provides some general estimates for the rates. Recall that, given a function $g$ on a Banach space $X$, a vector $x^{\ast} \in X^{\ast}$ is a Fréchet subgradient of $g$ at $x$ if
\[
g(u) - g(x) \geq \langle x^{\ast}, u - x \rangle + o(\|u - x\|).
\]
The collection $\partial Fg(x)$ of all such vectors is Fréchet subdifferential of $g$ at $x$. The sequential norm-to-weak$^{\ast}$ upper limit of $\partial F(g(u))$ as $u \to x$ is the limiting Fréchet subdifferential of $g$ at $x$. We shall denote it $\partial g(x)$.

Before we continue, we have to emphasize that Fréchet sudifferentials and limiting Fréchet subdifferentials are not always sufficient to get the results. There are some important Banach spaces (like $C$, the space of continuous function with uniform norm) in which a continuous concave function may have Fréchet subdifferential identically empty and certainly of no use. But as long as we consider a finite dimensional situation or work with convex objects in general Banach space (which are precisely the situations we shall discuss in the paper), the Fréchet subdifferetial is perfectly adequate. In particular, for a convex function the Fréchet subdifferential, the limiting Fréchet subdifferential and all other reasonable subdifferentials coincide with the subdifferential in the sense of convex analysis. For more information we can refer
to [7, 12, 19] (finite dimensional subdifferentials), [2, 14] (Fréchet subdifferentials), [9] (general subdifferential calculus).

Whichever subdifferential we take, we can associate with it two concepts: a normal cone to a set at a point belonging to the set:

\[ N(C, x) = \partial \delta_S(x), \]

where \( \delta_S \) is the indicator of \( S \) which is the function equal to zero on \( S \) and \(+\infty\) outside of \( S \) and a coderivative of a set-valued mapping from a Banach space to another at a certain points of its graph. Namely, if \( F : X \rightrightarrows Y \) and \( \bar{y} \in F(\bar{x}) \) then the coderivative of \( F \) at \((\bar{x}, \bar{y})\) is a set-valued mapping from \( Y^* \) to \( X^* \) defined as follows

\[ D^*_{F}(\bar{x}, \bar{y})(y^*) = \{ x^* \in X^* : (x^*, -y^*) \in N(\text{Graph } F, (\bar{x}, \bar{y})) \}. \]

It is in terms of coderivatives that an estimate for the rates of surjection and metric regularity can be given:

\[ \text{sur}_F(x|y) \geq \liminf_{(x,y) \to (\bar{x}, \bar{y})} \inf \{ ||x^*|| : x^* \in D^*F(x, y)(y^*), ||y^*|| = 1 \}. \tag{1} \]

Two circumstances should be taken into account in connection with this estimate. First is that it holds with any adequate subdifferential. The second, more essential to the content of the paper, is that this is still an estimate whose precision is not a priori clear.

Therefore looking for situation when exact analytic expression for the rates can be given is a worthy problem to consider. In this paper we review four such situations. The first is a known case of finite dimensional mappings considered in §2. The other three all relate to mappings with convex graph. In §3 another known result - surjection rate of a closed convex process at the origin – is considered. In §4 we consider “constraint systems” and get an exact formula for its rate of surjection at any point of the graph and in the last section we apply the latter result to constraint systems arising in linear semi-infinite programming. All references bibliographical comments will be give at corresponding section.s

2 Finite dimensional maps

We start by considering set-valued mappings between finite dimensional spaces. So let \( X \) and \( Y \) be finite dimensional Banach spaces, let \( F : X \rightrightarrows Y \) be a set-valued mapping with a closed graph and let \( \bar{\pi} \in F(\bar{x}) \).

**Theorem 1.** \( \text{sur}_F(\bar{x}|\bar{y}) = \inf \{ ||x^*|| : x^* \in D^*F(\bar{x}, \bar{y})(y^*), ||y^*|| = 1 \}. \)
The proof of the theorem is a consequence of the two lemmas below.

**Lemma 1.** Assume that \((z, w) \in \text{Graph } F\) and there are \(m > 0\) and \(\lambda > 0\) such that \(\|x^*\| \geq m\) whenever \(x^* \in D^*F(x, y)(y^*)\) for some \(x, y, y^*\) satisfying

\[
(1) \quad \|x - z\| < \lambda, \quad \|y - w\| < m\lambda, \quad \|y^*\| = 1.
\]

Then \(\text{Sur} F(z | w)(t) \geq mt\) if \(0 < t < \lambda/2\).

**Proof.** Assuming the contrary, we shall find a positive \(t < \lambda/2\) and a \(v \in \mathbb{R}^m\) with \(\xi = \|v - w\| < mt\) which does not belong to any \(F(x), \|x - z\| \leq t\)

Consider the function \(\varphi(x) = d(v, F(x)) + k\|x - z\|\) with \((\xi/t) < k < m\). We have that for any \(x\) with \(\|x - z\| = t\)

\[
\varphi(z) = d(v, F(z)) \leq \|v - w\| < kt \leq \varphi(x)
\]

Therefore there is a \(x_0\) such that \(\|x_0 - z\| < t\) and \(\varphi\) attains an unconditional local minimum on the \(t\)-ball around \(z\) at \(x_0\).

Let \(y_0 \in F(x_0)\) be the nearest to \(v\). Then \(\|y_0 - w\| < 2kt < m\lambda\) and the function \(f(x, y) = \|y - v\| + k\|x - z\|\) attains a local minimum at \((x_0, y_0)\) on the graph of \(F\). This means that there are \(\|u^*\|\) and \(\|y^*\|\) with norms not exceeding one such that

\[
0 \in (ku^*, y^*) + N(\text{Graph } F, (x_0, y_0))
\]

and \(\langle y^*, y_0 - v \rangle = \|y_0 - v\|, \langle u^*, x_0 - z \rangle = \|x_0 - z\|\). We observe that \(v \notin F(x_0)\) as \(\|x_0 - z\| < t\). Therefore \(y_0 \neq v\) so that \(\|y^*\| = 1\). Setting \(x^* = -ku^*\), we see that \(x^* \in D^*F(x_0, y_0)(y^*)\) and \(\|x^*\| = k < m\) in contradiction with the assumptions.

**Lemma 2.** Suppose that for some \((z, w) \in \text{Graph } F\)

\[
\lim_{\lambda \to 0} \frac{1}{\lambda} \text{Sur}(z, w)(\lambda) = m > 0
\]

and \(x^* \in D^*_F(z, w)(y^*)\). Then \(\|x^*\| \geq m\|y^*\|\).

**Proof.** Fix an \(\varepsilon > 0\) and let \(\lambda > 0\) be such that \(\text{Sur}(x, y)(t) > (m - \varepsilon)t\) for \(0 < t < \lambda\). We have

\[
(2) \quad \langle x^* | x - z \rangle - \langle y^* | y - w \rangle \leq r(\|x - z\| + \|y - w\|), \quad \forall (x, y) \in \text{Graph } F,
\]

where \(t^{-1}r(t) \to 0\) as \(t \to +0\). Take a positive \(t < \lambda\) and a \(y\) such that \(\|y - w\| = (m - \varepsilon)t\) and \(\langle y^*, y - w \rangle = -(m - \varepsilon)t\|y^*\|\). Then there is an \(x\) such that \(\|x - z\| \leq t\) and \(y \in F(x)\). For these \(x, y\), (2) implies:

\[
-\|x^*\|t \leq -(m - \varepsilon)t\|y^*\| + r((m - \varepsilon + 1)t).
\]
This means that \( \|x^*\| \geq (m-\varepsilon)\|y^*\| - t^{-1}r((m-\varepsilon+1)t) \) and the result follows as both \( t \) and \( \varepsilon \) can be chosen arbitrarily small.

**Proof of Theorem 1.**

Denote

\[ m = \inf\{\|x^*\| : x^* \in D^*F(\bar{x},\bar{y})(y^*), \|y^*\| = 1\} \]

Since the domain and the range spaces of \( F \) are finite dimensional, the set-valued mapping \((x, y, y^*) \to D^*F(x|y)(y^*)\) from Graph \( F \times Y^* \) is upper semi-continuous. Therefore (taking into account compactness of the unit sphere in \( Y^* \)), we conclude that

\[ \liminf_{(x,y) \to \text{Graph } F(\bar{x},\bar{y})} \inf\{\|x^*\| : x^* \in D^*F(x|y)(y^*), \|y^*\| = 1\} = m, \]

and Lemma 1 implies that \( \text{sur } F(\vec{x}|\vec{y}) \geq m \). The opposite inequality follows from the definitions (of the limiting subdifferential and the rate of surjection) and Lemma 2.

**Remark 1.** Lemma 1, as it is stated here, was proved in [7] (Theorem 8). But even a more general result can be found already in [6] (the last statement of Theorem 11.9). Lemma 2 can be distilled from the proof of Theorem 5.2 in [12]. Theorem 1 got its final form, as it is stated here, in [13].

### 3 Convex processes

Let now \( X \) and \( Y \) be arbitrary Banach spaces. Recall that a convex process from \( X \) into \( Y \) is a set-valued mapping whose graph is a closed convex cone in \( X \times Y \). Clearly \( 0 \in \mathcal{A}(0) \). The latter means that \( \mathcal{A} \) is homogeneous: \( \mathcal{A}(\lambda x) = \lambda \mathcal{A}(x) \) for \( \lambda > 0 \) and superadditive: \( \mathcal{A}(x) + \mathcal{A}(u) \subset \mathcal{A}(x+u) \). So let \( \mathcal{A} : X \to Y \) is a convex process. The quantity

\[ \|\mathcal{A}\| = \sup_{\|x\| \leq 1} \inf\{\|y\| : y \in \mathcal{A}(x)\} \]

is called the norm (or inner norm) of \( \mathcal{A} \) and

\[ C(\mathcal{A}) = \inf_{\|y^*\| = 1} \sup\{(y^*,y) : y \in \mathcal{A}(x), \|x\| \leq 1\} \]

the Banach constant of \( \mathcal{A} \).

We note the following consequence of the superadditivity property which justifies the use of the term “norm” in this context:

\[ \sigma(\mathcal{A}(x), \mathcal{A}(u)) \leq \|\mathcal{A}\|\|x - u\|, \tag{1} \]
where \( \sigma(P, Q) = \sup \{d(x, Q) : x \in P\} \) is the deviation of \( P \) from \( Q \). (Indeed,

\[
\sup_{y \in \mathcal{A}(x)} d(y, \mathcal{A}(u - x) + \mathcal{A}(x)) \leq \sup_{y \in \mathcal{A}(x)} d(y, \mathcal{A}(u - x) + \mathcal{A}(x)) = \sup_{y \in \mathcal{A}(x)} d(0, \mathcal{A}(u - x) + \mathcal{A}(x) - y) \leq d(0, \mathcal{A}(u - x)) \leq \|\mathcal{A}\|\|u - x\|.
\]

With every convex process we can associate two other convex processes: the inverse \( \mathcal{A}^{-1} : Y \to X \) defined by \( \mathcal{A}^{-1}(y) = \{x : y \in \mathcal{A}(x)\} \) and the dual \( \mathcal{A}^* : Y^* \to X^* \) defined by

\[
\mathcal{A}^*(y^*) = \{x^* \in X^* : \langle x^*, x \rangle - \langle y^*, y \rangle \leq 0, \forall (x, y) \in \text{Graph} \mathcal{A}\}.
\]

**Theorem 2.** For any convex process \( \mathcal{A} \) with \( \text{int} \mathcal{A}(X) \neq \emptyset \)

\[
\text{sur} \mathcal{A}(0|0) = \text{Sur} \mathcal{A}(0|0)(1) = \|\mathcal{A}^{-1}\|^{-1} = C(\mathcal{A}) = \inf_{\|y^*\|=1} d(0, \mathcal{A}^*(y^*))
\]

**Proof.** (a) Clearly, \( \text{Sur} \mathcal{A}(0|0)(\lambda) = \lambda \text{Sur} \mathcal{A}(0|0)(1) \) by homogeneity. So for any \( u, v \in \mathcal{A}(u) \) and \( \lambda > 0 \) we have

\[
\text{Sur} \mathcal{A}(u|v)(\lambda) = \sup \{r \geq 0 : v + rB_Y \subset \mathcal{A}(u + \lambda B_X)\} \\
\geq \sup \{r \geq 0 : v + rB_Y \subset \mathcal{A}(u) + \lambda B_X\} \\
\geq \sup \{r \geq 0 : v + rB_Y \subset v + \lambda B_X\} \\
= \lambda \text{Sur} \mathcal{A}(0|0)(1).
\]

Now the first equality in the statement is immediate from the definition of the rate of surjection.

(b) If \( r < \text{Sur} \mathcal{A}(0|0)(1) \), then for every \( y \) with \( \|y\| \leq r \) there is an \( x \) with \( \|x\| \leq 1 \) such that \( y \in \mathcal{A}(x) \). Thus for any \( y \) with \( \|y\| \leq 1 \) we have

\[
\inf \{\|x\| : y \in \mathcal{A}(x)\} \leq r^{-1}.
\]

This implies that \( \text{Sur} \mathcal{A}(0|0)(1) \leq \|\mathcal{A}^{-1}\|^{-1} \).

On the other hand, if \( r > \text{Sur} \mathcal{A}(0|0)(1) \), then there is a \( y \) with \( \|y\| \leq r \) which does not belong to \( \mathcal{A}(x) \) whenever \( \|x\| \leq 1 \). This means that \( \inf \{\|x\| : y \in \mathcal{A}(x)\} \geq r^{-1} \) and this, in view of (a) completes the proof of the second equality.

The proofs of the first two equalities do not make any use of the qualification condition \( \text{int} A(X) \neq \emptyset \) and even of convexity of \( \mathcal{A} \). The only source of these two equalities is the homogeneity property. For the other two (involving duality) the qualification condition is already needed and we each time shall separately consider two possible cases: non-degenerate, when \( A(X) = Y \), and degenerate, when \( A(X) \) is a convex cone with non-empty interior but not coinciding with the entire space.

Note that in the non-degenerate case the condition \( A(X) = Y \) implies (through the standard application of the Baire category theorem) that \( 0 \in \)
int \mathcal{A}(B_X) \) which means that \( \text{Sur}\mathcal{A}(0)(1) > 0 \). In the degenerate case, clearly, \( \text{Sur}\mathcal{A}(0)(1) = 0 \).

(c) Consider first the non-degenerate case. If \( \mathcal{A}(B_X) \) contains a ball of radius \( r \), then, of course, for every linear functional \( y^* \) we have \( \sup\{\langle y^*, y \rangle : y \in \mathcal{A}(x) \} \geq r\|y^*\| \). If, on the other hand, \( y \) does not belong to \( \mathcal{A}(B_X) \), then (as this set is obviously convex) it can be separated from the set by a \( y^* \) with the unit norm, that is for this specific \( y^* \), we have \( \langle y^*, y \rangle \geq r \) whenever the ball of radius \( r \) belongs to \( \mathcal{A}(B_X) \). Whence the third equality.

Suppose now that \( A(X) \neq Y \). As the interior of \( A(X) \) is nonempty, there is a \( y^* \) with \( \|y^*\| = 1 \) and \( \langle y^*, y \rangle \leq 0 \) for all \( y \in \mathcal{A}(X) \). On the other hand, zero of course belongs to \( \mathcal{A}(0) \), hence \( C(A) = 0 \).

(d) Again we first consider the non-degenerate case. Let \( x^* \in \mathcal{A}(y^*) \), \( \|y^*\| = 1 \), and let \( r < \text{Sur}(0)(1) \). Then there are \( y \) and \( x \) such that \( y \in \mathcal{A}(x) \), \( \|y\| = -r \) and \( \langle y^*, y \rangle = r \). Then \( -\|x^*\| \leq \langle y^*, y \rangle = -r \). This implies that \( d(0, \mathcal{A}(y^*)) \geq \text{Sur}(0)(1) \) in \( \|y^*\| = 1 \). (We apply the standard convention that \( \inf \emptyset = \infty \).

Conversely, let \( y \not\in \mathcal{A}(B_X) \). As zero belongs to the interior of the set on the right, for any \( \varepsilon > 0 \) the distance from \((1 + \varepsilon)y\) to the set is positive, say equal to \( \delta = \delta(\varepsilon) \). Therefore the set \( Q = B_X \times ((1 + 2\varepsilon)y + \varepsilon \delta B_Y \) cannot meet \( \text{Graph} \mathcal{A} \). We can therefore separate the sets by some \( 0 \neq (x^*, v^*) \in X^* \times Y^* \), say

\[
0 \geq \inf_{(u,v) \in \text{Graph} \mathcal{A}} (\langle x^*, u \rangle + \langle v^*, v \rangle) \\
\geq \sup_{(u,v) \in Q} (\langle x^*, u \rangle + \langle v^*, v \rangle) = \|x^*\| + (1 + 2\varepsilon)\langle v^*, y \rangle + \varepsilon \delta \|v^*\| > -\infty.
\]

We observe that \( v^* \neq 0 \) for otherwise we would have \( x^* = 0 \) as well. Thus we can assume that \( \|v^*\| = 1 \). Furthermore, setting \( y^* = -v^* \), we see that \( x^* \in \mathcal{A}^*(y^*) \). Finally,

\[
\|x^*\| \leq \langle (1 + 2\varepsilon)\|y\| + \varepsilon \delta(\varepsilon) \rangle.
\]

This is true for any \( y \not\in \mathcal{A}(B_X) \) and any \( \varepsilon > 0 \), and we get

\[
\inf_{\|y^*\| = 1} d(0, \mathcal{A}(y^*)) \leq \inf\{\|y\| : y \not\in \mathcal{A}(B_X) \} = \text{Sur}(0)(1).
\]

In the degenerate case the same argument as in (c) shows that there is a non-zero \( y^* \) such that \( \langle y^*, y \rangle \geq 0 \) whenever \( y \in \mathcal{A}(x) \) for some \( x \). This means that there is a \( y^* \) with \( \|y^*\| = 1 \) and such that \( 0 \in \mathcal{A}^*(y^*) \). This completes the proof of the theorem.

\textbf{Remark 2.} The norm of a convex process was introduced in [15] and the Banach constant in [6]. The first two equalities are mentioned e.g. in [4].
The equality of the rate of surjection and the last quantity in the statement is also proved in [4], although it goes back to the duality properties of convex processes (surjection vs. non-singularity) studied in [1]. We note also that the inequality \( \text{Sur}A(0\mid 0)(1) \geq C(A) \) holds for any homogeneous set-valued mapping (see e.g. [6]).

In all earlier appearances of the equalities collected in the statement of Theorem 2 the emphasis is done on the non-degenerate case. But adding the degenerate case is reasonable because in cases when the cone \( K \) itself has non-empty interior (as, say in Theorem 4 below) the results become unconditional and not needing any additional constraint qualification.

4 Constraint systems

Our next goal is to compute the rate of surjection of the set-valued mapping

\[
F(x) = \begin{cases} 
Ax - b - K, & \text{if } x \in C; \\
\emptyset & \text{if } x \notin C,
\end{cases}
\]

where \( A \) is a bounded linear operator \( X \to Y \), \( K \subset Y \) is a convex closed cone and \( C \subset X \) is a closed set.

Such mappings are typically associated with constraints in mathematical programming like \( Ax \geq b, x \in C \) with \( K \) being the cone defining an order in \( Y \).

Our purpose will be to prove the following theorem.

**Theorem 3.** Assume that \( \overline{y} \in F(\overline{x}) \) and \( \text{int}(A(C) - K) \neq \emptyset \). Then

\[
\text{sur}F(\overline{x}\mid \overline{y}) = \lim_{\lambda \to +0} \inf \{(\|x^* + Ay^*\| + \frac{1}{\lambda}(\langle y^*, b + \overline{y} - A\overline{x}\rangle + s_{C-\overline{x}}(x^*)) : x^* \in X^*, \|y^*\| = 1, y^* \in K^o}\}.
\]

Here \( s_C(x^*) = \sup \{\langle x^*, x \rangle : x \in C\} \) is the support function of \( C \).

**Proof.** To begin with, we observe that the rate of surjection does not change if we add the same quantity to all values, or if we translate the mapping by a certain vector in \( X \), that is the rate of surjection of \( F \) at \( (\overline{x}, \overline{y}) \) coincides with the rate of surjection of the mapping \( x \to F(x + \overline{x}) - \overline{y} \) at \( (0, 0) \). The new mapping has the same form as the original mapping, but \( b \) with replaced by \( b + \overline{y} - A\overline{x} \) and \( C \) by \( C - \overline{x} \). So we can assume without loss of generality that \( \overline{x} = 0 \) and \( \overline{y} = 0 \).

(a) The first step of our proof will be **homogenization** of the mapping. Namely, consider the set

\[
Q = \{(t, x, y) : t > 0, x \in tC, y \in Ax - tb - K\}.
\]
This is a convex cone. Let $\mathcal{A}$ be a convex process from $\mathbb{R} \times X$ into $Y$ whose graph is the closure of $Q$. Observe that $(t, x, y) \in \text{cl}(Q) \setminus Q$ only if $t = 0$. We have $(\eta, u^*) \in \mathcal{A}^*(y^*)$ if and only if $t\eta + \langle u^*, u \rangle - \langle y^*, y \rangle \leq 0$ whenever $(t, u, y) \in Q$. This amounts to

$$t(\eta + \langle u^*, w \rangle - \langle A^*y^*, w \rangle + \langle y^*, b \rangle + \langle y^*, z \rangle \leq 0$$

if $t > 0$, $w \in C$ and $z \in K$. It follows that $y^* \in K^\circ$ (the polar cone to $K$) and $\eta + \langle y^*, b \rangle + \langle u^* - A^*y^*, w \rangle \leq 0$ for all $w \in C$, that is that

$$\eta + \langle y^*, b \rangle + sc(u^* - A^*y^*) \leq 0. \tag{2}$$

Fix a $\lambda > 0$ and consider the following norm in $\mathcal{R} \times X$:

$$\|(t, x)\|_{\lambda} = \max\{\lambda|t|, \|x\|\}$$

and let $r_\lambda$ stand for $\text{Sur} \mathcal{A}(0|0)(1)$ corresponding to this norm in $\mathcal{R} \times X$. By Theorem 2, $r_\lambda$ is equal to the lower bound of $\|u^*\| - \lambda^{-1}|\eta|$ over the set of all $(\eta, u^*)$ satisfying (2) with some $y^* \in K^\circ, \|y^*\| = 1$. We note that the support function of $C$ is nonnegative as $0 \in C$. Furthermore, $b \in -K$, as long as we assume that $0 \in F(0)$, and therefore $\langle y^*, b \rangle \geq 0$. Thus by (2) for any fixed $u^*$ and $y^* \in K^\circ$ the lower bound with respect to $|\eta|$ (subject to the constraint) is attained for $\eta = - (\langle y^*, b \rangle + sc(u^* - A^*y^*))$ and therefore

$$r_\lambda = \inf_{\|y^*\| = 1, y^* \in K^\circ} \inf_{u^*} \left(\|u^*\| + \frac{1}{\lambda}(\langle y^*, b \rangle + sc(u^* - A^*y^*))\right). \tag{3}$$

Setting $x^* = u^* - A^*y^*$ we see that $r_\lambda$ coincides with the right quantity in the statement of the theorem.

Thus to prove the theorem we have to show that

$$\text{sur}F(0|0) = \lim_{\lambda \to +0} r_\lambda. \tag{4}$$

(b) It is easy to see that $\text{sur}F(0|0) \leq \lim_{\lambda \to 0} r_\lambda$. Indeed, denote for simplicity $\text{sur}F(0|0)$ by $\alpha$, take an $\varepsilon > 0$ and choose $\delta > 0$ so small that

$$\frac{1}{\lambda}\text{sur}F(u|v)(\lambda) > \alpha - \varepsilon$$

if $\max\{\|u\|, \|v\|, \lambda\} < \delta$ and $v \in F(u)$. Then, given a $y$ with $\|y\| \leq a - \varepsilon$ we find an $u$ with $\|u\| \leq \lambda$ such that $\lambda y \in F(u)$. Set $x = u/\lambda$, then $y \in Ax - (1/\lambda)b - K$ which means that $y \in \mathcal{A}(\lambda^{-1}, x)$. We have $\|(\lambda^{-1}, x)\|_{\lambda} = 1$, hence $y$ belongs to the image of the unit ball (in the $\lambda$-norm) under $\mathcal{A}$. As this is true for any $\lambda < \delta$, we see that $\alpha - \varepsilon \leq \lim r_\lambda$ hence $\alpha \leq \lim r_\lambda$. (We write “lim” because, as follows from (3), $r_\lambda$ do not decrease as $\lambda \to +0$.)
(c) In the degenerate case $0 \notin \text{int}(A(C) - b - K)$, zero is not in the interior of $A(\mathbb{R}_+ \times X)$ and therefore $r_\lambda = 0$ for all $\lambda$. Therefore to prove the opposite inequality in (4), we only need to consider the non-degenerate situation when $b \in \text{int}(A(C) - K)$.

For any $y$ consider the convex process $A_y$, obtained by replacing $b$ by $b + y$ in the definition of $A$. If the norm of $y$ is sufficiently small, then $y \in \text{int}(A(C) - K)$ and therefore $A_y(\mathbb{R} \times X) = Y$ and, by Theorem 2, $r_\lambda^{-1}(y) = \|A_y^{-1}\|_\lambda < \infty$ for such $y$. As above, the subscript $\lambda$ says that the norm with respect to the $\lambda$-norm in $\mathbb{R} \times X$ is considered, and $r_\lambda(y)$ is obtained from (3) by replacing $b$ by $b + y$. It is clear that $r_\lambda(y)$ depends continuously on $y$ in a neighborhood of zero.

For any $x, y$ set

$$\rho(x, y) = d(y, F(x)) = d(0, Ax - b - y - K) = d(Ax - b - y, K).$$

Clearly, this is a continuous function. We note further that $y \in A(t, x)$ implies $y + K \subset A(t, x)$ and that $Ax - b - y \in A_y(1, x)$ for all $x \in C$. This along with the deviation inequality (1) implies

(5)

$$d_\lambda((1, x), A_y^{-1}(0)) \leq \sigma_\lambda(A_y^{-1}(Ax - b - y), A_y^{-1}(K)) \leq \|A_y^{-1}\|_\lambda d_\lambda(Ax - b - y, K) = \|A_y^{-1}\|_\lambda \rho(x, y).$$

Now take $x, y$ and $\delta > 0$ small enough to guarantee in addition that

$$0 < (\|A_y^{-1}\|_\lambda + \delta)\rho(x, y) < \lambda.$$

Set for simplicity $\gamma(x, y, \delta) = (\|A_y^{-1}\|_\lambda + \delta)\rho(x, y)$. By (5) there is $(t, u) \in A_y^{-1}(0)$ such that

$$\|(1 - t, x - u)\|_\lambda < \gamma(x, y, \delta) < \lambda.$$

It follows that $|1 - t| < 1$ and, consequently, $t = 1 - (1 - t) > 0$. The latter in turn implies that $(1, u/t) \in A_y^{-1}(0)$, that is $(u/t) \in F^{-1}(y)$. It follows that

$$d(x, F^{-1}(y)) = \frac{1}{t}||tx - u|| \leq \frac{1}{t}(||(x - u)|| + (1 - t)||x||) \leq \frac{1}{1 - (1/\lambda)\gamma(x, y, \delta)}\gamma(x, y, \delta)(1 + \frac{||x||}{\lambda})$$

$$= k(x, y, \delta)(\|A_y^{-1}\|_\lambda + \delta)d(y, F(x)),$$

where $k(x, y, \delta) \to 1$ when $x \to 0$, $y \to 0$, $\delta \to 0$.

Thus, $F$ is metrically regular at 0 for 0 and for any $\lambda > 0$ the rate of metric regularity is not greater than $\|A_y^{-1}\|_\lambda = r_\lambda^{-1}$. Therefore $F$ is surjective at 0.
for 0 and \( \text{sur} F(0|0) \geq r_\lambda \). As this is true for any \( \lambda > 0 \), the result follows.

\[ \text{Remark 3.} \text{ Theorem 3 is a new result. But the study of the regularity problem for constraint systems was initiated by Robinson as early as in mid-70s [16, 17]. In the first of the paper he showed that } d(x, F^{-1}(\bar{y})) \leq (\|A\|_1 + O(\|x - \bar{x}\|)d(\bar{y}, F(x)) \text{. The second paper (in which smooth nonlinear operators are considered) effectively contains an extension of this estimate to systems depending on a parameter: if } A = A(p) \text{ continuously depends on a parameter } p \text{ and for a certain reference value } \bar{p} \text{ of the parameter } \bar{y} + b \in \text{int}(A(\bar{p}(C) - K), \text{ then } d(x, F^{-1}(p)(\bar{y}))) \leq (\|A(p)\|_1 + O(\|x - \bar{x}\| + \xi(p))d(\bar{y}, F(x)), \text{ where } \xi(p) \to 0 \text{ as } p \to \bar{p}. \]

As a corollary of this result, one can easily deduce that (in the non-parametric case) \( F \) is metrically regular with the estimate

\[ d(x, F^{-1}(y)) \leq (\|A\|_1 + O(\|x - \bar{x}\| + \xi(y))d(y, F(x)), \text{ where } \xi(y) \to 0 \text{ as } y \to \bar{y}. \]

That time however the concept of metric regularity was not yet formulated and the latter estimate is absent in Robinson’s paper.

It was in the Robinson’s proofs that the idea of homogenization and passing to the regularity problem for convex processes was first developed. And although Robinson’s proofs are essentially bound up with the \( \| \cdot \|_1 \)-norm in \( R \times X \) they can be rearranged for other norms as well as to getting the “full” metric regularity without using any additional parameters. That is what we have actually done in the part (c) of the proof where the lower estimate for the rate of surjection has been established. The upper estimate (and certainly the equality) is proved here for the first time.

5 Constraint systems of semi-infinite programming

Here we shall consider a version of a constraint system set-valued mapping which is associated with semi-infinite programming. Namely let \( a(t) \) and \( b(t) \) be continuous functions with values in \( \mathbb{R}^n \) and \( \mathbb{R}^n \) respectively, and let \( C \subset \mathbb{R}^n \) be a convex closed set. Let finally, \( Y = C[0,1] \) be the space of continuous functions on \([0,1]\) and \( K \) the cone of non-positive continuous functions. We consider the set-valued mapping

\[ F(x) = \begin{cases} a(\cdot)^T x - b(\cdot) - K, & x \in C; \\ \emptyset, & x \notin C. \end{cases} \]
We observe that \( Y^* \) is the space of Radon measures on \([0, 1]\) and the intersection of \( K^\circ \) and the unit sphere is the collection of probability measures on the segment. Note finally that the interior of \( K \) is nonempty. So we can apply Theorem 3 which leads to the following conclusion.

**Proposition 1.** Suppose that \( \overline{x} \in C \) and \( a^T(t)\overline{x} \leq b(t) \) for all \( t \) (that is \( 0 \in F(\overline{x}) \)). Then

\[
\text{sur} F(\overline{x}|0) = \lim_{\lambda \to 0} \inf \left\{ \|x^* + \int_0^1 a(t)d\mu\| \right. \\
+ \frac{1}{\lambda} \left( \int_0^1 (b(t) - a^T(t)\overline{x})d\mu + s_{C-\overline{x}}(x^*) : x^* \in \mathbb{R}^n, \int_0^1 d\mu = 1 \right). \tag{6}
\]

It is possible to calculate the limit in this case, thanks to the weak*-compactness of the set of probability measures and the finite dimensionality of the domain space. Indeed, if \( \text{sur} F(\overline{x}|0) \) is finite, then for any \( \lambda > 0 \), the infimum in the right-hand part of (6) is attained at some \( x^*_\lambda, \mu_\lambda \) and the sequence of \( x^*_\lambda \) is obviously bounded. So let \((x^*, \mu)\) be any limit point of \((x^*_\lambda, \mu_\lambda)\) as \( \lambda \to 0 \). Then (as \( s_{C-\overline{x}} \) is a nonnegative function and \( b(t) - a^T(t)\overline{x} \geq 0 \) for all \( t \)), we must have

\[
\int_0^1 (b(t) - a^T(t)\overline{x})d\mu = 0, \quad s_{C-\overline{x}}(x^*) = 0,
\]

which means that \( \mu \) is supported on the set \( T \subset [0, 1] \) on which \( b(t) = a^T\overline{x} \) and \( x^* \) belongs to the normal cone to \( C \) at \( \overline{x} \). Thus by (6)

\[
\text{sur} F(\overline{x}|0) \geq \lim_{\lambda \to 0} \inf \left\{ \|x^*_\lambda + \int_0^1 a(t)d\mu_\lambda\| \right.
\geq \left\{ \|x^* + \int_0^1 a(t)d\mu\| : \mu \geq 0, \int_0^1 d\mu = 1, \supp \mu \subset T; \ x^* \in N(C, \overline{x}) \right\}. \tag{7}
\]

The opposite inequality is immediate from (6)

By the Carathéodory theorem the infimum in (7) can be attained on measures consisting of at most \( n + 1 \) unit mass, and we get the following final result.

**Theorem 4.** Under the assumptions of Proposition 1 the rate of surjection of
$F$ at $x$ for zero is equal to the minimum of

$$\|x^* + \sum_{k=1}^{n+1} \lambda_k a(t_k)\|$$

on the set of all $2n+3$-tuples $(x^*, t_1, ..., t_{n+1}, \lambda_1, ..., \lambda_{n+1})$ such that

$x^* \in N(C, \bar{x})$, $t_k \in [0, 1]$; $b(t_k) = a(t_k)$; $\lambda_k \geq 0$ ($k = 1, ..., n+1$); $\sum_{k=1}^{n+1} \lambda_k = 1$.

**Corollary.** Suppose in addition that

$$C = \{x \in \mathbb{R}^n : e_i^T x = c_i, i = 1, ..., m\}$$

for some $e_i \in \mathbb{R}^n$ and $c_i \in \mathbb{R}$. Then

$$\text{sur} F(\bar{x} | 0) = \inf \left\{ \left\| \sum_{i=1}^m \mu_i e_i + \sum_{k=1}^{n+1} \lambda_k a(t_k) \right\| : b(t_k) = a(t_k); \lambda_k \geq 0; \sum_{k=1}^{n+1} \lambda_k = 1 \right\}.$$ 

**Remark 5.** The last corollary is one of the main result of [3]. It seems that other results of [3] can be deduced from it without much difficulty. It is not clear, however, whether the techniques used in that paper which substantially relies on the presence of equality constraints can be used to prove Theorem 4.

**Conclusion.**

We have presented in the paper precise formulas to compute the rate of surjection (metric regularity) for four different types of set-valued mappings. Two of the results (Theorems 1 and 2) are basically known, two others (Theorems 3 and 4) are rather new. In the last two cases we have considered set-valued mappings associated with linear operators. However, extensions to smooth nonlinear mapping is straightforward in each of the two cases thanks to the well known fact that adding a strictly differentiable mapping with zero derivative does not change the rates of surjection and metric regularity. All we need to get corresponding results is to replace the linear operator by the derivative of the mapping.

**References**


