

Predictive Control: for Linear and Hybrid Systems ISBN 9781139061759

- II Multiparametric Programming
- 5 Multiparametric Nonlinear Programming

5 Multiparametric Nonlinear Programming

The operations research community has addressed parameter variations in mathematical programs at two levels: *sensitivity analysis*, which characterizes the change of the solution with respect to small perturbations of the parameters, and *parametric programming*, where the characterization of the solution for a full range of parameter values is studied. In this chapter we introduce the concept of multiparametric programming and recall the main results of nonlinear multiparametric programming. The main goal is to make the reader aware of the complexities of general multiparametric nonlinear programming. Later in this book we will use multiparametric programming to characterize and compute the state feedback solution of optimal control problems. There we will only make use of Corollary 5.1 for multiparametric linear programs and Corollary 5.2 for multiparametric quadratic programs, which show that these specific programs are "well behaved."

5.1 Introduction to Multiparametric Programs

Consider the mathematical program

$$J^*(x) = \inf_{z} \quad J(z, x)$$

subj. to $g(z, x) \le 0$

where z is the optimization vector and x is a vector of parameters. We are interested in studying the behavior of the value function $J^*(x)$ and the optimizer $z^*(x)$ as we vary the parameter x. Mathematical programs where x is a scalar are referred to as *parametric programs*, while programs where x is a vector are referred to as *multiparametric programs*.

There are several reasons to look for efficient solvers of multiparametric programs. Typically, mathematical programs are affected by uncertainties due to factors that are either unknown or that will be decided later. Parametric programming systematically subdivides the space of parameters into characteristic regions, which depict the feasibility and corresponding performance as a function of the uncertain parameters, and hence provide the decision maker with a complete map of various outcomes.

Our interest in multiparametric programming arises from the field of system theory and optimal control. For example, for discrete-time dynamical systems, finite time constrained optimal control problems can be formulated as mathematical programs where the cost function and the constraints are functions of the initial state of the dynamical system. In particular, Zadeh and Whalen [293] appear to have been the first ones to express the optimal control problem for constrained discrete-time linear systems as a linear program. We can interpret the initial state as

a parameter. By using multiparametric programming we can characterize and compute the solution of the optimal control problem explicitly as a function of the initial state.

We are further motivated by the *model predictive control* (MPC) technique. MPC is very popular in the process industry for the automatic regulation of process units under operating constraints, and has attracted a considerable research effort in the last two decades. MPC requires an optimal control problem to be solved on-line in order to compute the next command action. This mathematical program depends on the current sensor measurements. The computation effort can be moved off-line by solving multiparametric programs, where the control inputs are the optimization variables and the measurements are the parameters. The solution of the parametric program problem is a *control law* describing the control inputs as function of the measurements. MPC and its multiparametric solution are discussed in Chapter 12.

In the following we will present several examples that illustrate the parametric programming problem and hint at some of the issues that need to be addressed by the solvers.

Example 5.1 Consider the parametric quadratic program

$$J^*(x) = \min_{z} \quad J(z, x) = \frac{1}{2}z^2 + 2xz + 2x^2$$

subj. to $z \le 1 + x$,

where $x \in \mathbb{R}$. Our goals are:

- 1. to find $z^*(x) = \arg\min_z J(z, x)$,
- 2. to find all x for which the problem has a solution, and
- 3. to compute the value function $J^*(x)$.

The Lagrangian is

$$L(z, x, u) = \frac{1}{2}z^{2} + 2xz + 2x^{2} + u(z - x - 1)$$

and the KKT conditions are (see Section 2.3.3 for KKT conditions for quadratic programs)

$$z + 2x + u = 0 (5.1a)$$

$$u(z - x - 1) = 0 (5.1b)$$

$$u \ge 0 \tag{5.1c}$$

$$z - x - 1 \le 0. \tag{5.1d}$$

Consider (5.1) and the two strictly complementary cases:

$$z + 2x + u = 0
A. z - x - 1 = 0
 u \ge 0$$

$$\Rightarrow \begin{cases} z^* = x + 1 \\ J^* = \frac{9}{2}x^2 + 3x + \frac{1}{2} \\ x \le -\frac{1}{3} \end{cases}$$

$$z + 2x + u = 0
B. z - x - 1 < 0
 u = 0$$

$$\Rightarrow \begin{cases} z^* = -2x \\ J^* = 0 \\ x > -\frac{1}{3} \end{cases}$$

$$(5.2)$$

This solution is depicted in Figure 5.1.

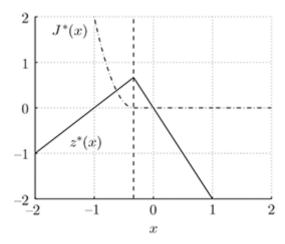


Figure 5.1 Example 5.1. Optimizer $z^*(x)$ and value function $J^*(x)$ as a function of the parameter x.

The above simple procedure, which required nothing but the solution of the KKT conditions, yielded the optimizer $z^*(x)$ and the value function $J^*(x)$ for all values of the parameter x. The set of admissible parameter values was divided into two *critical regions*, defined by $x \le -\frac{1}{3}$ and $x > -\frac{1}{3}$. In the region $x \le -\frac{1}{3}$ the inequality constraint is active and the Lagrange multiplier is greater or equal than zero, in the other region $x > -\frac{1}{3}$ the inequality constraint is not active and the Lagrange multiplier is equal to zero.

In general, when there are more than one inequality constraints, a critical region is defined by the set of inequalities that are active in the region. Throughout a critical region the conditions for optimality derived from the KKT conditions do not change. For our example, in each critical region the optimizer $z^*(x)$ is affine and the value function $J^*(x)$ is quadratic. Thus, considering all x, $z^*(x)$ is piecewise affine and $J^*(x)$ is piecewise quadratic. Both $z^*(x)$ and $J^*(x)$ are continuous, but $z^*(x)$ is not continuously differentiable.

In much of this book we will be interested in two questions: how to find the value function $J^*(x)$ and the optimizer $z^*(x)$ and what are their structural properties, e.g., continuity, differentiability and convexity. Such questions have been addressed for general nonlinear multiparametric programming by several authors in the past (see [18] and references therein), by making use of quite involved mathematical theory based on the continuity of point-to-set maps. The concept of point-to-set maps will not be much used in this book. However, it represents a key element for a rigorous mathematical description of the properties of a nonlinear multiparametric program and hence a few key theoretical results for nonlinear multiparametric programs based on the point-to-set map formalism, which will be discussed in this chapter.

5.2 General Results for Multiparametric Nonlinear Programs

Consider the nonlinear mathematical program dependent on a parameter *x* appearing in the cost function and in the constraints

$$J^*(x) = \inf_{z} \quad J(z, x)$$
subj. to $g(z, x) \le 0$, (5.3)

where $z \in \mathcal{Z} \subseteq \mathbb{R}^s$ is the optimization vector, $x \in \mathcal{X} \subseteq \mathbb{R}^n$ is the parameter vector, $J : \mathbb{R}^s \times \mathbb{R}^n \to \mathbb{R}^s$

 \mathbb{R} is the cost function and $g: \mathbb{R}^s \times \mathbb{R}^n \to \mathbb{R}^{ng}$ are the constraints. We denote by $g_i(z, x)$ the *i*-th component of the vector-valued function g(z, x).

A small perturbation of the parameter x in the mathematical program (5.3) can cause a variety of results. Depending on the properties of the functions J and g the solution $z^*(x)$ may vary smoothly or change abruptly as a function of x. Denote by 2^Z the set of subsets of \mathcal{Z} . We denote by R the point-to-set map which assigns to a parameter $x \in \mathcal{X}$ the (possibly empty) set R(x) of feasible variables $z \in \mathcal{Z}$, $R: \mathcal{X} \mapsto 2^Z$

$$R(x) = \{ z \in \mathcal{Z} : g(z, x) \le 0 \},$$
 (5.4)

by $\mathcal{K}^{\,*}$ the set of feasible parameters

$$\mathcal{K}^* = \{ x \in \mathcal{X} : \ R(x) \neq \emptyset \},\tag{5.5}$$

by $J^*(x)$ the real-valued function that expresses the dependence of the minimum value of the objective function over \mathcal{K}^* on x

$$J^*(x) = \inf_{z} \{ J(z, x) : z \in R(x) \}, \tag{5.6}$$

and by $Z^*(x)$ the point-to-set map which assigns the (possibly empty) set of optimizers $z^* \in 2^Z$ to a parameter $x \in \mathcal{X}$

$$Z^*(x) = \{ z \in R(x) : J(z, x) = J^*(x) \}.$$
(5.7)

 $J^*(x)$ will be referred to as optimal value function or simply *value function*, $Z^*(x)$ will be referred to as the *optimal set*. If $Z^*(x)$ is a singleton for all x, then $z^*(x) = Z^*(x)$ will be called *optimizer function*. We remark that R and Z^* are set-valued functions. As discussed in the notation section, with abuse of notation $J^*(x)$ and $Z^*(x)$ will denote both the functions and the value of the functions at the point x. The context will make clear which notation is being used.

The book by Bank and coauthors [18] and Chapter 2 of [107] describe conditions under which the solution of the nonlinear multiparametric program (5.3) is locally well behaved and establish properties of the optimal value function and of the optimal set. The description of such conditions requires the definition of continuity of point-to-set maps. Before introducing this concept we will show through two simple examples that continuity of the constraints $g_i(z, x)$ with respect to z and x is not enough to imply any "regularity" of the value function and the optimizer function.

Example 5.2 [18, p. 12]

Consider the following problem:

$$J^*(x) = \min_z \quad x^2 z^2 - 2x(1-x)z$$

subj. to $z \ge 0$
 $0 \le x \le 1$. (5.8)

Cost and constraints are continuous and continuously differentiable. For $0 < x \le 1$ the optimizer function is $z^* = (1-x)/x$ and the value function is $J^*(x) = -(1-x)^2$. For x = 0, the value function is $J^*(x) = 0$ while the optimal set is $Z^* = \{z \in \mathbf{R} : z \ge 0\}$. Thus, the value function is *discontinuous* at 0 and the optimal set is single-valued for all $0 < x \le 1$ and set-valued for x = 0.

Example 5.3 Consider the following problem:

$$J^*(x) = \inf_{\substack{z \\ \text{subj. to}}} z$$

subj. to $zx \ge 0$
 $-10 \le z \le 10$
 $-10 \le x \le 10$, (5.9)

where $z \in \mathbb{R}$ and $x \in \mathbb{R}$. For each fixed x the set of feasible z is a segment. The point-to-set map R(x) is plotted in Figure 5.2(a). The function $g_1 : (z, x) \mapsto zx$ is continuous. Nevertheless, the value function $J^*(x) = z^*(x)$ has a discontinuity at the origin as can be seen in Figure 5.2(b).

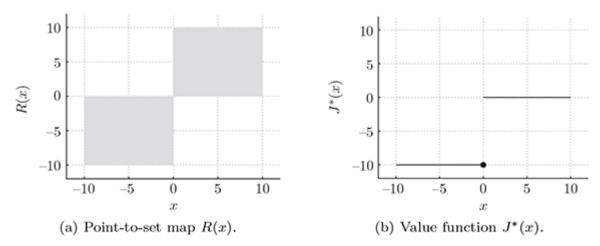


Figure 5.2 Example 5.3. Point-to-set map and value function.

Example 5.4 Consider the following problem:

$$J^{*}(x) = \inf_{\substack{z_{1}, z_{2} \\ \text{subj. to}}} -z_{1}$$

$$\sup_{\substack{g_{1}(z_{1}, z_{2}) + x \leq 0 \\ g_{2}(z_{1}, z_{2}) + x \leq 0,}} (5.10)$$

where examples of the functions $g_1(z_1, z_2)$ and $g_2(z_1, z_2)$ are plotted in Figures 5.3(a)–5.3(c). Figures 5.3(a)–5.3(c) also depict the point-to-set map $R(x) = \{[z_1, z_2] \in \mathbb{R}^2 | g_1(z_1, z_2) + x \le 0, g_2(z_1, z_2) + x \le 0\}$ for three fixed x. Starting from $x = \overline{x_1}$, as x increases, the domain of feasibility in the space z_1 , z_2 shrinks; at the beginning it is connected (Figure 5.3(a)), then it becomes

disconnected (Figure 5.3(b)) and eventually connected again (Figure 5.3(c)). No matter how smooth one chooses the functions g_1 and g_2 , the value function $J^*(x) = -z_1^*(x)$ will have a discontinuity at $x = \overline{x}_3$.

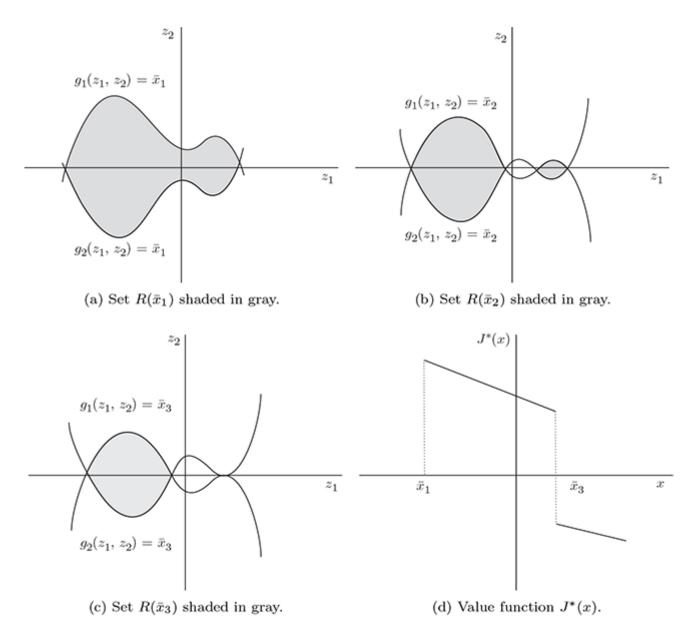


Figure 5.3 Example 5.4. Problem (5.10). (a)–(c) Projections of the point-to-set map R(x) for three values of the parameter x: $\overline{x}_1 < \overline{x}_2 < \overline{x}_3$; (d) Value function $J^*(x)$.

Examples 5.2, 5.3 and 5.4 show the case of simple and smooth constraints which lead to a discontinuous behavior of value function and in Examples 5.3 and 5.4 we also observe discontinuity of the optimizer function. The main causes are: in Example 5.2 the feasible vector space \mathcal{Z} is unbounded ($z \ge 0$),

in Examples 5.3 and 5.4 the feasible point-to-set map R(x) (defined in (5.4)) is discontinuous, as defined precisely below.

In the next sections we discuss both cases in detail.

Continuity of Point-to-Set Maps

Consider a point-to-set map $R: x \in \mathcal{X} \mapsto R(x) \in 2^{\mathbb{Z}}$. We give the following definitions of open and closed maps according to Hogan [152]:

Definition 5.1 The point-to-set map R(x) is open at a point $\overline{x} \in \mathcal{K}^*$ if for all sequences $\{x^k\} \subset \mathcal{K}^*$ with $x^k \to \overline{x}$ and for all $\overline{z} \in R(\overline{x})$ there exists an integer m and a sequence $\{z^k\} \in \mathcal{Z}$ such that $z^k \in R(x^k)$ for $k \ge m$ and $z^k \to \overline{z}$.

Definition 5.2 The point-to-set map R(x) is closed at a point $\overline{x} \in \mathcal{K}^*$ if for each pair of sequences $\{x^k\} \subset \mathcal{K}^*$, and $z^k \in R(x^k)$ with the properties

$$x^k \to \bar{x}, \ z^k \to \bar{z}.$$

it follows that $\overline{z} \in R(\overline{x})$.

We define the continuity of a point-to-set map according to Hogan [152] as follows:

Definition 5.3 The point-to-set map R(x) is continuous at a point \overline{x} in K^* if it is both open and closed at \overline{x} . R(x) is continuous in K^* if R(x) is continuous at every point x in K^* .

The definitions above are illustrated through two examples.

Example 5.5 Consider

$$R(x) = \{z \in \mathbb{R} : z \in [0,1] \text{ if } x < 1, z \in [0,0.5] \text{ if } x \ge 1.\}$$

The point-to-set map R(x) is plotted in Figure 5.4. It is easy to see that R(x) is not closed but open. In fact, if one considers a sequence $\{x^k\}$ that converges to $\overline{x} = 1$ from the left and extracts the sequence $\{z^k\}$ plotted in Figure 5.4 converging to $\overline{z} = 0.75$, then $\overline{z} \notin R(\overline{x})$ since R(1) = [0, 0.5].

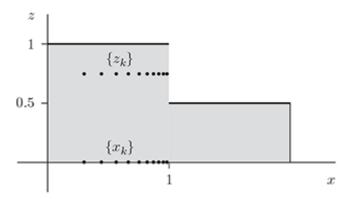


Figure 5.4 Example 5.5. Open and not closed point-to-set map R(x).

Example 5.6 Consider

$$R(x) = \{z \in \mathbb{R} : z \in [0, 1] \text{ if } x \le 1, z \in [0, 0.5] \text{ if } x > 1\}$$

The point-to-set map R(x) is plotted in Figure 5.5. It is easy to verify that R(x) is closed but not open. Choose $\overline{z} = 0.75 \in R(\overline{x})$. Then, for any sequence $\{x^k\}$ that converges to $\overline{x} = 1$ from the right, one is not able to construct a sequence $\{z^k\} \in \mathcal{Z}$ such that $z^k \in R(x^k)$ and $z^k \to \overline{z}$. In fact, such sequence z^k will always be bounded between 0 and 0.5.

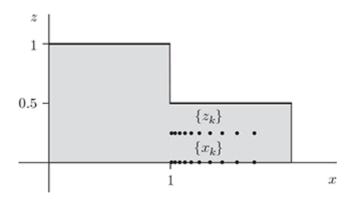


Figure 5.5 Example 5.6. Closed and not open point-to-set map R(x).

Remark 5.1 We remark that "upper semicontinuous" and "lower semicontinuous" definitions of point-to-set map are sometimes preferred to open and closed definitions [47, p. 109]. In [18, p. 25], nine different definitions for the continuity of point-to-set maps are introduced and compared. We will not give any details on this subject and refer the interested reader to [18, p. 25].

The examples above are only illustrative. In general, it is difficult to test if a set is closed or open by applying the definitions. Several authors have proposed sufficient conditions on g_i which imply the continuity of R(x). In the following we introduce a theorem which summarizes the main results of [254, 93, 152, 47, 18].

Theorem 5.1 If Z is convex, if each component $g_i(z, x)$ of g(z, x) is continuous on $Z \times \overline{x}$ and convex in z for each $\overline{x} \in X$ and if there exists a \overline{z} such that $g(\overline{z}, \overline{x}) < 0$, then R(x) is continuous at \overline{x} .

The proof is given in [152, Theorems 10 and 12]. An equivalent proof can be also derived from [18, Theorem 3.1.1 and Theorem 3.1.6]. ■

Remark 5.2 Note that convexity in z for each x is not enough to imply the continuity of R(x) everywhere in \mathcal{K}^* . In [18, Example 3.3.1 on p. 53] an example illustrating this is presented. We remark that in Example 5.3 the origin does not satisfy the last hypothesis of Theorem 5.1.

Remark 5.3 If the assumptions of Theorem 5.1 hold at each $\overline{x} \in \mathcal{X}$ then one can extract a continuous single-valued function (often called a "continuous selection") $r : \mathcal{X} \mapsto \mathbb{R}$ such that r(x)

 $\in R(x)$, $\forall x \in \mathcal{X}$, provided that \mathcal{Z} is finite-dimensional. Note that convexity of R(x) is a critical assumption [18, Corollary 2.3.1]. The following example shows a point-to-set map R(x) not convex for a fixed x which is continuous but has no continuous selection [18, p. 29]. Let Λ be the unit disk in \mathbb{R}^2 , define R(x) as

$$x \in \Lambda \mapsto R(x) = \left\{ z \in \Lambda : \|z - x\|_2 \ge \frac{1}{2} \right\}. \tag{5.11}$$

It can be shown that the point-to-set map R(x) in (5.11) is continuous according to Definition 5.3. In fact, for a fixed $\overline{x} \in \Lambda$ the set $R(\overline{x})$ is the set of points in the unit disk outside the disk centered in \overline{x} and of radius 0.5 (next called the *half disk*); small perturbations of \overline{x} yield small translations of the half disk inside the unit disk for all $\overline{x} \in \Lambda$. However R(x) has no continuous selection. Assume that there exists a continuous selection $r: x \in \Lambda \mapsto r(x) \in \Lambda$. Then, there exists a point x^* such that $x^* = r(x^*)$. Since $r(x) \in R(x)$, $\forall x \in \Lambda$, there exists a point x^* such that $x^* \in R(x^*)$. This is not possible since for all $x^* \in \Lambda$, $x^* \notin R(x^*)$ (recall that $R(x^*)$ is set of points in the unit disk outside the disk centered in x^* and of radius 0.5).

Remark 5.4 Let Λ be the unit disk in \mathbb{R} , define R(x) as

$$x \in \Lambda \mapsto R(x) = \left\{ z \in \Lambda : |z - x| \ge \frac{1}{2} \right\}.$$
 (5.12)

R(x) is closed and not open and it has no continuous selection.

Remark 5.5 Based on [18, Theorem 3.2.1-(I) and Theorem 3.3.3], the hypotheses of Theorem 5.1 can be relaxed for affine $g_i(z, x)$. In fact, affine functions are weakly analytic functions according to [18, p. 47]. Therefore, we can state that if \mathcal{Z} is convex, if each component $g_i(z, x)$ of g(z, x) is an affine function, then R(x) is continuous at \overline{x} for all $\overline{x} \in \mathcal{K}^*$.

Properties of the Value Function and Optimal Set

Consider the following definition

Definition 5.4 A point-to-set map R(x) is said to be uniformly compact near \overline{x} if there exist a neighborhood N of \overline{x} such that the closure of the set $\bigcup_{x \in N} R(x)$ is compact.

Now we are ready to state the two main theorems on the continuity of the value function and of the optimizer function.

Theorem 5.2 [152, Theorem 7] Consider problem (5.3)–(5.4). If R(x) is a continuous point-to-set map at \overline{x} and uniformly compact near \overline{x} and if J is continuous on $\overline{x} \times R(\overline{x})$, then J^* is continuous at \overline{x} .

Theorem 5.3 [152, Corollary 8.1] Consider problem (5.3)–(5.4). If R(x) is a continuous point-to-set map at \overline{x} , J is continuous on $\overline{x} \times R(\overline{x})$, Z^* is nonempty and uniformly compact near \overline{x} , and Z^* (\overline{x}) is single valued, then Z^* is continuous at \overline{x} .

Remark 5.6 Equivalent results of Theorems 5.2 and 5.3 can be found in [47, p. 116] and [18, Chapter 4.2].

Example 5.7 Example 5.2 revisited

Consider Example 5.2. The feasible map R(x) is unbounded and therefore it does not satisfy the assumptions of Theorem 5.2 (since it is not uniformly compact). Modify Example 5.2 as follows:

$$J^*(x) = \min_z \quad x^2 z^2 - 2x(1-x)z$$

subj. to $0 \le z \le M$
 $0 \le x \le 1$ (5.13)

with $M \ge 0$. The solution can be computed immediately. For $1/(1 + M) < x \le 1$ the optimizer function is $Z^* = (1 - x)/x$ and the value function is $J^*(x) = -(1 - x)^2$. For $0 < x \le 1/(1 + M)$, the value function is $J^*(x) = x^2M^2 - 2Mx(1 - x)$ and the optimizer function is $Z^* = M$. For X = 0, the value function is $J^*(x) = 0$ while the optimal set is $Z^* = \{z \in \mathbf{R} : 0 \le z \le M\}$.

No matter how large we choose M, the value function and the optimal set are continuous for all $x \in [0, 1]$.

Example 5.8 Example 5.3 revisited

Consider Example 5.3. The feasible map R(x) is not continuous at x = 0 and therefore it does not satisfy the assumptions of Theorem 5.2. Modify Example 5.3 as follows:

$$J^*(x) = \inf_{z} z$$
subj. to $zx \ge -\varepsilon$

$$-10 \le z \le 10$$

$$-10 \le x \le 10,$$

$$(5.14)$$

where $\varepsilon > 0$. The value function and the optimal set are depicted in Figure 5.2 for $\varepsilon = 1$. No matter how small we choose ε , the value function and the optimal set are continuous for all $x \in [-10, 10]$

The following corollaries consider special classes of parametric problems.

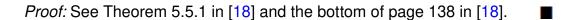
Corollary 5.1 (mp-LP) Consider the special case of the multiparametric program (5.3). where the objective and the constraints are linear

$$J^*(x) = \min_{z} \quad c'z$$

$$subj. \ to \quad Gz \le w + Sx,$$

$$(5.15)$$

and assume that there exists an \overline{x} and $z^*(\overline{x})$ with a bounded cost $J^*(\overline{x})$. Then, \mathcal{K}^* is a nonempty polyhedron, $J^*(x)$ is a continuous and convex function on \mathcal{K}^* and the optimal set $Z^*(x)$ is a continuous point-to-set map on \mathcal{K}^* .



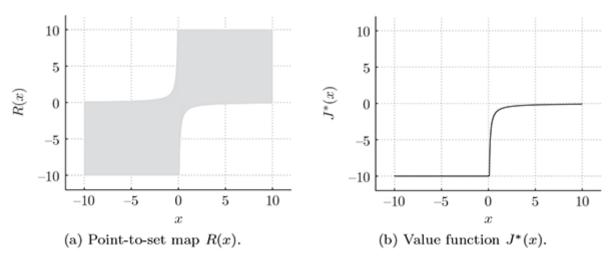


Figure 5.6 Example 5.8. Point-to-set map and value function.

Corollary 5.2 (mp-QP) Consider the special case of the multiparametric program (5.3). where the objective is quadratic and the constraints are linear

$$J^*(x) = \min_{z} \frac{1}{2}z'Hz + z'F$$

$$subj. \ to \quad Gz \le w + Sx,$$

$$(5.16)$$

and assume that H > 0 and that there exists $(\overline{z}, \overline{x})$ such that $G\overline{z} \le w + S\overline{x}$. Then, K^* is a nonempty polyhedron, $J^*(x)$ is a continuous and convex function on K^* and the optimizer function $z^*(x)$ is continuous in K^* .

Proof: See Theorem 5.5.1 in [18] and the bottom of page 138 in [18].

Remark 5.7 We remark that Corollary 5.1 requires the existence of optimizer $z^*(\overline{x})$ with a bounded cost. This is implicitly guaranteed in the mp-QP case since in Corollary 5.2 the matrix H is assumed to be strictly positive definite. Moreover, the existence of an optimizer $z^*(\overline{x})$ with a bounded cost guarantees that $J^*(x)$ is bounded for all x in \mathcal{K}^* . This has been proven in [115, p. 178, Theorem 1] for the mp-LP case and it is immediate to prove for the mp-QP case.

Remark 5.8 Both Corollary 5.1 (mp-LP) and Corollary 5.2 (mp-QP) could be formulated stronger: J^* and Z^* are even Lipschitz-continuous. J^* is also piecewise affine (mp-LP) or piecewise quadratic (mp-QP), and for the mp-QP $z^*(x)$ is piecewise affine. For the linear case, Lipschitz continuity is known from Walkup-Wets [283] as a consequence of Hoffman's theorem. For the quadratic case, Lipschitz continuity follows from Robinson [253], as e.g., shown by Klatte and Thiere [178]. The "piecewise" properties are consequences of local stability analysis of parametric optimization, e.g., [107, 18, 200] and are the main focus of the next chapter.