

# P-FACTOR-APPROACH TO DEGENERATE OPTIMIZATION PROBLEMS

O. A. Brezhneva,<sup>1</sup> and A. A. Tret'yakov<sup>2</sup>

<sup>1</sup>*Department of Mathematics and Statistics, Miami University, Oxford, OH 45056, USA, brezhnoa@muohio.edu,* <sup>2</sup>*Center of the Russian Academy of Sciences, Vavilova 40, Moscow, GSP-1, Russia and University of Podlasie in Siedlce, 3 Maja, Siedlce, Poland, tret@ap.siedlce.pl*

**Abstract** The paper describes and analyzes an application of the  $p$ -regularity theory to *nonregular, (irregular, degenerate)* nonlinear optimization problems. The  $p$ -regularity theory, also known as the *factor-analysis of nonlinear mappings*, has been developing successfully for the last twenty years. The  $p$ -factor-approach is based on the construction of a  $p$ -factor-operator, which allows us to describe and analyze nonlinear problems in the degenerate case.

First, we illustrate how to use the  $p$ -factor-approach to solve degenerate optimization problems with equality constraints, in which the Lagrange multiplier associated with the objective function might be equal to zero. We then present necessary and sufficient optimality conditions for a degenerate optimization problem with inequality constraints. The  $p$ -factor-approach is also used for solving mathematical programs with equilibrium constraints (MPECs). We show that the constraints are 2-regular at the solution of the MPEC. This property allows us to localize the minimizer independently of the objective function. The same idea is applied to some other nonregular nonlinear programming problems and allows us to reduce these problems to a regular system of equations without an objective function.

**keywords:** Lagrange optimality conditions, degeneracy,  $p$ -regularity

## 1. Introduction

The main goal of this paper is to describe and analyze an application of the  $p$ -regularity theory to *nonregular, (irregular, degenerate)* nonlinear optimization problems. In the first part of the paper, we recall some definitions of the  $p$ -regularity theory [2, 3]. In the second part, we illustrate how to use the  $p$ -factor-approach to solve degenerate optimization problems with equality constraints, in which the Lagrange multiplier associated with the objective function might be equal to zero. In the third part, we present necessary and sufficient optimality conditions for a degenerate optimization problem with inequality constraints. In

Please use the following format when citing this chapter:

Author(s) [insert Last name, First-name initial(s)], 2006, in IFIP International Federation for Information Processing, Volume 199, System Modeling and Optimization, eds. Ceragioli F., Dontchev A., Furuta H., Marti K., Pandolfi L., (Boston: Springer), pp. [insert page numbers].

the last part of the paper we consider mathematical programs with equilibrium constraints (MPECs).

**Notation.** Let  $\mathcal{L}(X, Y)$  be the space of all continuous linear operators from  $X$  to  $Y$  and for a given linear operator  $\Lambda : X \rightarrow Y$ , we denote its image by  $\text{Im}\Lambda = \{y \in Y \mid y = \Lambda x \text{ for some } x \in X\}$ . Also,  $\Lambda^* : Y^* \rightarrow X^*$  denotes the adjoint of  $\Lambda$ , where  $X^*$  and  $Y^*$  denote the dual spaces of  $X$  and  $Y$ , respectively. Let  $p$  be a natural number and let  $B : X \times X \times \dots \times X$  (with  $p$  copies of  $X$ )  $\rightarrow Y$  be a continuous symmetric  $p$ -multilinear mapping. The  $p$ -form associated to  $B$  is the map  $B[\cdot]^p : X \rightarrow Y$  defined by  $B[x]^p = B(x, x, \dots, x)$ , for  $x \in X$ .

If  $F : X \rightarrow Y$  is a differentiable mapping, its derivative at a point  $x \in X$  will be denoted by  $F'(x) : X \rightarrow Y$ . If  $F : X \rightarrow Y$  is of class  $C^p$ , we let  $F^{(p)}(x)$  be the  $p$ th derivative of  $F$  at the point  $x$  (a symmetric multilinear map of  $p$  copies of  $X$  to  $Y$ ) and the associated  $p$ -form, also called the  $p$ th-order mapping, is  $F^{(p)}(x)[h]^p = F^{(p)}(x)(h, h, \dots, h)$ .

## 2. The $p$ -factor-operator and the $p$ -regular mappings

In this section, we recall some definitions of the  $p$ -regularity theory [2, 3].

Consider a sufficiently smooth mapping  $F$  from a Banach space  $X$  to a Banach space  $Y$ . The mapping  $F$  is called *regular* at some point  $\bar{x} \in X$  if

$$\text{Im } F'(\bar{x}) = Y.$$

We are interested in the case when the mapping  $F$  is *nonregular (irregular, degenerate)* at  $\bar{x}$ , i.e., when

$$\text{Im } F'(\bar{x}) \neq Y. \quad (1)$$

We recall the definition of the  $p$ -regular mapping and of the  $p$ -factor-operator. We construct a  $p$ -factor-operator under an assumption that the space  $Y$  is decomposed into a direct sum

$$Y = Y_1 \oplus \dots \oplus Y_p, \quad (2)$$

where  $Y_1 = \text{cl Im } F'(\bar{x})$ ,  $Y_i = \text{cl Sp } (\text{Im } P_{Z_i} F^{(i)}(\bar{x})[\cdot]^i)$ ,  $i = 2, \dots, p-1$ ,  $Y_p = Z_p$ ,  $Z_i$  is a closed complementary subspace for  $(Y_1 \oplus \dots \oplus Y_{i-1})$  with respect to  $Y$ ,  $i = 2, \dots, p$ , and  $P_{Z_i} : Y \rightarrow Z_i$  is the projection operator onto  $Z_i$  along  $(Y_1 \oplus \dots \oplus Y_{i-1})$  with respect to  $Y$ ,  $i = 2, \dots, p$ .

Define the mappings [2]

$$f_i(x) : X \rightarrow Y_i, \quad f_i(x) = P_{Y_i} F(x), \quad i = 1, \dots, p, \quad (3)$$

where  $P_{Y_i} : Y \rightarrow Y_i$  is the projection operator onto  $Y_i$  along  $(Y_1 \oplus \dots \oplus Y_{i-1} \oplus Y_{i+1} \oplus \dots \oplus Y_p)$  with respect to  $Y$ ,  $i = 1, \dots, p$ .

The  $p$ -factor-operator plays the central role in the  $p$ -regularity theory. The number  $p$  is chosen as the minimum number for which (2) holds. We give the following definition of the  $p$ -factor-operator.

DEFINITION 1 *The linear operator  $\Psi_p(h) \in \mathcal{L}(X, Y_1 \oplus \dots \oplus Y_p)$ , defined by*

$$\Psi_p(h) = f'_1(\bar{x}) + f''_2(\bar{x})[h] + \dots + f_p^{(p)}(\bar{x})[h]^{p-1}, \quad h \in X,$$

*is called a  $p$ -factor-operator of the mapping  $F(x)$  at the point  $\bar{x}$ .*

Now we are ready to introduce another very important definition in the  $p$ -regularity theory.

DEFINITION 2 *We say that the mapping  $F$  is  $p$ -regular at  $\bar{x}$  along an element  $h$  if  $\text{Im} \Psi_p(h) = Y$ .*

The following definition is a specific form of Definition 1 for the case of  $p = 2$  and  $F : \mathcal{R}^n \rightarrow \mathcal{R}^n$ .

DEFINITION 3 *A linear operator  $\Psi_2(h) : \mathcal{R}^n \rightarrow \mathcal{R}^n$ ,*

$$\Psi_2(h) = F'(\bar{x}) + P^\perp F''(\bar{x})h, \quad h \in \mathcal{R}^n, \quad \|h\| = 1,$$

*is said to be the 2-factor-operator, where  $P^\perp$  is a matrix of the orthoprojector onto  $(\text{Im} F'(\bar{x}))^\perp$ , which is an orthogonal complementary subspace to the image of the first derivative of  $F$  evaluated at  $\bar{x}$ .*

The next definition is a specific form of Definition 2.

DEFINITION 4 *The mapping  $F$  is called 2-regular at  $\bar{x}$  along an element  $h$  if*

$$\text{Im} \Psi_2(h) = \mathcal{R}^n.$$

### 3. Degenerate optimization problems with equality constraints

In this section, we consider the nonlinear optimization problem with equality constraints:

$$\underset{x \in X}{\text{minimize}} \quad f(x) \quad \text{subject to} \quad F(x) = 0, \quad (4)$$

where  $f : X \rightarrow R$ . We will denote a local solution of (4) by  $\bar{x}$ . We assume that  $F, f$  are  $C^{p+1}$  in some neighborhood of  $\bar{x}$  and that the mapping  $F : X \rightarrow Y$  is nonregular at  $\bar{x}$ , i.e., the condition (1) holds.

The Lagrangian for problem (4) is defined as

$$L(x, \lambda_0, \lambda) = \lambda_0 f(x) + \langle \lambda, F(x) \rangle, \quad (5)$$

where  $(\lambda_0, \lambda) \in (\mathcal{R} \times Y^*) \setminus \{0\}$  is a generalized Lagrange multiplier.

The classical first-order Euler-Lagrange necessary optimality conditions for problem (4) state that there exists a generalized Lagrange multiplier  $(\lambda_0, \bar{\lambda})$  such that

$$\begin{aligned} \lambda_0 f'(\bar{x}) + (F'(\bar{x}))^* \bar{\lambda} &= 0, \\ F(\bar{x}) &= 0, \\ \lambda_0^2 + \|\bar{\lambda}\|^2 &= 1. \end{aligned}$$

In other words, the point  $(\bar{x}, \lambda_0, \bar{\lambda})$  is a solution of the following system of equations:

$$\mathcal{L}(x, \lambda_0, \lambda) = \begin{pmatrix} \lambda_0 f'(x) + (F'(x))^* \lambda \\ F(x) \\ \lambda_0^2 + \|\lambda\|^2 - 1 \end{pmatrix} = 0. \tag{6}$$

We are interested in the case, when the Lagrange multiplier  $\lambda_0$  might be equal to zero. In this case, the mapping  $\mathcal{L}$  and the system (6) are degenerate at  $(\bar{x}, 0, \bar{\lambda})$ . However, if the mapping  $\mathcal{L}$  is  $p$ -regular at  $(\bar{x}, 0, \bar{\lambda})$  with respect to some vector  $h$ , then solving system (6) can be reduced to solving a regular system of equations. This result is stated in the following Theorem 5.

Before we give the theorem, introduce the notation  $z = (x, \lambda_0, \lambda)$  and the functions  $l_i(z)$  associated with the mapping  $\mathcal{L}(z)$  introduced in (6). We define functions  $l_i(z)$  by (3) with  $f_i(x) = l_i(z)$  and  $F(x) = \mathcal{L}(z)$ . We also define  $\bar{\Psi}_p(h)$  to be a  $p$ -factor-operator of the mapping  $\mathcal{L}(z)$  at the point  $\bar{z} = (\bar{x}, 0, \bar{\lambda})$ .

**THEOREM 5** *Let  $\bar{x}$  be a solution of (4) and let  $(0, \bar{\lambda})$  be a generalized Lagrange multiplier such that  $(\bar{x}, 0, \bar{\lambda})$  is a solution of (6). Assume that the mapping  $\mathcal{L}(z)$ , defined in (6) is  $p$ -regular at the point  $\bar{z} = (\bar{x}, 0, \bar{\lambda})$  with respect to some vector  $h$ . Assume also that  $\text{Ker} \bar{\Psi}_p(h) = \{0\}$ . Then the following system has a locally unique regular solution  $\bar{z}$ :*

$$\bar{\mathcal{L}}(z) = l_1(z) + l'_2(z)[h] + \dots + l_p^{(p-1)}[h]^{p-1} = 0. \tag{7}$$

**Example.** Consider the problem

$$\underset{(x,y)}{\text{minimize } y} \quad \text{subject to} \quad F(x, y) = x^2 - y^3 = 0, \tag{8}$$

This problem has a solution  $(\bar{x}, \bar{y})^T = (0, 0)^T$  and  $F'(0, 0) = (0, 0)$ . The Lagrangian for (8) is defined as

$$L(x, y, \lambda_0, \lambda) = \lambda_0 y + (x^2 - y^3)\lambda.$$

Then the system (6) for problem (8) has the form

$$\mathcal{L}(x, y, \lambda_0, \lambda) = \begin{pmatrix} 2x\lambda \\ \lambda_0 - 3y^2\lambda \\ x^2 - y^3 \\ \lambda_0^2 + \lambda^2 - 1 \end{pmatrix} = 0.$$

This system has a unique degenerate solution  $(\bar{x}, \bar{y}, \lambda_0, \bar{\lambda}) = (0, 0, 0, 1)$ .

Since  $\mathcal{L}(x, y, \lambda_0, \lambda)$  is 3-regular at  $(0, 0, 0, 1)^T$  with respect to the vector  $h = (0, 1, 0, 0)^T$ , we get by Theorem 5 that the system (7) is defined as

$$\bar{\mathcal{L}}(x, y, \lambda_0, \lambda) = \begin{pmatrix} 2x\lambda \\ \lambda_0 - 3y^2\lambda \\ -3y^2 - 6y \\ \lambda_0^2 + \lambda^2 - 1 \end{pmatrix} = 0.$$

This system has a locally unique regular solution  $(0, 0, 0, 1)^T$ .

#### 4. Optimality conditions for degenerate optimization problems

In this section, we consider the nonlinear optimization problem with inequality constraints

$$\underset{x \in X}{\text{minimize}} \ f(x) \quad \text{subject to} \quad g(x) = (g_1(x), \dots, g_m(x))^T \geq 0, \quad (9)$$

where  $f, g_i : X \rightarrow \mathcal{R}^1$  and  $X$  is a Banach space. We will denote a local solution of (9) by  $\bar{x}$ .

For some  $p \geq 2$ , we say that we have the *completely degenerate case* if

$$g_i^{(r)}(\bar{x}) = 0, \quad r = 1, \dots, p-1, \quad i \in I(\bar{x}) = \{i \mid g_i(\bar{x}) = 0\}. \quad (10)$$

Introduce the sets

$$H_p(\bar{x}) = \{h \in X \mid g_i^{(p)}(\bar{x}) [h]^p \geq 0, i \in I(\bar{x})\}, \quad (11)$$

$$I_p(\bar{x}, h) = \{i \in I(\bar{x}) \mid g_i^{(p)}(\bar{x}) [h]^p = 0, h \in H_p(\bar{x})\},$$

and

$$H_\alpha = \{h \in H_p(\bar{x}) \mid |g_i^{(p)}(\bar{x}) [h]^p| \leq \alpha, i \in I(\bar{x}), \|h\| = 1\}.$$

Introduce the definition.

DEFINITION 6 A mapping  $g(x)$  is called strongly  $p$ -regular at  $\bar{x}$  if there exists  $\alpha > 0$  such that  $\sup_{h \in H_\alpha} \|\{\Psi_p(h)\}^{-1}\| < \infty$ .

Let fix some element  $h \in X$  and introduce the  $p$ -factor-Lagrange function [3]:

$$\mathcal{L}_p(x, h, \lambda(h)) = f(x) - \sum_{i=1}^m \lambda_i(h) g_i^{(p-1)}(x)[h]^{p-1}. \quad (12)$$

THEOREM 7 Let  $\bar{x}$  be a local solution to problem (9). Assume that there exists a vector  $h$ ,  $\|h\| = 1$ ,  $h \in H_p(\bar{x})$ , such that the vectors  $\{g_i^{(p)}(\bar{x})[h]^{p-1}$ ,  $i \in I_p(\bar{x}, h)\}$  are linearly independent. Then there exist Lagrange multipliers  $\bar{\lambda}_i(h)$  such that

$$\mathcal{L}_{p_x}(\bar{x}, h, \bar{\lambda}(h)) = f'(\bar{x}) - \sum_{i=1}^m \bar{\lambda}_i(h) (g_i^{(p)}(\bar{x})[h]^{p-1}) = 0,$$

$$g_i(\bar{x}) \geq 0, \quad \bar{\lambda}_i(h) \geq 0, \quad \bar{\lambda}_i(h) g_i(\bar{x}) = 0, \quad i = 1, \dots, m.$$

Furthermore, suppose that the mapping  $g(x)$  is strongly  $p$ -regular at  $\bar{x}$ . If there exist  $w > 0$  and  $\bar{\lambda}(h)$  such that

$$\mathcal{L}_{p_x}(\bar{x}, h, \bar{\lambda}(h)) = 0 \quad \text{and} \quad \mathcal{L}_{p_{xx}}\left(\bar{x}, h, \frac{2\bar{\lambda}(h)}{p(p+1)}\right) [h]^2 \geq w \|h\|^2 \quad (13)$$

for all  $h \in H_p(\bar{x})$ , then  $\bar{x}$  is an isolated solution to problem (9).

## 5. Mathematical programs with equilibrium constraints (MPECs)

The MPEC considered in this section is a mathematical program with nonlinear complementary problem (NCP) constraints:

$$\underset{x}{\text{minimize}} \quad f(x) \quad \text{subject to} \quad g(x) \geq 0, \quad x \geq 0, \quad \langle g(x), x \rangle = 0, \quad (14)$$

where  $f : \mathcal{R}^n \rightarrow \mathcal{R}$ ,  $g : \mathcal{R}^n \rightarrow \mathcal{R}^n$  are twice continuously differentiable functions. We are interested in the case when the strict complementarity condition does *not* hold at the solution  $\bar{x}$ , i.e., when there exists at least one index  $j$  such that  $g_j(\bar{x}) = 0$  and  $\bar{x}_j = 0$ .

We show that under a certain condition, the problem (14) can be reduced to solving a system of nonlinear equations, which is independent of the objective function.

By introducing the slack variables  $s_j^2$  and  $y_j^2$ , we reduce (14) to the problem with only equality constraints in the form:

$$F(x, s, y) = \begin{pmatrix} g_1(x) - s_1^2 \\ \dots \\ g_n(x) - s_n^2 \\ x_1 - y_1^2 \\ \dots \\ x_n - y_n^2 \\ s_1 y_1 \\ \dots \\ s_n y_n \end{pmatrix} = 0. \quad (15)$$

System (15) is a system of the  $3n$  equations in the  $3n$  unknowns  $x$ ,  $s$ , and  $y$ . The corresponding Jacobian is given by

$$F'(x, s, y) = \begin{pmatrix} g'_1(x) & -2s_1(e^1)^T & 0 \\ \dots & \dots & \dots \\ g'_n(x) & -2s_n(e^n)^T & 0 \\ (e^1)^T & 0 & -2y_1(e^1)^T \\ \dots & \dots & \dots \\ (e^n)^T & 0 & -2y_n(e^n)^T \\ 0 & y_1(e^1)^T & s_1(e^1)^T \\ \dots & \dots & \dots \\ 0 & y_n(e^n)^T & s_n(e^n)^T \end{pmatrix}, \quad (16)$$

where  $e^1, \dots, e^n$  denotes the standard basis in  $\mathcal{R}^n$ . If there exists an index  $j$  such that  $y_j = 0$  and  $s_j = 0$  (the strict complementarity condition does not hold), then the Jacobian matrix (16) is singular.

Assume that we can identify the set  $I_0$  of the weakly active constraint indices, that is,  $I_0 = \{i = 1, \dots, n \mid g_i(\bar{x}) = 0 \text{ and } \bar{x}_i = 0\}$ . There are different techniques to identify the set  $I_0$ , for example, ones described in [1] or in [4].

Define the vector  $h = (h_1, \dots, h_n)$  as

$$h_i = \begin{cases} 1, & i \in I_0 \\ 0, & i \notin I_0 \end{cases}$$

and a vector  $\bar{h} \in \mathcal{R}^{3n}$  as  $\bar{h}^T = (0_n^T, h^T, h^T)$ .

From the explicit form of the Jacobian (16), the orthoprojector  $P^\perp$  onto  $(\text{Im } F'(x, s, y))^\perp$  in  $\mathcal{R}^{3n}$  is a diagonal matrix  $P^\perp = \text{diag}(p_j)_{j=1}^{3n}$  that is constant in some neighborhood of  $(\bar{x}, \bar{s}, \bar{y})$  and that is given by

$$p_i = \begin{cases} 1, & i = 2n + j, \text{ and } j \in I_0 \\ 0, & \text{otherwise.} \end{cases}$$

Construct the mapping

$$\Psi(x, s, y) = F(x, s, y) + P^\perp F'(x, s, y)\bar{h}.$$

Without loss of generality, we assume that  $I_0 = \{1, \dots, r\}$ . Then

$$\Psi(x, s, y) = \begin{pmatrix} g_1(x) - s_1^2 \\ \dots \\ g_n(x) - s_n^2 \\ x_1 - y_1^2 \\ \dots \\ x_n - y_n^2 \\ s_1 y_1 + s_1 + y_1 \\ \dots \\ s_r y_r + s_r + y_r \\ s_{r+1} y_{r+1} \\ \dots \\ s_n y_n \end{pmatrix} = 0. \quad (17)$$

The Jacobian of (17) is nonsingular at  $(\bar{x}, \bar{s}, \bar{y})$ . Consequently the system (17) has a locally unique regular solution  $(\bar{x}, \bar{s}, \bar{y})$ . Thus, we have reduced the solution of the problem (14) to solving system (17) that is independent of the objective function  $f(x)$ .

## References

- [1] F. Facchinei, A. Fisher, and C. Kanzow, On the accurate identification of active constraints, *SIAM J. Optim.*, 9 (1998), pp. 14–32.
- [2] A. F. Izmailov, and A. A. Tret'yakov, *Factor-analysis of nonlinear mappings*, Nauka, Moscow, 1994 (in Russian).
- [3] A. A. Tret'yakov, and J. E. Marsden, Factor-analysis of nonlinear mappings:  $p$ -regularity theory, *Communications on Pure and Applied Analysis*, 2 (2003), pp. 425–445.
- [4] S. J. Wright, An algorithm for degenerate nonlinear programming with rapid local convergence, *SIAM Journal on Optimization*, 15 (2005), pp. 673–696.