

New results and proofs in nonlinear optimization

Agnieszka Prusińska^{1*}, Krzysztof Szkatuła^{2†} and
Alexey Tret'yakov^{1†}

^{1*}Faculty of Science, University of Siedlce, ul. Konarskiego 2, Siedlce,
21-100, Poland.

²Systems Research Institute, Polish Academy of Sciences, ul. Newelska
6, Warsaw, 01-447, Poland.

*Corresponding author(s). E-mail(s): aprus@uws.edu.pl;

Contributing authors: szkatulk@ibspan.waw.pl;

alexey.tretiyakov@uws.edu.pl;

[†]These authors contributed equally to this work.

Abstract

This paper presents several new perspectives on fundamental results in nonlinear optimization and mathematical programming. Elementary proofs are developed for classical optimality results, including the Lagrange multiplier theorem, the Karush–Kuhn–Tucker theorem, Farkas’ lemma, and the cone closure theorem. The proposed arguments avoid a number of traditional analytical tools, such as the Implicit Function Theorem, separation theorems, cone closure techniques, penalty methods, and limiting procedures, while preserving mathematical rigor.

Keywords: optimization theory, mathematical programming, Karush–Kuhn–Tucker conditions, Lagrange multipliers, Farkas’ lemma, Newton’s method, degenerate optimization, p -regularity

MSC Classification: 65K05 , 49M15 , 90C30 , 65K10 , 49M37 , 90C46

1 Introduction

Theoretical foundations of nonlinear optimization have been extensively studied over the last several decades and constitute one of the cornerstones of modern applied mathematics. Classical optimality conditions, including the Lagrange multiplier theorem,

the Karush–Kuhn–Tucker (KKT) theorem, Farkas’ lemma, and second-order optimality conditions, play a central role in nonlinear programming, optimal control, variational analysis, and numerical optimization.

The standard derivations of these results are based on sophisticated analytical tools. For example, the classical proofs of the Lagrange multiplier theorem frequently employ the Implicit Function Theorem or Lyusternik’s theorem [4, 17, 20]. Similarly, the proof of the KKT theorem is traditionally based on separation arguments, Farkas’ lemma, cone duality, or limiting techniques [4, 20, 23]. These methods have become standard in optimization theory and form the basis of many modern textbooks.

The role of Farkas’ lemma extends well beyond linear programming. Together with cone separation results, it provides one of the fundamental tools for establishing first-order optimality conditions, constraint qualifications, duality theory, and numerical algorithms. The relationship between Farkas’ lemma and cone closure has therefore been investigated in numerous studies and constitutes an essential component of modern optimization theory.

Another important research direction concerns optimization problems with degenerate constraints and degenerate systems of nonlinear equations. Classical Newton-type methods generally require nonsingularity of the Jacobian matrix and therefore become ineffective in degenerate situations. To overcome these difficulties, several generalized Newton methods have been developed, including approaches based on the theory of p -regularity and higher-order approximations [5, 10, 24].

The theory of p -regularity has proved to be an effective framework for the analysis of singular nonlinear mappings and degenerate optimization problems. Its applications include higher-order optimality conditions, nonlinear equations, constrained optimization, and generalized Newton-type methods. In particular, the two-factor Newton method provides quadratic convergence under appropriate regularity assumptions even when the classical Jacobian matrix is singular [3, 5, 10].

Recent years have also witnessed renewed interest in Newton-type methods for convex optimization and nonlinear systems. Several authors have investigated convergence properties, stabilization techniques, and modifications of Newton iterations for both convex quadratic programming and nonlinear equations [6, 18, 21, 22, 25].

The present paper differs from the above approaches in several respects. Rather than developing another numerical algorithm, we revisit several fundamental results of optimization theory from a unified perspective. First, elementary proofs are proposed for the Lagrange multiplier theorem, the Karush–Kuhn–Tucker theorem, Farkas’ lemma, and the cone closure theorem. The proposed arguments avoid a number of analytical tools commonly used in classical derivations while preserving mathematical rigor. Second, we investigate the equivalence of optimization problems with equality and inequality constraints through the framework of p -regularity. Third, the applicability of the two-factor Newton method to degenerate Kuhn–Tucker systems is established. Finally, new solvability properties of Newton systems associated with smooth convex optimization problems are discussed together with representative degenerate optimization examples.

The remainder of the paper is organized as follows. Section 2 presents a new elementary proof of the classical Lagrange multiplier theorem. Section 3 develops an

elementary proof of the Karush–Kuhn–Tucker theorem based on a projection argument. Sections 4 and 5 investigate Farkas’ lemma and the cone closure theorem, leading to a generalized version of Farkas’ lemma with weaker assumptions. Section 6 discusses the equivalence of optimization problems with equality and inequality constraints and establishes new second-order optimality conditions. Sections 7 and 8 apply the theory of p -regularity to degenerate nonlinear systems and degenerate Kuhn–Tucker systems through the 2-factor Newton method. Section 9 investigates new properties of smooth convex functions and derives convergence results for Newton’s method. Finally, Section 10 illustrates the theoretical developments with representative degenerate optimization problems, while Section 11 summarizes the main conclusions.

2 The elementary proof of the Lagrange multipliers theorem

The classical proof of the Lagrange multiplier theorem relies on advanced analytical tools. In this section we present a substantially simpler proof based solely on the Weierstrass theorem (see, e.g. [12]).

We consider the following equality-constrained optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \varphi(x) \\ \text{subject to } F(x) = 0, \end{aligned} \tag{1}$$

where

$$F = (f_1, \dots, f_m)^\top, \quad F, \varphi \in C^2(\mathbb{R}^n), \quad f_i : \mathbb{R}^n \rightarrow \mathbb{R}, \quad i = 1, \dots, m.$$

The following theorem gives the classical first-order necessary optimality conditions for problem (1) (cf. Bertsekas [4], Ioffe and Tikhomirov [17]).

Theorem 1 *Suppose that $\varphi, F \in C^2(\mathbb{R}^n)$ and that the Jacobian matrix $F'(x^*)$ has full row rank.*

If x^* is a local solution of problem (1), then

$$\varphi'(x^*) \in \text{Lin}\{f'_1(x^*), \dots, f'_m(x^*)\}.$$

Proof Assume, to obtain a contradiction, that

$$\varphi'(x^*) \notin \text{Lin}\{f'_i(x^*) \mid i = 1, \dots, m\}.$$

Consequently,

$$\varphi'(x^*) = h + \bar{h}, \text{ where } h \in \text{Ker } F'(x^*), \quad h \neq 0, \quad \bar{h} \in \text{Lin}\{f'_i(x^*) \mid i = 1, \dots, m\}.$$

In this case $\langle \varphi'(x^*), h \rangle \neq 0$, and suppose moreover that $\langle \varphi'(x^*), h \rangle > 0$. Then for $\alpha \in (0, \epsilon)$ and for every $\xi \in \mathbb{R}^m$, such that $\|\xi\| = 1$,

$$\left\| F\left(x^* - \alpha \cdot h + \alpha^{3/2} \cdot F'(x^*)^T \cdot \xi\right) \right\|^2 \geq c \cdot \alpha^3 \text{ and } \|F(x^* - \alpha \cdot h)\|^2 \leq c \cdot \alpha^4,$$

where $\epsilon > 0$ is sufficiently small and $c > 0$ is independent constant. By the Weierstrass theorem, there exists $\eta : \mathbb{R}^1 \rightarrow \mathbb{R}^m$, such that:

$$\min_{\xi \in \mathbb{R}^m} \left\| F \left(x^* - \alpha h + \alpha^{3/2} F'(x^*)^T \xi \right) \right\|^2 = \left\| F \left(x^* - \alpha h + \alpha^{3/2} F'(x^*)^T \eta(\alpha) \right) \right\|^2$$

and $\|\eta(\alpha)\| \leq 1$. Hence,

$$\left(\left\| F \left(x^* - \alpha \cdot h + \alpha^{3/2} \cdot F'(x^*)^T \cdot \eta(\alpha) \right) \right\|^2 \right)' = 0_m$$

and, as the consequence,

$$2F' \left(x^* - \alpha h + \alpha^{3/2} F'(x^*)^T \eta(\alpha) \right) \cdot F'(x^*)^T \cdot F \left(x^* - \alpha h + \alpha^{3/2} F'(x^*)^T \eta(\alpha) \right) = 0_m.$$

Therefore,

$$F \left(x^* - \alpha \cdot h + \alpha^{3/2} \cdot F'(x^*)^T \cdot \eta(\alpha) \right) = 0_m$$

and

$$\varphi \left(x^* - \alpha \cdot h + \alpha^{3/2} \cdot F'(x^*)^T \cdot \eta(\alpha) \right) = \varphi(x^*) - \alpha \cdot \langle \varphi'(x^*), h \rangle + o(\alpha) < \varphi(x^*).$$

This contradiction completes the proof. $\square$

The above proof is constructive and relies only on elementary arguments together with the Weierstrass theorem. In contrast to the standard derivations, it does not require the Implicit Function Theorem or Lyusternik's theorem (see, e.g., [17], [4], [20]).

3 New elementary proof of the Karush-Kuhn-Tucker theorem

The Karush–Kuhn–Tucker theorem is a cornerstone of nonlinear optimization. Existing proofs typically employ separation arguments or Farkas' lemma. We present an alternative proof based on a simple projection argument.

Consider the following mathematical programming problem:

$$\begin{aligned} \varphi(x^*) &= \min_{x \in \mathbb{R}^n} \varphi(x) \\ \text{subject to } & g_i(x) \leq 0, \quad i = 1, \dots, m, \\ \text{where } & x^* \in \mathbb{R}^n, \quad g_i \in \mathcal{C}^1(\mathbb{R}^n). \end{aligned} \tag{2}$$

Karush-Kuhn-Tucker (KKT) theorem is the fundamental result characterizing the properties of a solution of the problem (2). The KKT theorem establishes first-order necessary optimality conditions for problem (2).

Classical proofs of the KKT theorem rely on advanced mathematical tools such as the Implicit Function Theorem, the cone closure theorem, Farkas' lemma, separation theorems, and limiting arguments. The proof presented in this paper is considerably simpler and is based on the following lemma.

Lemma 1 Let $\mathcal{K} \subset \mathbb{R}^n$ be a closed convex set such that $0 \notin \mathcal{K}$. Let $p = \text{Proj}_{\mathcal{K}}(0)$ denote the projection of the origin onto $\mathcal{K}$. Then the following inequality

$$\langle z - p, p \rangle \geq 0 \text{ for all } z \in \mathcal{K} \tag{3}$$

holds.

Since the proof of Lemma 1 is straightforward, it is omitted. The KKT theorem can be stated as follows.

Theorem 2 *Let x^* be the solution of the problem (2), where $\varphi, g_i \in C^1(\mathbb{R}^n)$, $i = 1, \dots, m$. Then there exist nonnegative multipliers $\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*$, fulfilling the following conditions*

$$\lambda_0^* \cdot \varphi'(x^*) + \sum_{i=1}^m \lambda_i^* \cdot g_i'(x^*) = 0_m \text{ and } \sum_{i=0}^m \lambda_i^* = 1, \lambda_i^* \cdot g_i(x^*) = 0, i = 1, \dots, m. \quad (4)$$

Classical proofs of this theorem are based on concepts such as the Implicit Function Theorem, the cone closure theorem, Farkas' lemma, separation theorems, and limiting arguments. In contrast, the proof presented here is based solely on the straightforward lemma 1, which is based solely on the projection property.

Proof Let us denote the active set of constraints by $I(x^*) = \{i \in 1, \dots, m \mid g_i(x^*) = 0\}$. Obviously, $1 \leq |I(x^*)| \leq m$. The condition $\lambda_i^* \cdot g_i(x^*) = 0$ means that

$$\text{for } i \in \{1, \dots, m\}, \lambda_i^* = 0 \text{ when } g_i(x^*) < 0.$$

Let us denote

$$\mathcal{K} = \left\{ y \in \mathbb{R}^n \mid y = \lambda_0 \cdot \varphi'(x^*) + \sum_{i=1}^m \lambda_i \cdot g_i'(x^*), \lambda_i \geq 0, i = 0, \dots, m, \sum_{i=0}^m \lambda_i = 1 \right\}.$$

Our goal is to prove that

$$0_n \in \mathcal{K}, \quad (5)$$

which is equivalent to (4). Suppose, to obtain a contradiction, that $0 \notin \mathcal{K}$. Since $\varphi'(x^*) \in \mathcal{K}$, the following inequalities

$$\langle \varphi'(x^*), -p \rangle \leq -\|p\|^2 < 0 \quad (6)$$

hold and because $g_i'(x^*) \in \mathcal{K}$, $i = 1, \dots, m$ the following inequalities

$$\langle g_i'(x^*), -p \rangle \leq -\|p\|^2 < 0 \quad (7)$$

hold, as well. Consequently,

$$\begin{aligned} \varphi(x^* - \alpha \cdot p) &< \varphi(x^*) \text{ and} \\ g_i(x^* - \alpha \cdot p) &< g_i(x^*) < 0, i = 1, \dots, m, \end{aligned}$$

for sufficiently small $\alpha > 0$. This contradicts the local optimality of x^* for problem (2). $\square$

4 Farkas' lemma and the cone closure theorem

The so-called theorems of the alternative play a crucial role in the derivation of optimality conditions and in the design of numerical methods. One of these results can be stated as follows (see, e.g. [10]).

Theorem 3 (Farkas' lemma) *Let $\eta, \xi_i \in \mathbb{R}^n$, $i = 1, \dots, m$. Assume that for every $x \in \mathbb{R}^n$ satisfying $\langle \xi_i, x \rangle \geq 0$, $i = 1, \dots, m$, the following inequality holds (Farkas' condition)*

$$\langle \eta, x \rangle \geq 0. \quad (8)$$

Then

$$\eta \in \text{cone} \{ \xi_i, i = 1, \dots, m \},$$

where $\text{cone} \{ \xi_i, i = 1, \dots, m \} = \{ y \in \mathbb{R}^n \mid y = \sum_{i=1}^m \alpha_i \cdot \xi_i^*, \alpha_i \geq 0, i = 1, \dots, m \}$.

Proof (See [10]) Let us consider the following function

$$\Phi(x) = \langle \eta, x \rangle + \sum_{i=1}^m (\langle \xi_i, x \rangle_-)^2, \text{ where } a_- = \begin{cases} a, & a \leq 0 \\ 0, & a > 0 \end{cases}, a \in \mathbb{R}^n.$$

If Farkas' condition holds, then there exists $y^* = \operatorname{argmin}_{x \in \mathbb{R}^n} \Phi(x)$.

By the first-order necessary optimality condition (Fermat's theorem), we have either $\Phi'(y^*) = 0_n$ or $\eta + \sum_{i=1}^m 2\langle \xi_i, y^* \rangle_- \xi_i = 0_n$. Defining $-\lambda_i^* := 2\langle \xi_i, y^* \rangle_-$, we obtain $\eta = \sum_{i=1}^m \lambda_i^* \cdot \xi_i^*$, $i = 1, \dots, m$, i.e. $\eta \in \operatorname{Cone} \{\xi_i, i = 1, \dots, m\}$. $\square$

Remark 1 The explicit form of the Farkas multipliers follows directly from the proof of Theorem 3.

Existing proofs of Theorem 3 are typically based on separation theorems, limiting arguments, and the cone closure theorem. Proofs based on the cone closure theorem are generally more involved. Below, we present a shorter proof based on Farkas' lemma. The proof relies solely on Farkas' lemma.

Theorem 4 (Cone closure theorem) *The cone*

$$\mathcal{K} = \{y \in \mathbb{R}^n \mid y = \sum_{i=1}^m \lambda_i^* \cdot a_i, a_i \in \mathbb{R}^n, \lambda_i \geq 0, i = 1, \dots, m\}$$

is closed.

Proof (See [10]) Let $\{y_i\} \subset \mathcal{K}, i = 1, \dots$ be a sequence such that $\|y_i\| \leq c$ and $y_i \rightarrow y$ when $i \rightarrow \infty$. It remains to show that $y \in \mathcal{K}$. Indeed, for every $x \in \mathbb{R}^n$ satisfying $\langle a_j, x \rangle \geq 0, j = 1, \dots, m$, it follows that $\langle y_i, x \rangle \geq 0, i = 1, \dots$. Passing to the limit as $i \rightarrow \infty$, we obtain for every $x \in \mathbb{R}^n$ such that $\langle a_j, x \rangle \geq 0, j = 1, \dots, m$ it follows that

$$\lim_{i \rightarrow \infty} \langle y_i, x \rangle = \langle y, x \rangle \geq 0.$$

By Farkas' lemma, $y \in \operatorname{cone} \{a_j, j = 1, \dots, m\}$. Hence, $\mathcal{K}$ is a closed set. $\square$

Therefore, Farkas' lemma and the cone closure theorem are equivalent. This establishes a direct relationship between two fundamental results of optimization theory.

5 Redundancy of conditions in the Farkas' lemma

Most fundamental optimality results for optimization problems with inequality constraints are derived using Farkas' lemma. However, condition (8) requiring that every $x \in \mathbb{R}^n$ satisfying $\langle \xi, x \rangle \geq 0, i = 1, \dots, m$ also satisfies $\langle \eta, x \rangle \geq 0$ is stronger than necessary. It is sufficient to verify the condition on a finite set of vectors z_j . If condition (8) holds for this finite set, that is $\forall z_j, j = 1, \dots, k$ whenever $\langle \xi_i, z_j \rangle \geq 0, i = 1, \dots, m \Rightarrow \langle \eta, z_j \rangle \geq 0, j = 1, \dots, k$, then $\eta = \sum_{i=1}^m \alpha_i \xi_i, \alpha_i > 0, i = 1, \dots, m$. For example,

let $\xi_1 = (0, 1)^\top$, $\xi_2 = (1, 0)^\top$, $\eta = (1, 1)^\top$, $z_1 = (1, 0)^\top$, $z_2 = (0, 1)^\top$. Then $\langle \xi_i, z_j \rangle \geq 0$, $i = 1, 2$ and $\langle \eta, z_j \rangle \geq 0$, $j = 1, 2 \Rightarrow \eta = \alpha_1 \xi_1 + \alpha_2 \xi_2$, $\alpha_1 = 1$, $\alpha_2 = 1$.

The finite set $\{z_j\}$ is constructed as follows. Let $\text{span}\{\xi_i \mid i = 1, \dots, m\} \triangleq \mathbb{R}^p$, $p \leq n$, $\xi_i \in \mathbb{R}^n$, $\eta \in \mathbb{R}^n$. Then decompose $\mathbb{R}^n = \mathbb{R}^p \oplus \mathbb{R}^{n-p}$ and let $y_1, \dots, y_{n-p}$ be a basis of $\mathbb{R}^{n-p}$. Furthermore,

1) For η to belong to $\text{cone}\{\xi_i \mid i = 1, \dots, m\}$, it is necessary that $\eta \in \text{span}\{\xi_i \mid i = 1, \dots, m\}$, which is equivalent to $\langle \eta, y_j \rangle = 0$, $j = 1, \dots, n-p$. Equivalently, this condition may be expressed as: if $\langle \xi_i, \pm y_j \rangle \geq 0$ then $\langle \eta, \pm y_j \rangle \geq 0$, $j = 1, \dots, n-p$.

2) Assume now that $\eta \in \text{span}\{\xi_i \mid i = 1, \dots, m\}$. Next construct in $\mathbb{R}^p$ the convex cone $\mathcal{K}_p := \{z \in \mathbb{R}^p \mid \langle \xi_i, z \rangle \geq 0, i = 1, \dots, m\}$. Let $x_1, \dots, x_r$ be its extreme (one-dimensional) edges. Since the cone is polyhedral, it has finitely many extreme rays, denoted by $x_1, x_2, \dots, x_r$. We set $z_1 = x_1, \dots, z_r = x_r, z_{r+1} = y_1, z_{r+2} = -y_1, \dots, z_{r+2(n-p)-1} = y_{n-p}, z_{r+2(n-p)} = -y_{n-p}$, that is, we define the finite set $\{z_j\}$, $j = 1, \dots, k$, $k = r + 2(n-p)$.

Theorem 5 (Generalized Farkas' Lemma) *Let $\langle \xi_i, z_j \rangle \geq 0$, $i = 1, \dots, m$, $\Rightarrow \langle \eta, z_j \rangle \geq 0$, $j = 1, \dots, k$, where k is defined above. Then*

$$\eta = \sum_{i=1}^m \alpha_i \xi_i. \quad (9)$$

Proof (See [12]) Since $\eta \in \text{span}\{\xi_i \mid i = 1, \dots, m\}$. We prove that $\eta = \sum_{i=1}^m \alpha_i \xi_i$, $\alpha_i \geq 0$, $i = 1, \dots, m$. Indeed,

$$\forall x: \langle \xi_i, x \rangle \geq 0, i = 1, \dots, m \Rightarrow x \in \mathbb{R}^n \cap \mathbb{R}^p \Rightarrow x \in \mathcal{K}_p.$$

Therefore, $x = \beta_1 x_1 + \dots + \beta_r x_r$, $z_1 = x_1, \dots, z_r = x_r$, $\beta_j \geq 0$, $j = 1, \dots, r$. Consequently,

$$\langle \eta, x_1 \rangle \geq 0, \dots, \langle \eta, x_r \rangle \geq 0 \Rightarrow \langle \eta, \beta_1 x_1 + \dots + \beta_r x_r \rangle = \langle \eta, x \rangle \geq 0.$$

Thus, the assumptions of Farkas' lemma are satisfied, which yields (9). $\square$

6 On the Reduction of Optimization Problems with Inequality Constraints to Equality-Constrained Problems

When solving optimization problems with inequality constraints of the form (2), it is often convenient to transform the inequality constraints into equality constraints by introducing artificial variables.

The strict complementarity (SCS, cf. Evtushenko [9]) is commonly introduced to ensure that the optimality conditions for problems with inequality and equality constraints are equivalent. Otherwise, the two sets of optimality conditions are generally not equivalent, and such a reduction is of limited practical value (cf. [4]). To explain this, we introduce the Lagrangian for the problem (2):

$$L(x, \lambda) = \varphi(x) + \sum_{i=1}^m \lambda_i g_i(x).$$

The first-order necessary Kuhn–Tucker optimality conditions are

$$L'_x(x, \lambda) = \varphi'(x) + \sum_{i=1}^m \lambda_i g'_i(x) \quad (10)$$

and the complementary condition

$$\lambda_i g_i(x) = 0, \quad \lambda_i \geq 0, \quad i = 1, \dots, m. \quad (11)$$

The second-order sufficient optimality conditions for problem (2) are as follows. if at a feasible point x^* , $(g_i(x^*) \leq 0, i = 1, \dots, m)$, the necessary conditions (10) and (11) are satisfied for some λ^* and

$$\langle L''_{xx}(x^*, \lambda^*)h, h \rangle > 0, \quad \forall h \neq 0, \quad (12)$$

such that $\langle g'_i(x^*, h) \rangle \leq 0, i \in I(x^*) = \{i \in \{1, \dots, m\} \mid g_i^*(x) = 0\}$ then x^* is a local minimum point of the problem (2), cf. Polyak [23].

For equality-constrained optimization problems, feasibility of the candidate solution is automatically enforced by the constraints themselves. This property does not hold for the inequality-constrained formulation (2), since conditions (10) and (11) alone do not guarantee feasibility. Therefore, we reformulate problem (2) as an equality-constrained optimization problem by introducing artificial variables $s = (s_1, \dots, s_m)^T$

$$\begin{aligned} \varphi(x^*) &= \min_{x \in \mathbb{R}^n} \varphi(x) \\ \text{subject to } &g_i(x) + s_i^2 = 0, \quad i = 1, \dots, m. \end{aligned} \quad (13)$$

The problems (2) and (13) are equivalent (cf. [11]). The necessary optimality conditions for the problem (13) are as follows.

$$\begin{aligned} \varphi'(x) + \sum_{i=1}^m \lambda_i g'_i(x) &= 0_n, \\ \lambda_1 \begin{pmatrix} 2s_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + \lambda_m \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 2s_m \end{pmatrix} &= 0_m = 2\lambda s, \\ g_i(x) + s_i^2 &= 0, \quad i = 1, \dots, m, \end{aligned} \quad (14)$$

where $\lambda s = (\lambda_1 s_1, \dots, \lambda_m s_m)^T$. The system (14) does not determine the signs of the multipliers λ_i . However, feasibility of the solution follows directly from system (14). Let us denote $\bar{g}_i(x, s) = g_i(x) + s_i^2, i = 1, \dots, m$, and introduce the Lagrangian of the problem (13):

$$\bar{L}(x, s, \lambda) = \varphi(x) + \sum_{i=1}^m \lambda_i (g_i(x) + s_i^2). \quad (15)$$

Let $g(x) = (g_1(x), \dots, g_m(x))^T$, $s^2 = (s_1^2, \dots, s_m^2)^\top$. Assume that the constraint regularity condition (CRC) holds at the solution x^* if the vectors $\{g'_i(x^*)\}$, $i \in I(x^*)$ are linearly independent. Then the sufficient optimality conditions mean that the matrix $\bar{L}''_{(x,s)}(x^*, s^*, \lambda^*)$ is positive definite on the tangent space $\text{Ker } \bar{g}'(x^*, s^*)$ i.e.

$$\langle \bar{L}''_{xs}(x^*, s^*, \lambda^*)h, h \rangle \geq \alpha \|h\|^2, \quad \alpha > 0, \quad \forall h \in \text{Ker } \bar{g}'(x^*, s^*). \quad (16)$$

However, the matrix $\bar{L}''_{xs}(x^*, s^*, \lambda^*)$ is not positive definite on $\bar{g}'(x^*, s^*)$, if the condition SCS does not hold, that is. λ_i^* and s_i^* do not vanish simultaneously for $i = 1, \dots, m$. That is, the sufficient conditions (12) and (16) are not equivalent. This is because the second-order necessary conditions for equality-constrained optimization can be weakened, leading to weaker sufficient optimality conditions. Let us explain the following. Suppose that there exists $h_2 \in \text{Ker } F'(x^*)$ for the problem (8) such that

$$\langle L''_{xx}(x^*, \lambda^*)h_2, h_2 \rangle = 0 \quad (17)$$

$L(x, \lambda) = \varphi(x) + \sum_{i=1}^m \lambda_i f_i(x)$ and hence the second order sufficient conditions are not fulfilled. Let $\text{Ker } F'(x^*) = \mathbb{R}_1^n$ and $\mathbb{R}_2^n = (\mathbb{R}_1^n)^\perp$. Then $\mathbb{R}^n = \mathbb{R}_1^n + \mathbb{R}_2^n$. The following theorem holds (see, e.g. [10]).

Theorem 6 (Necessary and sufficient optimality conditions of a special kind) *Let $\varphi, F \in C^3(\mathbb{R}^n)$ and for any $h_2 \in \text{Ker } F'(x^*)$ ($\|h\| = 1$) such that*

$$L''_{xx}(x^*, \lambda^*)[h_2]^2 = 0 \text{ or } (L''_{xx}(x^*, \lambda^*)h_2 = 0),$$

the following conditions hold

$$\|F''(x^*)[h_2]^2\| \geq \alpha > 0, \quad (18)$$

and there exist $h_1 \in \mathbb{R}_2^n$ and $c(h_2) \geq c > 0$, $\|h_1\| = 1$ which satisfy the relation

$$F'(x^*)[h_1] + \frac{1}{2}F''(x^*)[c(h_2)h_2]^2 = 0. \quad (19)$$

Then, (necessity) if x^* is a solution of (8) then $\exists \lambda^*$ such that $L'_x(x^*, \lambda^*) = 0$ and

$$L''_{xx}(x^*, \lambda^*)[h_1]^2 \geq 0. \quad (20)$$

Moreover, (sufficiency) if $\forall h_1, h_2$ satisfying (18) and (19) there exists $\beta > 0$ such that $L'_x(x^*, \lambda^*) = 0$ and

$$L''(x^*, \lambda^*)[h_1]^2 \geq \beta > 0, \quad (21)$$

then x^* is the local minimum of the problem (8).

Remark 2 The element h_1 may not belong to $\text{Ker } F'(x^*)$.

Proof of Theorem 6. (see [10]) Necessity: We show that $\forall t \in (0, \epsilon)$, $\epsilon > 0$ and sufficiently small, there exists $\omega(t) : \mathbb{R} \rightarrow \mathbb{R}^n$, such that the arc

$$\gamma(t) = x^* + c_2 t h_2 + t^2 h_1 + \omega(t) \in M(x^*) = \{x \in \mathbb{R}^n \mid F(x) = F(x^*) = 0\}$$

and $\|\omega(t)\| = o(t^2)$. In fact

$$\begin{aligned} F(x^* + c_2 t h_2 + t^2 h_1) &= F(x^*) + F'(x^*)[c_2 t h_2 + t^2 h_1] + \\ &+ \frac{1}{2} F''(x^*)[c_2 t h_2 + t^2 h_1]^2 + o(t^2) = \\ &= t^2 F'(x^*) h_1 + \frac{1}{2} F''(x^*)[c_2 t h_2]^2 + o(t^2) = o(t^2). \end{aligned} \quad (22)$$

Since $F'(x^*)$ is nonsingular we can apply the contraction mapping principle (cf. [17]) to the mapping

$$\Phi_t(x) = x - (F'(x^*))^{-1} F(x^* + c_2 t h_2 + t^2 h_1 + x) \quad (23)$$

which implies the existence of $\omega(t)$ such that

$$F(x^* + c_2 t h_2 + t^2 h_1 + \omega(t)) = 0 \quad (24)$$

where

$$\|\omega(t)\| = o(t^2) \text{ and } c_2 > 0 \text{ independent constant.} \quad (25)$$

Next, we apply the standard technique to the relation

$L(x^* + c_2 t h_2 + t^2 h_1 + \omega(t), \lambda^*) \geq L(x^*, \lambda^*)$ and we get $L''_{xx}(x^*, \lambda^*)[h_1]^2 \geq 0$. This completes the proof of necessity.

For sufficiency, assume the contrary. There is a sequence of points $\{x_k\} \in M(x^*)$ and $x_k \rightarrow x^*$ for $k \rightarrow \infty$ such that $\varphi(x_k) < \varphi(x^*)$, using the proof of necessity and the standard technique of proving the sufficiency conditions, we come to a contradiction with the relation (21). $\square$

Example 1 ([12]) Let us consider the problem (13) in the form

$$\min \varphi(y) \quad (26)$$

for $F(x) = g(y) + s^2 = 0_m$. Here the optimization variable is (y, s) . To avoid confusion with the notation used previously, we replace x with y in the problem (2). If the point (x^*, s^*) , $s^* = 0$ is a solution to the problem (2), then for $h_2 = (0, \bar{s})^\top$ there exists $h_1 = (\bar{y}, 0)^\top$ such that

$$F'(x^*)[h_1] + \frac{1}{2} F''(x^*)[h_2]^2 = 0 \Leftrightarrow g'(y^*)\bar{y} + \frac{1}{2} 2\bar{s}^2 = 0. \quad (27)$$

Since the optimality conditions are local in nature, we will consider only active constraints, that is, $i \in I(x^*)$ and, without loss of generality, we consider $r = m$. The condition (18) is satisfied if $\bar{s} \neq 0$, the condition (19) is satisfied due to the regularity of constraints $\{g_i(x)\}$, $i = 1, \dots, m$ at the point x^* , with $c_2 = 1$. Thus, for problem (13) Theorem 6 applies since $L''_{xx}(x^*, \lambda^*)[h_2]^2 = 0$, although $h_2 = (0, \bar{s})^\top \in \text{Ker } F'(x^*) = \{(\bar{y}, \bar{s}) \mid g'(y^*)\bar{y} + 0 \cdot \bar{s} = 0\}$. Moreover, the following equivalence theorem holds.

Theorem 7 (Equivalence of Optimality Conditions) *Let $\varphi, g \in C^3(\mathbb{R}^n)$. Then the optimality conditions (10), (11), (12) and (21) are equivalent.*

Proof (See [10]) It is necessary to check only a sufficient part of the optimality conditions. Assume, that conditions (10) – (12) are fulfilled. Then, taking $\bar{s} \neq 0$, we get $\bar{y} = \{g'(y^*)\}^{-1}(-\bar{s}^2)$ from the condition (19). Moreover, $L''_{xx}(y^*, s^* \lambda^*)[y]^2 \geq \beta \|\bar{y}\|^2$, due to the sufficient conditions (12). Here $\bar{y}$ plays the role of h_1 in Theorem 6 and therefore (21) holds, and hence the sufficient conditions of Theorem 6 are satisfied. This proves one implication.

Now, suppose that the sufficient conditions (21) for the problem (26) are satisfied. Then $\forall h \ g'(y^*)h < 0_m$ there will be $\bar{s} \neq 0$ and, therefore, (21) holds for $h_1 = h$, i.e. $L''_{xx}(y^*, 0, \lambda^*)[h]^2 \geq \beta \|h\|^2$, which means that $L''_{xx}(y^*, \lambda^*)[h]^2 \geq \beta \|h\|^2$, where $L(y, \lambda)$ plays the role of the function $L(x, \lambda) = \varphi(x) + \sum_{i=1}^m \lambda_{ii} g_i(x)$ in the problem (2) and finally, the sufficient conditions (12) for the problem (2) are satisfied. This completes the proof. $\square$

7 The 2-Factor Method for Solving Degenerate Systems of Nonlinear Equations

Numerical methods for solving systems of nonlinear equations, particularly those arising from first-order optimality conditions, are typically based on Newton's method (cf. [9]). However, if the Jacobian matrix $\Phi'(\cdot)$ is singular at the solution x^* , then Newton's method whose iterative scheme is

$$x_{k+1} = x_k - \{F'(x_k)\}^{-1} F(x_k), \quad k = 0, 1, \dots, \quad (28)$$

may fail or become inapplicable. In this case, the 2-factor method can be applied. It guarantees convergence and, moreover, quadratic convergence under suitable assumptions. In the general setting, the 2-factor iteration takes the form

$$x_{k+1} = x_k - \{\Phi'(x_k) + P\Phi''(x_k)h\}^{-1} (\Phi(x_k) + P\Phi'(x_k)h), \quad k = 0, 1, \dots, \quad (29)$$

where P denotes the orthogonal projection onto $(\text{Im } \Phi'(x^*))^\perp$ and h is a fixed unit vector satisfying $\{\Phi'(x^*) + P\Phi''(x^*)h\}$ is nonsingular, i.e. the mapping $\Phi(\cdot)$ is 2-regular at x^* in the direction h . In this case, the following theorem holds (see, e.g. [10]).

Theorem 8 *Let $\Phi \in C^3(\mathbb{R}^n)$ and $\Phi(\cdot)$ be 2-regular at the solution x^* in the direction h . Then, for every $x_0 \in U_\varepsilon(x^*)$ the sequence $\{x_k\}$ generated by (29) converges to the solution x^* , and the quadratic convergence rate is valid:*

$$\|x_{k+1} - x^*\| = c \cdot \|x_k - x^*\|^2, \quad k = 0, 1, \dots, \quad (30)$$

where $\varepsilon > 0$ is sufficiently small where $c > 0$ is a constant independent of k .

The proof can be found in [3]. In the next section, we show how the 2-factor method can be applied to solve optimization problems with inequality constraints after introducing artificial variables.

8 The 2-Factor Newton Method for Degenerate Kuhn–Tucker Systems

When solving optimization problems with inequality constraints by introducing artificial variables, it is usually necessary to assume the strict complementarity condition (SCS) to ensure that the Kuhn–Tucker system is nonsingular. Otherwise, the applicability of Newton-type methods cannot, in general, be guaranteed. In this section, we

investigate the effectiveness of this approach from the point of view of applying the 2-factor method to solve systems of degenerate optimality conditions and prove that the resulting Kuhn–Tucker system is 2-regular. Accordingly, the problem (2)

$$\begin{aligned} \min \varphi(x) \\ g_i(x) \leq 0, \quad i = 1, \dots, m, \end{aligned}$$

is replaced by the equivalent problem (13)

$$\begin{aligned} \min \varphi(x) \\ g_i(x) + s_i^2 = 0, \quad i = 1, \dots, m. \end{aligned}$$

Consider the Lagrangian

$$L(x, s, \lambda) = \varphi(x) + \langle g(x) + s^2, \lambda \rangle, \quad \lambda \in \mathbb{R}^m,$$

and denote $u = (x, s, \lambda)$. Then,

$$L'(u) = \begin{pmatrix} L_x(u) \\ 2\lambda s \\ g(x) + s^2 \end{pmatrix}, \quad (31)$$

$$L''(u) = \begin{pmatrix} L_{xx}(u) & 0_{nm} & (g_x(x))^\top \\ 0_{mn} & 2\Lambda & 2S \\ g_x(x) & 2S & 0_{mm} \end{pmatrix}, \quad (32)$$

where $\Lambda = \text{diag}(\lambda)$, $S = \text{diag}(s)$. Consider the system of ℓ nonlinear equations

$$L'(u) = 0_\ell, \quad (33)$$

where $\ell = n + 2m$. Let $u^* = (x^*, s^*, \lambda^*)^\top \in \mathbb{R}^\ell$ be the solution of (33). If $\lambda^* \geq 0_m$, $L'(u^*) = 0_\ell$, then first-order necessary optimality conditions for the problem (2) are satisfied at the point (x^*, λ^*) . If, in addition, the second-order sufficient conditions (12) hold, then the point x^* is a local solution to the problem (2). However, the presence of the inequality $\lambda^* \geq 0_m$ does not yield a system consisting solely of equality constraints. Therefore, we replace the inequality $\lambda \geq 0_m$ by the equivalent system of equations

$$\lambda^\top - y^2 = 0_m, \quad (34)$$

where $y \in \mathbb{R}^m$, $y^2 = (y_1^2, \dots, y_m^2)^\top$ and consider the extended system

$$\Psi(z) = \begin{pmatrix} L'(u) \\ \lambda^\top - y^2 \end{pmatrix} = \begin{pmatrix} 0_\ell \\ 0_m \end{pmatrix} = 0_r, \quad (35)$$

where $z = (x, s, \lambda, y)^\top$, $r = n + 3m$. Thus, the inequality-constrained optimization problem is transformed into a system of nonlinear equations. Moreover, the necessary

Kuhn-Tucker conditions are satisfied at the solution (x^*, λ^*) , and now $\lambda^* \geq 0$. Let

$$\Psi'(z) = \begin{pmatrix} L_{xx}(u) & 0_{nm} & (g_x(x))^{\top} & 0_{nm} \\ 0_{mn} & 2\Lambda & 2S & 0_{mm} \\ g_x(x) & 2S & 0_{mm} & 0_{mm} \\ 0_{mn} & I_m & 0_{mm} & 2Y \end{pmatrix}, \quad (36)$$

where $Y = \text{diag}(y)$. Here I_m is the identity matrix of a dimension $m \times m$. Newton's method for solving system (35) has the form

$$z_{k+1} = z_k - (\Psi'(z_k))^{-1} \Psi(z_k). \quad (37)$$

If SCS is not satisfied at the solution z^* , then the matrix $\Psi'(z^*)$ is singular and the method (37) does not guarantee convergence. The so-called 2-factor Newton method for solving the system (35), which guarantees quadratic convergence, is constructed below. Following the approach of [3] and [5], the mapping

$$\Phi(z) = \Psi(z) + P\Psi'(z)h \quad (38)$$

is introduced and 2-factor operator

$$\Phi'(z) = \Psi'(z) + P\Psi''(z)h, \quad (39)$$

where the $r \times r$ matrix P denotes the orthogonal projection $\Psi''(z^*)h$ onto $(\text{Im } \Psi'(z^*))^{\perp}$. An r -dimensional vector h with unit norm is chosen from the condition that $\Phi'(z^*)$ is singular. Note that the matrix P can be chosen, in general, simply from the condition that the 2-factor operator $\Phi'(z^*)$ is nonsingular.

Theorem 9 *Let $\varphi, g \in C^3(\mathbb{R}^n)$, there exists z^* satisfying system (35). Assume that at (x^*, z^*) the sufficient optimality conditions for (12) and the constraint regularity condition are satisfied. Then 2-factor operator $\Phi'(z^*)$ is nonsingular.*

See [10], [12] We show that the mapping $\Psi(z)$ is 2-regular at the point z^* along $h^{\top} = (0_n^{\top}, 0_m^{\top}, e_0^{\top}, e_1^{\top})$, where $e_0 \in \mathbb{R}^m$ and each i -th component of e_0 equals one if $i \in I_0(x^*) = \{i \in I(x^*) \mid \lambda_i^* = 0, s_i^* = 0\}$, and zero otherwise. Moreover, $e_1 \in \mathbb{R}^m$ and each i -th component of the e_1 vector equals one if $\lambda_i^* = 0$, and zero otherwise. Define a matrix

$$\begin{aligned} \Phi'(z^*) &= \Psi'(z^*) + P\Psi''(z^*)h = \\ &= \begin{pmatrix} L_{xx}(z^*) & 0_{nm} & (g_x(x^*))^{\top} & 0_{nm} \\ 0_{mn} & 2\Lambda^* + 2J_0(x^*) & 2S^* & 0_{mm} \\ (g_x(x^*))^{\top} & 2S^* & 0_{mm} & 0_{mm} \\ 0_{mn} & I_m & 0_{mm} & 2Y^* + 2J_1(x^*) \end{pmatrix} \end{aligned} \quad (40)$$

where the $r \times r$ -diagonal matrix P has the i -th diagonal element equal to zero, if $1 \leq i \leq k$ or $k+m+1 \leq i \leq \ell$. For $k+1 \leq i \leq n+m$ the i -th diagonal element equals one if $i-n \in I_0(x^*)$, and zero otherwise. For $\ell+1 \leq i \leq r$ the i -th diagonal element equals one if $i-\ell \in I_0(x^*)$, and

zero otherwise. Moreover, the $m \times m$ diagonal matrix $J_0(x^*)$ has the j -th diagonal element equal to either zero or one, if $j \in I_0(x^*)$. In turn, the $m \times m$ diagonal matrix $J_1(x^*)$ has the j -th diagonal element equal to one, if $j \in I_0(x^*)$ and zero otherwise. It follows that for $h = (0^\top, 0^\top, e_0^\top, e_1^\top)^\top$ we get

$$P\Psi''(z^*)h = \begin{pmatrix} 0_{nn} & 0_{nm} & 0_{nm} & 0_{nm} \\ 0_{mn} & 2J_0(x^*) & 0_{nm} & 0_{mm} \\ 0_{mn} & 0_{mm} & 0_{mm} & 0_{mm} \\ 0_{mn} & 0_{mn} & 0_{mm} & 2J_1(x^*) \end{pmatrix}$$

which proves (19) and establishes the nonsingularity of the operator $\Phi'(z^*)$. $\square$

Consequently, the 2-factor method applied to the system (35) has the form

$$z_{k+1} = z_k - (\Psi'(z_k) + P\Psi''(z_k)h)^{-1} (\Psi(z_k) + P\Psi'(z_k)h), \quad k = 0, 1, \dots \quad (41)$$

Theorem 10 *Let $\Psi \in C^3(\mathbb{R}^r)$ and for $h \in \mathbb{R}^r$, $\|h\| \neq 0$. Assume that the operator $\Psi'(z^*) + P\Psi''(z^*)h$ is nonsingular. Then 2-factor method (41) converges locally to the solution z^* of the system (35) with a quadratic rate of convergence, i.e.*

$$\|z_{k+1} - z^*\| \leq c\|z_k - z^*\|^2, \quad k = 0, 1, \dots,$$

where $z_0 \in U_\epsilon(z^*)$, $\epsilon > 0$ is sufficiently small, and $c > 0$ is independent of k .

Proof The proof follows from the fact that the 2-factor method (41) is the usual Newton method applied to the system

$$\Psi(z) + P\Psi'(z)h = 0, \quad (42)$$

since z^* is simultaneously a solution of (35) and (42) all assumptions required for the local quadratic convergence of Newton's method are satisfied. $\square$

9 New Properties of Smooth Convex Functions and Newton's Method

Consider the unconstrained minimization problem for a smooth convex function $\varphi(x)$, where $x \in \mathbb{R}^n$, which is sufficiently smooth in a neighborhood of the minimizer x^* . Classical convergence analysis of Newton's method assumes that the Hessian matrix $\varphi''(x^*)$ is positive definite to obtain quadratic convergence (or, more generally, super-linear convergence). Newton's method has been extensively studied in the literature (see, for example, [7]). In this section, we show that even if the Hessian matrix is singular at the minimizer x^* , in a neighborhood of this point the range of the objective function $\varphi'(x)$ belongs to the image of its second derivative and, therefore, the Newton's system with respect to the direction of descent is resolvable for every point in this neighborhood, i.e.

$$\varphi''(x) \cdot z = -\varphi'(x) \quad (43)$$

is solvable with respect to z for any $x \in U_\epsilon(x)$, where $\epsilon > 0$ is sufficiently small. Thus, we can apply the Newton's iterative minimization scheme:

$$x_{k+1} = x_k - (\varphi''(x_k))^{-1} \cdot \varphi'(x_k), \quad k = 0, 1, \dots, \quad x_0 \in U_\epsilon(x^*), \quad (44)$$

where $(\varphi''(x_k))^{-1}$ denotes any right inverse of the Hessian operator. Moreover, the sequence generated by (44) converges linearly to the minimizer which appears not to have been reported previously in this form. This new topological property of convexity (43) and extremality offers a new perspective at numerical optimization methods.

Theorem 11 *Let $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ be a sufficiently smooth convex function. Assume that all derivatives involved are uniformly bounded in some neighborhood $U_\epsilon(x^*)$ of an isolated minimizer x^* in argument and with respect to the order of differentiation, then for any point $x \in U_\epsilon(x^*)$ the system*

$$\varphi''(x) \cdot z = -\varphi'(x)$$

is solvable with respect to z .

Example 2 For a nonconvex function, Theorem 11 is not valid. If $\varphi(x) = x_1^2 + (x_2 - x_1)^2$ at the points $x_2 = 4 \cdot x_1^2$ belonging to an arbitrary sufficiently small neighborhood $U_\epsilon(0)$ the system (43) is not solvable.

For Newton's method (44), the following theorem holds.

Theorem 12 *If the assumptions of Theorem 11 are fulfilled then for the iterative sequence $\{x_k\}$ generated by iteration (44), the convergence is linear, i.e.*

$$\|x_{k+1} - x^*\| \leq q \cdot \|x_k - x^*\|, \quad k = 0, 1, \quad (45)$$

where $q \in (0, 1)$ and x_0 is any point belonging to a sufficiently small neighborhood $U(x^*)$.

Proofs of theorems 11 and 12 can be found in [7].

10 Examples of Degenerate Applied Problems

In this section, we present two representative degenerate applied problems, one of which arises in operations research, the so-called Perfect Partition Problem, whereas the second is the classical Rod Bending Problem arising in elasticity theory. We show that those well-known problems are degenerate at their solution points.

10.1 The Perfect Partition Problem

In operations research, number theory, and theoretical computer science, the Perfect Partition Problem is a well-known NP-complete problem (see [15]). However, owing to the existence of pseudo-polynomial-time dynamic programming algorithms and numerous heuristic methods, that solve many instances, either exactly or approximately, it has often been referred to as "the easiest hard problem" (cf. [19]). The Perfect Partition Problem consists in partitioning the index set $I = \{1, 2, \dots, n\}$ of the integer-valued vector $A = \{a_1, a_2, \dots, a_n\}$ into two disjoint subsets I_1 and I_2 such that

$$\sum_{i \in I_1} a_i = \sum_{j \in I_2} a_j, \quad \text{where } I_1 \cup I_2 = I \text{ and } I_1 \cap I_2 = \emptyset. \quad (46)$$

Let us assume that the Perfect Partition Problem (46) has a solution. In this case (46) is equivalent to finding a nontrivial solution of the system of quadratic equations.

$$F(x) = \begin{pmatrix} x_2^2 - x_1^2 \\ \vdots \\ x_n^2 - x_1^2 \\ a_1 x_1 + \dots + a_n x_n \end{pmatrix}, \quad F(x) = 0. \quad (47)$$

The Jacobian matrix $F'(x^*)$ is singular because for the solution x^* is implied by the following identity

$$F'(x) = \text{diag}(2, \dots, 2, 1) \cdot F(x).$$

Consequently, the classical Newton method cannot be applied to find the solution x^* . Under suitable assumptions, however, the mapping F is 2-regular along $h \in \mathbb{R}^n$, $\|h\| = 1$, i.e. matrix $F''(x^*) \cdot h$ is nonsingular. In this case 2-factor method described in the chapter 7 can be applied to find the solution of the system (47) in the form of an iterative scheme:

$$x_{k+1} = x_k - \{F'(x_k) + PF''(x_k) \cdot h\}^{-1} (F(x_k) + PF'(x_k) \cdot h), \quad k = 0, 1, \dots,$$

which converges to x^* in a sufficiently small neighborhood, i.e. $x_0 \in U_\epsilon(x^*)$, where $\epsilon > 0$ is small.

10.2 Rod Bending Problem

Consider the Rod Bending Problem from elasticity theory. The problem is governed by the equation

$$F(x, \delta) = x'' + (1 + \delta)(x + x^2) = 0, \quad (48)$$

where

$$F : C^2[0, \pi] \times \mathbb{R} \rightarrow C[0, \pi] \text{ and } x(0) = x(\pi) = 0.$$

The objective is to determine a solution of (48) $x(\delta)$ depending on the small parameter $\delta > 0$. The point $(x^*, \delta^*) = (0, 0)$ is the trivial solution of the problem (48). The first derivative of the mapping $F(x, \delta)$ is as follows

$$F'(x, \delta) = [F'_x(x, \delta), F'_\delta(x, \delta)] = \left[\frac{d^2}{dx^2} + (1 + \delta)(1 + 2x), x + x^2 \right].$$

In this case $F'_x(0, 0) = \frac{d^2}{dx^2} + 1$ and $F'_\delta(0, 0) = 0$. The Fréchet derivative $F'(x, \delta)$ is singular at the point $(x^*, \delta^*) = (0, 0)$. The boundary value problem $x'' + x = \sin t$, $x(0) = x(\pi) = 0$ has no solution and $F'_x(0, 0)$ is not surjective. Therefore, neither the classical Newton method, the Poincaré small-parameter method, nor other numerical methods based on the nonsingularity of $F'(0, 0)$ are applicable. However, the 2-factor operator $F'(0, 0) + PF''(0, 0)h$ is not singular in the direction of a vector $h \neq 0$, i.e. the mapping F is 2-regular along h and the 2-factor method described in Section 7 can therefore be applied for finding a solution $x(\delta)$ of (48) which depends on a small parameter $\delta > 0$.

11 Concluding remarks

This paper presented alternative proofs of several classical results in nonlinear optimization and established new connections between optimality conditions, Farkas' lemma, cone closure, and p-regularity theory. These developments provide a unified perspective on several fundamental topics in optimization and suggest new possibilities for the analysis of degenerate optimization problems and the development of Newton-type methods.

Declarations

Funding

Open access funding provided by University of Siedlce. Not supported by any fund.

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Data availability

Not applicable.

Materials availability

All of the material is owned by the authors and no permissions are required.

Code availability

Not applicable.

Author contribution

All authors contributed equally to this work. All authors reviewed the manuscript.

References

- [1] Anitescu M (2002) A superlinearly convergent sequential quadratically constrained quadratic programming algorithm for degenerate nonlinear programming. SIAM J. Optim., 12(4).

- [2] Barvinok A, Rudelson M (2022) When a system of real quadratic equations has a solution. *Adv. Math.*, 403.
- [3] Belash K, Tret'yakov A (1988) Methods for solving degenerate problems. *USSR Computational Mathematics and Mathematical Physics*, Elsevier, 28(4:90-94.
- [4] Bertsekas D (2016) *Nonlinear Programming* 3rd Edition. (Athena Scientific, Belmont, Massachusetts).
- [5] Brezhneva OA, Tret'yakov A (2007) The p-factor-Lagrange methods for degenerate nonlinear programming. *Numerical Functional Analysis and Optimization*, 28(9-10):1051-1086.
- [6] De Marchi A (2022) On a primal-dual newton proximal method for convex quadratic programs. *Computational Optimization and Applications*, 81.
- [7] Denisov D, Evtushenko Y, Tret'yakov A (2021) Some properties of smooth convex functions and Newton's method. *Doklady Mathematics*, 103:76–80.
- [8] Dostal Z, Brzobohaty T, Horak D, Kozubek T, Vodstrcil T (2014) On r-linear convergence of semi-monotonic inexact augmentedlagrangians for bound and equality constrained quadratic programming problems with application. *Comput. Math. Appl.*, 67.
- [9] Y. Evtushenko Y (1985) *Numerical optimization techniques*. Translations Series in Mathematics and Engineering, (Springer-Verlag Berlin, Heidelberg).
- [10] Evtushenko Y, Malkova V, Tret'yakov A (2020) P-Regularity Theory: Applications to Optimization. In: Olenov, N., Evtushenko, Y., Khachay, M., Malkova, V. (eds) *Optimization and Applications. OPTIMA 2020. Lecture Notes in Computer Science*, 12422: 97–109, 416 (Springer, Cham).
- [11] Evtushenko Y, Tret'yakov AA (2013) pth-order approximation of the solution set of nonlinear equations. *Computational Mathematics and Mathematical Physics* (Pleiades Publishing, Ltd.), 53(12).
- [12] Evtushenko Y, Tret'yakov AA (2020) A New View of some Fundamental Results in Optimization. *Computational Mathematics and Mathematical Physics*, 60:1412–1421.
- [13] Facchinei F, Fischer A, Kanzow C (1998) On the accurate identification of active constraints. *SIAM J. Optim.*, 9(1):14–32.
- [14] Fletcher R (1993) Resolving degeneracy in quadratic programming. *Ann. Oper. Res.*, 46.
- [15] Garey M, Johnson D (1979) *Computers and Intractability: A Guide to the Theory of NP-Completeness*. (Freeman, San Francisco).

- [16] Gould N, Orban D, Robinson D (2013) Trajectory-following methods for large-scale degenerate convex quadratic programming. *Math. Program. Comput.*, 5.
- [17] Ioffe AD, Tikhomirov VM (1979) Theory of extremal problems. *Studies in Mathematics and its Applications*, vol 6 (North-Holland publishing Company, Amsterdam New York Oxford).
- [18] Li N, Zhi L (2022) Improved two-step newton's method for computing simple multiple zeros of polynomial systems. *Numer. Algorithms*, 91.
- [19] Mertens S (2005) *Computational Complexity and Statistical Physics*, chapter The Easiest Hard Problem: Number Partitioning. (Oxford University Press).
- [20] Nocedal J, Wright S (2006) *Numerical optimization*, 2nd edition (Springer. New York).
- [21] Permenter F (2023) Log-domain interior-point methods for convex quadratic programming. *Optimization Letters*, 17(7).
- [22] Poirier A, Torres J (2023) Approximating roots by quadratic iteration. *Proyecciones J. Math.*, 42.
- [23] Polyak B (1987) *Introduction to Optimization*. Translations Series in Mathematics and Engineering (Optimization Software, Inc., Publications Division, New York).
- [24] Tret'yakov A, Marsden J (2003) Factor-analysis of nonlinear mappings: p-regularity theory. *Commun. Pure Appl. Anal.*, 2.
- [25] Yamakawa Y, Takayuki O (2022) A stabilized sequential quadratic semidefinite programming method for degenerate nonlinear semidefinite programs. *Comput. Optim. Appl.*, 83.