Appendix A
Miscellaneous Material

A.1 Linear Algebra and Linear Inequalities

The following statement can be regarded as a variant of the celebrated Farkas Lemma (e.g., [27, Theorem 2.201]).

Lemma A.1. *For any* $A \in \mathbf{R}^{l \times n}$ *and* $B \in \mathbf{R}^{m \times n}$*, for the cone*

$$C = \{x \in \mathbf{R}^n \mid Ax = 0,\, Bx \leq 0\}$$

it holds that

$$C^\circ = \{x \in \mathbf{R}^n \mid x = A^{\mathrm{T}}y + B^{\mathrm{T}}z,\, y \in \mathbf{R}^l,\, z \in \mathbf{R}^m_+\}.$$

Lemma A.1 can be derived as a corollary of the Motzkin Theorem of the Alternatives [186, p. 28], stated next.

Lemma A.2. *For any* $A \in \mathbf{R}^{l \times n}$*,* $B \in \mathbf{R}^{m \times n}$*,* $B_0 \in \mathbf{R}^{m_0 \times n}$*, one and only one of the following statements holds: either there exists* $x \in \mathbf{R}^n$ *such that*

$$Ax = 0, \quad Bx \leq 0, \quad B_0 x < 0,$$

or there exists $(y,\, z,\, z^0) \in \mathbf{R}^l \times \mathbf{R}^m \times \mathbf{R}^{m_0}$ *such that*

$$A^{\mathrm{T}}y + B^{\mathrm{T}}z + B_0^{\mathrm{T}}z^0 = 0, \quad z \geq 0, \quad z^0 \geq 0, \quad z^0 \neq 0.$$

The following simplified version of Lemma A.2, convenient in some applications, is known as the Gordan Theorem of the Alternatives.

Lemma A.3. *For any* $B \in \mathbf{R}^{m_0 \times n}$*, one and only one of the following two alternatives is valid: either there exists* $x \in \mathbf{R}^n$ *such that*

$$B_0 x < 0,$$

or there exists $z^0 \in \mathbf{R}^m$ *such that*

$$B_0^{\mathrm{T}}z^0 = 0, \quad z^0 \geq 0, \quad z^0 \neq 0.$$

A.F. Izmailov and M.V. Solodov, *Newton-Type Methods for Optimization and Variational Problems*, Springer Series in Operations Research and Financial Engineering, DOI 10.1007/978-3-319-04247-3,

The following is Hoffman's lemma giving a (global) error bound for linear systems (e.g., [27, Theorem 2.200]).

Lemma A.4. *For any* $A \in \mathbf{R}^{l\times n}$, $a \in \mathbf{R}^l$, *and* $B \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, *assume that the set*

$$S = \{x \in \mathbf{R}^n \mid Ax = a,\, Bx \leq b\}$$

is nonempty.

Then there exists $c > 0$ *such that*

$$\mathrm{dist}(x,\, S) \leq c(\|Ax - a\| + \|\max\{0,\, Bx - b\}\|) \quad \forall\, x \in \mathbf{R}^n.$$

The following is the finite-dimensional version of the classical Banach Open Mapping Theorem.

Lemma A.5. *For any* $A \in \mathbf{R}^{l\times n}$, *if* $\mathrm{rank}\, A = l$, *then there exists* $c > 0$ *such that for any* $B \in \mathbf{R}^{l\times n}$ *close enough to* A, *and any* $y \in \mathbf{R}^l$, *the equation*

$$Bx = y$$

has a solution $x(y)$ *such that*

$$\|x(y)\| \leq c\|y\|.$$

This result is complemented by the more exact characterization of invertibility of small perturbations of a nonsingular matrix; see, e.g., [103, Theorem 2.3.4].

Lemma A.6. *Let* $A \in \mathbf{R}^{n\times n}$ *be a nonsingular matrix.*

Then any matrix $B \in \mathbf{R}^{n\times n}$ *satisfying the inequality* $\|B - A\| < 1/\|A^{-1}\|$ *is nonsingular, and*

$$\|B^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2\|B - A\|}{1 - \|A^{-1}\|\|B - A\|}.$$

Lemma A.7 below is well known; it is sometimes called the Finsler Lemma [81], or the Debreu Lemma [54]. The similar in spirit Lemma A.8, on the other hand, is not standard, so we have to give its proof (from [150]). As the proofs of Lemmas A.7 and A.8 are somehow related, it makes sense to provide both.

Lemma A.7. *Let* $H \in \mathbf{R}^{n\times n}$ *be any symmetric matrix and* $A \in \mathbf{R}^{l\times n}$ *any matrix such that*

$$\langle H\xi,\, \xi\rangle > 0 \quad \forall\, \xi \in \ker A \setminus \{0\}. \tag{A.1}$$

Then the matrix $H + cA^{\mathrm{T}}A$ *is positive definite for all* $c \geq 0$ *large enough.*

Proof. We argue by contradiction. Suppose that there exist $\{c_k\} \subset \mathbf{R}$ and $\{\xi^k\} \subset \mathbf{R}^n$ such that $c_k \to +\infty$ as $k \to \infty$, and for all k it holds that $\|\xi^k\| = 1$ and

$$\langle (H + c_k A^{\mathrm{T}} A)\xi^k, \, \xi^k \rangle \leq 0. \tag{A.2}$$

Without loss of generality, we may assume that $\{\xi^k\} \to \xi$, with some $\xi \in \mathbf{R}^n \setminus \{0\}$. Dividing (A.2) by c_k and passing onto the limit as $k \to \infty$, we obtain that

$$0 \geq \langle A^{\mathrm{T}} A\xi, \, \xi \rangle = \|A\xi\|^2,$$

i.e., $\xi \in \ker A$.

On the other hand, since for each k it holds that

$$\langle A^{\mathrm{T}} A\xi^k, \, \xi^k \rangle = \|A\xi^k\|^2 \geq 0,$$

the inequality (A.2) implies that $\langle H\xi^k, \, \xi^k \rangle \leq 0$. Passing onto the limit as $k \to \infty$, we obtain that $\langle H\xi, \, \xi \rangle \leq 0$, in contradiction with (A.1). □

Another result, as already commented somewhat similar in nature to the Debreu–Finsler Lemma, is stated next.

Lemma A.8. *Let $H \in \mathbf{R}^{n \times n}$ and $A \in \mathbf{R}^{l \times n}$ be such that*

$$H\xi \notin \operatorname{im} A^{\mathrm{T}} \quad \forall\, \xi \in \ker A \setminus \{0\}. \tag{A.3}$$

Then for any $C > 0$, any $\tilde{H} \in \mathbf{R}^{n \times n}$ close enough to H, and any $\tilde{A} \in \mathbf{R}^{l \times n}$ close enough to A, the matrix $\tilde{H} + c(A + \Omega)^{\mathrm{T}} \tilde{A}$ is nonsingular for all $c \in \mathbf{R}$ such that $|c|$ is large enough, and for all $\Omega \in \mathbf{R}^{l \times n}$ satisfying $\|\Omega\| \leq C/|c|$.

Proof. Suppose the contrary, i.e., that there exist sequences $\{H_k\} \subset \mathbf{R}^{n \times n}$, $\{A_k\} \subset \mathbf{R}^{l \times n}$, $\{\Omega_k\} \subset \mathbf{R}^{l \times n}$, $\{c_k\} \subset \mathbf{R}$ and $\{\xi^k\} \subset \mathbf{R}^n \setminus \{0\}$, such that $\{H_k\} \to H$, $\{A_k\} \to A$, $|c_k| \to \infty$ as $k \to \infty$, and for all k it holds that $\|\Omega_k\| \leq C/|c_k|$ and

$$H_k \xi^k + c_k (A + \Omega_k)^{\mathrm{T}} A_k \xi^k = 0. \tag{A.4}$$

We can assume, without loss of generality, that $\|\xi^k\| = 1$ for all k, and $\{\xi^k\} \to \xi$, with some $\xi \in \mathbf{R}^n \setminus \{0\}$. Then since the right-hand side in

$$A^{\mathrm{T}} A_k \xi^k = -\frac{1}{c_k} H_k \xi^k - \Omega_k^{\mathrm{T}} A_k \xi^k,$$

tends to zero as $k \to \infty$, it must hold that $A^{\mathrm{T}} A\xi = 0$.

It is thus established that $A\xi \in \ker A^{\mathrm{T}}$, and since $A\xi \in \operatorname{im} A = (\ker A^{\mathrm{T}})^{\perp}$, this shows that $A\xi = 0$. Thus, $\xi \in \ker A \setminus \{0\}$.

On the other hand, (A.4) implies that the inclusion

$$H_k \xi^k + c_k \Omega_k^{\mathrm{T}} A_k \xi^k = -c_k A^{\mathrm{T}} A_k \xi^k \in \operatorname{im} A^{\mathrm{T}}$$

holds for all k, where the second term in the left-hand side tends to zero as $k \to \infty$ because $\{c_k \Omega_k\}$ is bounded and $\{A_k \xi^k\} \to A\xi = 0$. Hence, $H\xi \in \operatorname{im} A^{\mathrm{T}}$ by the closedness of $\operatorname{im} A^{\mathrm{T}}$. This gives a contradiction with (A.3). □

We complete this section by the following fact concerned with the existence of the inverse of a block matrix; see [243, Proposition 3.9].

Lemma A.9. *If $A \in \mathbf{R}^{n\times n}$ is a nonsingular matrix, $B \in \mathbf{R}^{n\times m}$, $C \in \mathbf{R}^{m\times n}$, $D \in \mathbf{R}^{m\times m}$, then for the matrix*

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

it holds that

$$\det M = \det A \det(D - CA^{-1}B).$$

Under the assumptions of Lemma A.9, the matrix $D - CA^{-1}B$ is referred to as the *Schur complement* of A in M.

A.2 Analysis

Our use of the big-O and little-o notation employs the following conventions. For a mapping $F : \mathbf{R}^n \to \mathbf{R}^m$ and a function $\varphi : \mathbf{R}^n \to \mathbf{R}_+$, and for a given $\bar{x} \in \mathbf{R}^n$, we write $F(x) = O(\varphi(x))$ as $x \to \bar{x}$ if there exists $c > 0$ such that $\|F(x)\| \le c\varphi(x)$ for all $x \in \mathbf{R}^n$ close enough to $\bar{x}$. We write $F(x) = o(\varphi(x))$ as $x \to \bar{x}$ if for every $\varepsilon > 0$ no matter how small, it holds that $\|F(x)\| \le \varepsilon\varphi(x)$ for all $x \in \mathbf{R}^n$ close enough to $\bar{x}$. For sequences $\{x^k\} \subset \mathbf{R}^n$ and $\{t_k\} \subset \mathbf{R}_+$, by $x^k = O(t_k)$ as $k \to \infty$ we mean that there exists $c > 0$ such that $\|x^k\| \le ct_k$ for all k large enough. Accordingly, $x^k = o(t_k)$ as $k \to \infty$ if for every $\varepsilon > 0$ no matter how small, it holds that $\|x^k\| \le \varepsilon t_k$ for all k large enough. For a sequence $\{\tau_k\} \subset \mathbf{R}$, we write $\tau^k \le o(t_k)$ as $k \to \infty$ if for any $\varepsilon > 0$ no matter how small it holds that $\tau_k \le \varepsilon t_k$ for all k large enough.

Concerning convergence rate estimates, the terminology is as follows. Let a sequence $\{x^k\} \subset \mathbf{R}^n$ be convergent to some $\bar{x} \in \mathbf{R}^n$. If there exist $q \in (0,\, 1)$ and $c > 0$ such that

$$\|x^k - \bar{x}\| \le cq^k$$

for all k large enough (or, in other words, $\|x^k - \bar{x}\| = O(q^k)$ as $k \to \infty$), then we say that $\{x^k\}$ has *geometric convergence rate*. If there exists $q \in (0,\, 1)$ such that

$$\|x^{k+1} - \bar{x}\| \le q\|x^k - \bar{x}\| \tag{A.5}$$

for all k large enough, then we say that $\{x^k\}$ has *linear convergence rate*. Linear rate implies geometric rate, but the converse is not true. If for every $q \in (0,\, 1)$, no matter how small, the inequality (A.5) holds for all k large enough, then we say that $\{x^k\}$ has *superlinear convergence rate*. To put it in other words, superlinear convergence means that

$$\|x^{k+1} - \bar{x}\| = o(\|x^k - \bar{x}\|)$$

as $k \to \infty$. A particular case of the superlinear rate is *quadratic convergence rate*, meaning that there exists $c > 0$ such that

$$\|x^{k+1} - \bar{x}\| \leq c\|x^k - \bar{x}\|^2$$

for all k large enough or, in other words,

$$\|x^{k+1} - \bar{x}\| = O(\|x^k - \bar{x}\|^2)$$

as $k \to \infty$. Unlike superlinear or quadratic rate, linear convergence rate depends on the norm: linear convergence rate in some norm in $\mathbf{R}^n$ does not necessarily imply linear convergence rate in a different norm.

We next state some facts and notions of differential calculus for mappings (generally vector-valued and with vector variable). It is assumed that the reader is familiar with differential calculus for scalar-valued functions in a scalar variable.

The mapping $F : \mathbf{R}^n \to \mathbf{R}^m$ is said to be *differentiable at* $x \in \mathbf{R}^n$ if there exists a matrix $J \in \mathbf{R}^{m\times n}$ such that for $\xi \in \mathbf{R}^n$ is holds that

$$F(x + \xi) = F(x) + J\xi + o(\|\xi\|)$$

as $\xi \to 0$. The matrix J with this property is necessarily unique; it coincides with the *Jacobian* $F'(x)$ (the matrix of first partial derivatives of the components of F at x with respect to all the variables), and it is also called the *first derivative of* F *at* x. The rows of the Jacobian are the *gradients* $F_1'(x), \ldots, F_m'(x)$ (vectors of first partial derivatives with respect to all variables) of the components of F at x.

The mapping $F : \mathbf{R}^n \to \mathbf{R}^m$ is (*continuously*) *differentiable on a set* $S \subset \mathbf{R}^n$ if it is differentiable at every point of some open set $O \subset \mathbf{R}^n$ such that $S \subset O$ (and the mapping $F'(\cdot)$ defined on O is continuous at every point of S).

The mapping $F : \mathbf{R}^n \to \mathbf{R}^m$ is said to be *twice differentiable at* $x \in \mathbf{R}^n$ if it is differentiable in a neighborhood of x, and the mapping $F'(\cdot)$ defined on this neighborhood is differentiable at x. The derivative $(F')'(x)$ of $F'(\cdot)$ at x can be regarded as a linear operator from $\mathbf{R}^n$ to $\mathbf{R}^{m\times n}$, or alternatively, as a bilinear mapping $F''(x) : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^m$ defined by

$$F''(x)[\xi^1, \xi^2] = ((F')'(x)\xi^1)\xi^2, \quad \xi^1, \xi^2 \in \mathbf{R}^n.$$

This bilinear mapping is necessarily symmetric, that is,

$$F''(x)[\xi^1, \xi^2] = F''(x)[\xi^2, \xi^1] \quad \forall\, \xi^1, \xi^2 \in \mathbf{R}^n.$$

The mapping in question is called the *second derivative of* F *at* x, and it is comprised by the *Hessians* $F_1''(x), \ldots, F_m''(x)$ (the matrices of second partial derivatives) of the components of F at x:

$$F''(x)[\xi^1, \xi^2] = (\langle F_1''(x)\xi^1, \xi^2\rangle, \ldots, \langle F_m''(x)\xi^1, \xi^2\rangle) \quad \forall\, \xi^1, \xi^2 \in \mathbf{R}^n.$$

Note that the symmetry of the bilinear mapping $F''(x)$ is equivalent to the symmetry of the Hessians of the components of F.

If F is twice differentiable at $\bar{x}$, then for $\xi \in \mathbf{R}^n$ it holds that

$$F(x+\xi) = F(x) + F'(x)\xi + \frac{1}{2}F'(x)[\xi, \xi] + o(\|\xi\|^2)$$

as $\xi \to 0$. This fact can be regarded as a particular case of the Taylor formula.

The mapping F is *twice* (*continuously*) *differentiable on a set* $S \subset \mathbf{R}^n$ if it is twice differentiable at every point of some open set $O \subset \mathbf{R}^n$ such that $S \subset O$ (and the mapping $F''(\cdot)$ defined on the set O is continuous at every point of S).

Furthermore, the mapping $F : \mathbf{R}^n \times \mathbf{R}^l \to \mathbf{R}^m$ is said to be *differentiable at* $(x, y) \in \mathbf{R}^n \times \mathbf{R}^l$ *with respect to* x if the mapping $F(\cdot, y)$ is differentiable at x. The derivative of the latter mapping at x is called the *partial derivative of* F *with respect to* x *at* (x, y), and it is denoted by $\frac{\partial F}{\partial x}(x, y)$.

Similarly, the mapping $F : \mathbf{R}^n \times \mathbf{R}^l \to \mathbf{R}^m$ is said to be *twice differentiable at* $(x, y) \in \mathbf{R}^n \times \mathbf{R}^l$ *with respect to* x if the mapping $F(\cdot, y)$ is twice differentiable at x. The second derivative of the latter mapping at x is called the *second partial derivative of* F *with respect to* x *at* (x, y), and it is denoted by $\frac{\partial^2 F}{\partial x^2}(x, y)$.

In this book, any of the itemized assertions in the next statement is referred to as a mean-value theorem. The first part of item (a) is a rather subtle and not widely known result; it was established in [198]. The other statements are fairly standard.

Theorem A.10. *For any* $F : \mathbf{R}^n \to \mathbf{R}^m$ *and any* $x^1, x^2 \in \mathbf{R}^n$, *the following assertions are valid:*

(a) If F *is continuous on* $[x^1, x^2] = \{tx^1 + (1-t)x^2 \mid t \in [0, 1]\}$ *and differentiable on* $(x^1, x^2) = \{tx^1 + (1-t)x^2 \mid t \in (0, 1)\}$, *then there exist* $t_i \in (0, 1)$ *and* $\theta_i \geq 0$, $i = 1, \ldots, m$, *such that* $\sum_{i=1}^m \theta_i = 1$ *and*

$$F(x^1) - F(x^2) = \sum_{i=1}^{m} \theta_i F'(t_i x^1 + (1-t_i)x^2)(x^1 - x^2),$$

and in particular,

$$\|F(x^1) - F(x^2)\| \leq \sup_{t \in (0,\, 1)} \|F'(tx^1 + (1-t)x^2)\| \|x^1 - x^2\|.$$

(b) If F is continuously differentiable on the line segment $[x^1, x^2]$, then

$$F(x^1) - F(x^2) = \int_0^1 F'(tx^1 + (1-t)x^2)(x^1 - x^2)\, dt.$$

The next fact is an immediate corollary of assertion (b) of Theorem A.10.

Lemma A.11. *For any $F : \mathbf{R}^n \to \mathbf{R}^m$ and any $x^1, x^2 \in \mathbf{R}^n$, if F is differentiable on the line segment $[x^1, x^2]$, with its derivative being Lipschitz-continuous on this segment with a constant $L > 0$, then*

$$\|F(x^1) - F(x^2) - F'(x^2)(x^1 - x^2)\| \leq \frac{L}{2}\|x^1 - x^2\|^2.$$

A.3 Convexity and Monotonicity

For a detailed exposition of finite-dimensional convex analysis, we refer to [235]. In this section we only recall some basic definitions and facts used in this book. For details on (maximal) monotone mappings and related issues, we refer to [17, 30] and [239, Chap. 12].

For a finite number of points $x^1, \ldots, x^m \in \mathbf{R}^n$, their *convex combinations* are points of the form $\sum_{i=1}^m t_i x^i$ with some $t_i \geq 0$, $i = 1, \ldots, m$, such that $\sum_{i=1}^m t_i = 1$. In particular, convex combinations of two points $x^1, x^2 \in \mathbf{R}^n$ are points of the form $tx^1 + (1-t)x^2$, $t \in [0, 1]$, and they form a line segment connecting x^1 and x^2.

A set $S \subset \mathbf{R}^n$ is said to be *convex* if for each pair of points $x^1, x^2 \in S$ all convex combinations of these points belong to S (equivalently, for any points $x^1, \ldots, x^m \in S$, where $m \geq 2$, all convex combinations of these points belong to S).

The *convex hull* of a set $S \subset \mathbf{R}^n$, denoted by $\operatorname{conv} S$, is the smallest convex set in $\mathbf{R}^n$ that contains S (equivalently, the set of all convex combinations of points in S).

By a (Euclidean) *projection* of a point $x \in \mathbf{R}^n$ onto a given set $S \subset \mathbf{R}^n$ we mean a point closest to x among all the points in S, i.e., any global solution of the optimization problem

$$\begin{array}{ll} \text{minimize} & \|y - x\| \\ \text{subject to} & y \in S. \end{array} \tag{A.6}$$

As the objective function in (A.6) is coercive, projection of any point onto any nonempty closed set in $\mathbf{R}^n$ exists. If, in addition, the set is convex, then the following holds.

Lemma A.12. *Let $S \subset \mathbf{R}^n$ be any nonempty closed convex set.*

Then the projection operator onto S, $\pi_S : \mathbf{R}^n \to S$, is well defined and single valued: for any point $x \in \mathbf{R}^n$ its projection $\pi_S(x)$ onto S exists and is unique. Moreover, $\bar{x} = \pi_S(x)$ if and only if

$$\bar{x} \in S, \quad \langle x - \bar{x},\, y - \bar{x} \rangle \leq 0 \quad \forall\, y \in S.$$

In addition, the projection operator is nonexpansive:

$$\|\pi_S(x^1) - \pi_S(x^2)\| \leq \|x^1 - x^2\| \quad \forall\, x^1,\, x^2 \in \mathbf{R}^n.$$

A set $C \subset \mathbf{R}^n$ is called a *cone* if for each $x \in C$ it contains all points of the form tx, $t \geq 0$. The *polar cone* to C is defined by

$$C^\circ = \{\xi \in \mathbf{R}^n \mid \langle \xi,\, x \rangle \leq 0\; \forall\, x \in C\}.$$

Lemma A.13. *For any nonempty closed convex cone $C \subset \mathbf{R}^n$ it holds that*

$$x = \pi_C(x) + \pi_{C^\circ}(x) \quad \forall\, x \in \mathbf{R}^n,$$

and in particular,

$$C^\circ = \{x \in \mathbf{R}^n \mid \pi_C(x) = 0\},$$

$$\pi_C(x - \pi_C(x)) = 0 \quad \forall\, x \in \mathbf{R}^n.$$

An important property concerns separation of (convex) sets by hyperplanes. The following separation theorem can be found in [235, Corollary 11.4.2].

Theorem A.14. *Let $S_1,\, S_2 \subset \mathbf{R}^n$ be nonempty closed convex sets, with at least one of them being also bounded (hence, compact).*

Then $S_1 \cap S_2 = \emptyset$ if and only if there exist $\xi \in \mathbf{R}^n \setminus \{0\}$ and $t \in \mathbf{R}$ such that

$$\langle \xi,\, x^1 \rangle < t < \langle \xi,\, x^2 \rangle \quad \forall\, x^1 \in S_1,\, x^2 \in S_2.$$

Given a convex set $S \subset \mathbf{R}^n$, a function $f : S \to \mathbf{R}$ is said to be *convex* (on the set S) if

$$f(tx^1 + (1-t)x^2) \leq tf(x^1) + (1-t)f(x^2) \quad \forall\, x^1,\, x^2 \in S,\; \forall\, t \in [0,\, 1].$$

Equivalently, f is convex if its *epigraph* $\{(x,\, t) \in S \times \mathbf{R} \mid f(x) \leq t\}$ is a convex set. It is immediate that a linear combination of convex functions with nonnegative coefficients is a convex function, and the maximum over a finite family of convex functions is a convex function.

Furthermore, f is said to be *strongly convex* (on S) if there exists $\gamma > 0$ such that

$$f(tx^1 + (1-t)x^2) \leq tf(x^1) + (1-t)f(x^2) - \gamma t(1-t)\|x^1 - x^2\|^2$$
$$\forall\, x^1,\, x^2 \in S,\; \forall\, t \in [0,\, 1].$$

The sum of a convex function and a strongly convex function is evidently strongly convex. The following are characterizations of convexity for smooth functions.

Proposition A.15. *Let $O \subset \mathbf{R}^n$ be a nonempty open convex set, and let $f : O \to \mathbf{R}$ be differentiable in O.*

Then the following items are equivalent:

(a) The function f is convex on O.
(b) $f(x^1) \geq f(x^2) + \langle f'(x^2),\, x^1 - x^2\rangle$ for all $x^1, x^2 \in O$.
(c) $\langle f'(x^1) - f'(x^2),\, x^1 - x^2\rangle \geq 0$ for all $x^1, x^2 \in O$.

If f is twice differentiable in O, then the properties above are further equivalent to

(d) The Hessian $f''(x)$ is positive semidefinite for all $x \in O$.

It is clear that a quadratic function is convex if and only if its (constant) Hessian is a positive semidefinite matrix (see item (d) in Proposition A.15). Moreover, a quadratic function is strongly convex if and only if its Hessian is positive definite.

Given a convex function $f : \mathbf{R}^n \to \mathbf{R}$, an element $a \in \mathbf{R}^n$ is called a *subgradient* of f at a point $x \in \mathbf{R}^n$ if

$$f(y) \geq f(x) + \langle a,\, y - x\rangle \quad \forall\, y \in \mathbf{R}^n.$$

The set of all the elements $a \in \mathbf{R}^n$ with this property is called the *subdifferential* of f at x, denoted by $\partial f(x)$.

Proposition A.16. *Let $f : \mathbf{R}^n \to \mathbf{R}$ be convex on $\mathbf{R}^n$.*

Then for each $x \in \mathbf{R}^n$ the subdifferential $\partial f(x)$ is a nonempty compact convex set. Moreover, f is continuous and directionally differentiable at every $x \in \mathbf{R}^n$ in every direction $\xi \in \mathbf{R}^n$, and it holds that

$$f'(x;\, \xi) = \max_{y \in \partial f(x)} \langle y,\, \xi\rangle.$$

For some further calculus rules for subdifferentials see Sect. 1.4.1, where they are presented in a more general (not necessarily convex) setting. Section 1.4.1 provides all the necessary material for the convex calculus in this book.

We complete this section by some definitions and facts concerned with the notion of monotonicity for (multi)functions. For a (generally) set-valued mapping Ψ from $\mathbf{R}^n$ to the subsets of $\mathbf{R}^n$, define its domain

$$\operatorname{dom} \Psi = \{x \in \mathbf{R}^n \mid \Psi(x) \neq \emptyset\}.$$

Then Ψ is said to be *monotone* if

$$\langle y^1 - y^2,\, x^1 - x^2\rangle \geq 0 \quad \forall\, y^1 \in \Psi(x^1),\ \forall\, y^2 \in \Psi(x^2),\ \forall\, x^1,\, x^2 \in \operatorname{dom} \Psi,$$

and *maximal monotone* if, in addition, its *graph* $\{(x, y) \in \mathbf{R}^n \times \mathbf{R}^n \mid y \in \Psi(x)\}$ is not contained in the graph of any other monotone set-valued mapping.

Some examples of maximal monotone mappings are: a continuous monotone function $F : \mathbf{R} \to \mathbf{R}$, the subdifferential multifunction $\partial f(\cdot)$ of a convex function $f : \mathbf{R}^n \to \mathbf{R}$, and the normal cone multifunction $N_S(\cdot)$ for a closed convex set $S \subset \mathbf{R}^n$.

The sum of two monotone mappings is monotone, and the sum of two maximal monotone mappings is maximal monotone if the domain of one intersects the interior of the domain of the other.

Furthermore, Ψ is said to be *strongly monotone* if there exists $\gamma > 0$ such that $\Psi - \gamma I$ is monotone, which is equivalent to the property

$$\langle y^1 - y^2,\, x^1 - x^2\rangle \geq \gamma \|x^1 - x^2\|^2 \quad \forall\, y^1 \in \Psi(x^1),\ \forall\, y^2 \in \Psi(x^2),\ \forall\, x^1,\, x^2 \in \operatorname{dom} \Psi.$$

In particular, the identity mapping I is strongly monotone, and the sum of a monotone mapping and a strongly monotone mapping is strongly monotone.

The following characterization of monotonicity for smooth mappings can be found, e.g., in [239, Proposition 12.3].

Proposition A.17. *Let $F : \mathbf{R}^n \to \mathbf{R}^n$ be differentiable on $\mathbf{R}^n$.*

Then F is monotone if and only if $F'(x)$ is positive semidefinite for all $x \in \mathbf{R}^n$.

We refer to [17, 30] and [239, Chap. 12] for other details on (maximal) monotone mappings and related issues.

References

1. W. Achtziger, C. Kanzow, Mathematical programs with vanishing constraints: optimality conditions and constraint qualifications. Math. Program. **114**, 69–99 (2007)
2. W. Achtziger, T. Hoheisel, C. Kanzow, A smoothing-regularization approach to mathematical programs with vanishing constraints. Comput. Optim. Appl. **55**, 733–767 (2013)
3. L.T.H. An, An efficient algorithm for globally minimizing a quadratic function under convex quadratic constraints. Math. Program. **87**, 401–426 (2000)
4. R. Andreani, E.G. Birgin, J.M. Martínez, M.L. Schuverdt, On Augmented Lagrangian methods with general lower-level constraints. SIAM J. Optim. **18**, 1286–1309 (2007)
5. R. Andreani, E.G. Birgin, J.M. Martínez, M.L. Schuverdt, Augmented Lagrangian methods under the constant positive linear dependence constraint qualification. Math. Program. **111**, 5–32 (2008)
6. M. Anitescu, Degenerate nonlinear programming with a quadratic growth condition. SIAM J. Optim. **10**, 1116–1135 (2000)
7. M. Anitescu, *Nonlinear Programs with Unbounded Lagrange Multiplier Sets* (Mathematics and Computer Science Division, Argonne National Laboratories, Argonne, 2000). Preprint ANL/MCS-P796-0200
8. M. Anitescu, *On solving mathematical programs with complementarity constraints as nonlinear programs* (Mathematics and Computer Science Division, Argonne National Laboratories, Argonne, 2000). Preprint ANL/MCS-P864-1200
9. M. Anitescu, A superlinearly convergent sequential quadratically constrained quadratic programming algorithm for degenerate nonlinear programming. SIAM J. Optim. **12**, 949–978 (2002)
10. M. Anitescu, On the rate of convergence of sequential quadratic programming with nondifferentiable exact penalty function in the presence of constrain degeneracy. Math. Program. **92**, 359–386 (2002)

A.F. Izmailov and M.V. Solodov, *Newton-Type Methods for Optimization and Variational Problems*, Springer Series in Operations Research and Financial Engineering, DOI 10.1007/978-3-319-04247-3,

11. M. Anitescu, On using the elastic mode in nonlinear programming approaches to mathematical programs with complementarity constraints. SIAM J. Optim. **15**, 1203–1236 (2005)
12. M. Anitescu, P. Tseng, S.J. Wright, Elastic-mode algorithms for mathematical programs with equilibrium constraints: global convergence and stationarity properties. Math. Program. **110**, 337–371 (2007)
13. A.V. Arutyunov, *Optimality Conditions: Abnormal and Degenerate Problems* (Kluwer Academic, Dordrecht, 2000)
14. A.V. Arutyunov, A.F. Izmailov, Sensitivity analysis for cone-constrained optimization problems under the relaxed constraint qualifications. Math. Oper. Res. **30**, 333–353 (2005)
15. C. Audet, P. Hansen, B. Jaumard, G. Savard, A branch and cut algorithm for nonconvex quadratically constrained quadratic programming. Math. Program. **87**, 131–152 (2000)
16. A.B. Bakushinskii, A regularization algorithm based on the Newton-Kantorovich method for solving variational inequalities. USSR Comput. Math. Math. Phys. **16**, 16–23 (1976)
17. H.H. Bauschke, P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics (Springer, New York, 2011)
18. D.P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods* (Academic Press, New York, 1982)
19. D.P. Bertsekas, *Nonlinear Programming*, 2nd edn. (Athena, Belmont, 1999)
20. S.C. Billups, Algorithms for complementarity problems and generalized equations. Ph.D. thesis. Technical Report 95-14. Computer Sciences Department, University of Wisconsin, Madison, 1995
21. E.G. Birgin, J.M. Martínez, Local convergence of an inexact-restoration method and numerical experiments. J. Optim. Theory Appl. **127**, 229–247 (2005)
22. E.G. Birgin, J.M. Martínez, Improving ultimate convergence of an augmented Lagrangian method. Optim. Method. Softw. **23**, 177–195 (2008)
23. P. Boggs, J. Tolle, Sequential quadratic programming. Acta Numer. **4**, 1–51 (1995)
24. B.T. Boggs, J.W. Tolle, A.J. Kearsley, A truncated SQP algorithm for large scale nonlinear programming problems, in *Advances in Optimization and Numerical Analysis*, ed. by S. Gomez, J.-P. Hennart. Math. Appl. vol 275 (Kluwer Academic, Dordrecht, 1994), pp. 69–77
25. J.F. Bonnans, Asymptotic admissibility of the unit stepsize in exact penalty methods. SIAM J. Control Optim. **27**, 631–641 (1989)
26. J.F. Bonnans, Local analysis of Newton-type methods for variational inequalities and nonlinear programming. Appl. Math. Optim. **29**, 161–186 (1994)
27. J.F. Bonnans, A. Shapiro, *Perturbation Analysis of Optimization Problems* (Springer, New York, 2000)

28. J.F. Bonnans, A. Sulem, Pseudopower expansion of solutions of generalized equations and constrained optimization. Math. Program. **70**, 123–148 (1995)
29. J.F. Bonnans, J.Ch. Gilbert, C. Lemaréchal, C. Sagastizábal, *Numerical Optimization: Theoretical and Practical Aspects*, 2nd edn. (Springer, Berlin, 2006)
30. R.S. Burachik, A.N. Iusem, *Set-Valued Mappings and Enlargements of Monotone Operators*. (Springer, New York, 2008)
31. R.S. Burachik, A.N. Iusem, B.F. Svaiter, Enlargement of monotone operators with applications to variational inequalities. Set-Valued Anal. **5**, 159–180 (1997)
32. R.S. Burachik, C.A. Sagastizábal, B.F. Svaiter, ε-Enlargements of maximal monotone operators: theory and applications, in *Reformulation – Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, ed. by M. Fukushima, L. Qi (Kluwer Academic, 1999), pp. 25–44
33. R.H. Byrd, R.B. Schnabel, G.A. Shultz, A trust region algorithm for nonlinearly constrained optimization. SIAM J. Numer. Anal. **24**, 1152–1170 (1987)
34. R.H. Byrd, N.I.M. Gould, J. Nocedal, R.A. Waltz, An algorithm for nonlinear optimization using linear programming and equality constrained subproblems. Math. Program. **100**, 27–48 (2004)
35. R. Byrd, M. Marazzi, J. Nocedal, On the convergence of Newton iterations to non-stationary points. Math. Program. **99**, 127–148 (2004)
36. R.H. Byrd, N.I.M. Gould, J. Nocedal, R.A. Waltz, On the convergence of successive linear-quadratic programming algorithms. SIAM J. Optim. **16**, 471–489 (2006)
37. R.H. Byrd, F.E. Curtis, J. Nocedal, An inexact SQP method for equality constrained optimization. SIAM J. Optim. **19**, 351–369 (2008)
38. R.H. Byrd, J. Nocedal, R.A. Waltz, Steering exact penalty methods for nonlinear programming. Optim. Method. Softw. **23**, 197–213 (2008)
39. R.H. Byrd, F.E. Curtis, J. Nocedal, An inexact Newton method for nonconvex equality constrained optimization. Math. Program. **122**, 273–299 (2010)
40. Y. Chen, M. Florian, The nonlinear bilevel programming problem: formulations, regularity and optimality conditions. Optimization **32**, 193–209 (1995)
41. C. Chin, R. Fletcher, On the global convergence of an SLP-filter algorithm that takes EQP steps. Math. Program. **96**, 161–177 (2003)
42. M.R. Celis, J.E. Dennis, R.A. Tapia, A trust region algorithm for nonlinear equality constrained optimization, in *Numerical Optimization*, ed. by P.T. Boggs, R.H. Byrd, R.B. Schnabel (SIAM, Philadelphia, 1985), pp. 71–82
43. F.H. Clarke, On the inverse function theorem. Pacisic J. Math. **64**, 97–102 (1976)

44. F.H. Clarke, *Optimization and Nonsmooth Analysis* (Wiley, New York, 1983)
45. A.R. Conn, N.I.M. Gould, Ph.L. Toint, *Trust-Region Methods* (SIAM, Philadelphia, 2000)
46. R.W. Cottle, J.-S. Pang, R.E. Stone, *The Linear Complementarity Problem* (SIAM, Philadelphia, 2009)
47. F.E. Curtis, J. Nocedal, A. Wächter, A matrix-free algorithm for equality constrained optimization problems with rank-deficient Jacobians. SIAM J. Optim. **20**, 1224–1249 (2009)
48. H. Dan, N. Yamashita, M. Fukushima, A superlinearly convergent algorithm for the monotone nonlinear complementarity problem without uniqueness and nondegeneracy conditions. Math. Oper. Res. **27**, 743–754 (2002)
49. A.N. Daryina, A.F. Izmailov, M.V. Solodov, A class of active-set newton methods for mixed complementarity problems. SIAM J. Optim. **15**, 409–429 (2004)
50. A.N. Daryina, A.F. Izmailov, M.V. Solodov, Mixed complementarity problems: regularity, error bounds, and Newton-type methods. Comput. Mathem. Mathem. Phys. **44**, 45–61 (2004)
51. A.N. Daryina, A.F. Izmailov, M.V. Solodov, Numerical results for a globalized active-set Newton method for mixed complementarity problems. Comput. Appl. Math. **24**, 293–316 (2005)
52. T. De Luca, F. Facchinei, C. Kanzow, A semismooth equation approach to the solution of nonlinear complementarity problems. Math. Program. **75**, 407–439 (1996)
53. T. De Luca, F. Facchinei, C. Kanzow, A theoretical and numerical comparison of some semismooth algorithms for complementarity problems. Comput. Optim. Appl. **16**, 173–205 (2000)
54. G. Debreu, Definite and semidefinite quadratic forms. Econometrica **20**, 295–300 (1952)
55. R.S. Dembo, S.C. Eisenstat, T. Steihaug, Inexact Newton methods. SIAM J. Numer. Anal. **19**, 400–408 (1982)
56. S. Dempe, *Foundations of Bilevel Programming* (Kluwer, Dordrecht, 2002)
57. J.E. Dennis, J.J. Moré, A characterization of superlinear convergence and its applications to quasi-Newton methods. Math. Comp. **28**, 549–560 (1974)
58. J.E. Dennis, R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (SIAM, Philadelphia, 1996)
59. J.E. Dennis, M. El-Alem, M.C. Maciel, A global convergence theory for general trust-region-based algorithms for equality constrained optimization. SIAM J. Optim. **7**, 177–207 (1997)
60. S.P. Dirkse, M.C. Ferris, The PATH solver: a non-monotone stabilization scheme for mixed complementarity problems. Optim. Method. Softw. **5**, 123–156 (1995)

61. A.L. Dontchev, R.T. Rockafellar, Characterizations of strong regularity for variational inequalities over polyhedral convex sets. SIAM J. Optim. **6**, 1087–1105 (1996)
62. A.L. Dontchev, R.T. Rockafellar, *Implicit Functions and Solution Mappings* (Springer, New York, 2009)
63. A.L. Dontchev, R.T. Rockafellar, Newton's method for generalized equations: a sequential implicit function theorem. Math. Program. **123**, 139–159 (2010)
64. H.G. Eggleston, *Convexity* (Cambridge University Press, Cambridge, 1958)
65. L.C. Evans, R.F. Gariepy, *Measure Theory and Fine Properties of Functions* (CRC Press, Boca Raton, 1992)
66. M. Fabian, D. Preiss, On the Clarke's generalized Jacobian, in *Proceedings of the 14th Winter School on Abstract Analysis*, Circolo Matematico di Palermo, Palermo, 1987. Rendiconti del Circolo Matematico di Palermo, Ser. II, Number 14, pp. 305–307
67. F. Facchinei, Minimization of SC^1 functions and the Maratos effect. Oper. Res. Lett. **17**, 131–137 (1995)
68. F. Facchinei, J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems* (Springer, New York, 2003)
69. F. Facchinei, J. Soares, A new merit function for nonlinear complementarity problems and a related algorithm. SIAM J. Optim. **7**, 225–247 (1997)
70. F. Facchinei, A. Fischer, C. Kanzow, Regularity properties of a semismooth reformulation of variational inequalities. SIAM J. Optim. **8**, 850–869 (1998)
71. F. Facchinei, A. Fischer, C. Kanzow, On the accurate identifcation of active constraints. SIAM J. Optim. **9**, 14–32 (1999)
72. F. Facchinei, A. Fischer, C. Kanzow, J.-M. Peng, A simply constrained optimization reformulation of KKT systems arising from variational inequalities. Appl. Math. Optim. **40**, 19–37 (1999)
73. D. Fernández, A quasi-Newton strategy for the sSQP method for variational inequality and optimization problems. Math. Program. **137**, 199–223 (2013)
74. D. Fernández, M.V. Solodov, On local convergence of sequential quadratically-constrained quadratic-programming type methods, with an extension to variational problems. Comput. Optim. Appl. **39**, 143–160 (2008)
75. D. Fernández, M. Solodov, Stabilized sequential quadratic programming for optimization and a stabilized Newton-type method for variational problems. Math. Program. **125**, 47–73 (2010)
76. D. Fernández, M. Solodov, Local convergence of exact and inexact augmented Lagrangian methods under the second-order sufficient optimality condition. SIAM J. Optim. **22**, 384–407 (2012)

77. D. Fernández, A.F. Izmailov, M.V. Solodov, Sharp primal superlinear convergence results for some Newtonian methods for constrained optimization. SIAM J. Optim. **20**, 3312–3334 (2010)
78. D. Fernández, E.A. Pilotta, G.A. Torres, An inexact restoration strategy for the globalization of the sSQP method. Comput. Optim. Appl. **54**, 595–617 (2013)
79. M.C. Ferris, F. Tin-Loi, On the solution of a minimum weight elastoplastic problem involving displacement and complementarity constraints. Comput. Method. Appl. Mech. Eng. **174**, 108–120 (1999)
80. M.C. Ferris, C. Kanzow, T.S. Munson, Feasible descent algorithms for mixed complementarity problems. Math. Program. **86**, 475–497 (1999)
81. P. Finsler, Über das vorkommen definiter und semidefiniter formen und scharen quadratischer formen. Comment. Math. Helv. **94**, 188–192 (1937)
82. A. Fischer, A special Newton-type optimization method. Optimization **24**, 296–284 (1992)
83. A. Fischer, Solution of monotone complementarity problems with locally Lipschitzian function. Math. Program. **76**, 513–532 (1997)
84. A. Fischer, Modified Wilson's method for nonlinear programs with nonunique multipliers. Math. Oper. Res. **24**, 699–727 (1999)
85. A. Fischer, Local behavior of an iterative framework for generalized equations with nonisolated solutions. Math. Program. **94**, 91–124 (2002)
86. A. Fischer, A. Friedlander, A new line search inexact restoration approach for nonlinear programming. Comput. Optim. Appl. **46**, 333–346 (2010)
87. R. Fletcher, A model algorithm for composite non-differentiable optimization problems. Math. Program. **17**, 67–76 (1982)
88. R. Fletcher, Second order corrections for non-differentiable optimization, in *Numerical Analysis*, ed. by D. Griffiths (Springer, Berlin, 1982), pp. 85–114
89. R. Fletcher, *Practical Methods of Optimization*, 2nd edn. (Wiley, New York, 1987)
90. R. Fletcher, E.S. de la Maza, Nonlinear programming and nonsmooth optimization by successive linear programming. Math. Program. **43**, 235–256 (1989)
91. R. Fletcher, S. Leyffer, Nonlinear programming without a penalty function. Math. Program. **91**, 239–269 (2002)
92. R. Fletcher, S. Leyffer, Solving mathematical programs with equilibrium constraints as nonlinear programs. Optim. Method. Softw. **19**, 15–40 (2004)
93. R. Fletcher, N. Gould, S. Leyffer, P. Toint, A. Wächter, Global convergence of trust-region and SQP-filter algorithms for general nonlinear programming. SIAM J. Optim. **13**, 635–659 (2002)
94. R. Fletcher, S. Leyffer, P.L. Toint, On the global convergence of a filter-SQP algorithm. SIAM J. Optim. **13**, 44–59 (2002)

95. R. Fletcher, S. Leyffer, D. Ralph, S. Scholtes, Local convergence of SQP methods for mathematical programs with equilibrium constraints. SIAM J. Optim. **17**, 259–286 (2006)
96. M.P. Friedlander, M.A. Saunders, A globally convergent linearly constrained Lagrangian method for nonlinear optimization. SIAM J. Optim. **15**, 863–897 (2005)
97. M. Fukushima, Z.-Q. Luo, P. Tseng, A sequential quadratically constrained quadratic programming method for differentiable convex minimization. SIAM J. Optim. **13**, 1098–1119 (2003)
98. J. Gauvin, A necessary and sufficient regularity conditions to have bounded multipliers in nonconvex programming. Math. Program. **12**, 136–138 (1977)
99. P.E. Gill, D.P. Robinson, A globally convergent stabilized SQP Method. SIAM J. Optim. **23**, 1983–2010 (2013)
100. P.E. Gill, W. Murray, M.H. Wright, *Numerical Linear Algebra and Optimization*, vol 1 (Addison Wesley, Redwood City, 1991)
101. P.E. Gill, W. Murray, M.A. Saunders, M.H. Wright, Some theoretical properties of an augmented Lagrangian merit function, in *Advances in Optimization and Parallel Computing*, ed. by P.M. Pardalos (North–Holland, Amsterdam, 1992), pp. 101–128
102. P.E. Gill, W. Murray, M.A. Saunders, SNOPT: an SQP algorithm for large-scale constrained optimization. SIAM Rev. **47**, 99–131 (2005)
103. G.H. Golub, C.F. Van Loan, *Matrix Computations*, 3rd edn. (The Johns Hopkins University Press, Baltimore, 1996)
104. M. Golubitsky, D.G. Schaeffer, *Singularities and Groups in Bifurcation Theory*, vol 1. (Springer, New York, 1985)
105. C.C. Gonzaga, E.W. Karas, M. Vanti, A globally convergent filter method for nonlinear programming. SIAM J. Optim. **14**, 646–669 (2003)
106. N.I.M. Gould, Some reflections on the current state of active-set and interior-point methods for constrained optimization. Numerical Analysis Group Internal Report 2003-1. Computational Science and Engineering Department, Rutherford Appleton Laboratory, Oxfordshire, 2003
107. N.I.M. Gould, D.P. Robinson, A second derivative SQP method: global convergence and practical issues. SIAM J. Optim. **20**, 2023–2048 (2010)
108. N.I.M. Gould, D.P. Robinson, A second derivative SQP method: local convergence and practical issues. SIAM J. Optim. **20**, 2049–2079 (2010)
109. N.I.M. Gould, D. Orban, Ph.L. Toint, Numerical methods for large-scale nonlinear optimization. Acta Numer. **14**, 299–361 (2005)
110. L. Grippo, F. Lampariello, S. Lucidi, A nonmonotone line search technique for Newton's method. SIAM J. Numer. Anal. **23**, 707–716 (1986)
111. O. Güler, F. Gürtuna, O. Shevchenko, Duality in quasi-Newton methods and new variational characterizations of the DFP and BFGS updates. Optim. Method. Softw. **24**, 45–62 (2009)
112. W.W. Hager, Stabilized sequential quadratic programming. Comput. Optim. Appl. **12**, 253–273 (1999)

113. W.W. Hager, M.S. Gowda, Stability in the presence of degeneracy and error estimation. Math. Program. **85**, 181–192 (1999)
114. S.-P. Han, A globally convergent method for nonlinear programming. J. Optim. Theory Appl. **22**, 297–309 (1977)
115. J. Han, D. Sun, Superlinear convergence of approximate Newton methods for LC^1 optimization problems without strict complementarity, in *Recent Advances in Nonsmooth Optimization*, vol 58, ed. by D.-Z. Du, L. Qi, R.S. Womersley (World Scientific, Singapore, 1993), pp. 353–367
116. J. Herskovits, *A Two-Stage Feasible Direction Algorithm Including Variable Metric Techniques for Nonlinear Optimization Problems*. Rapport de Recherche 118 (INRIA, Rocqencourt, 1982)
117. J. Herskovits, A two-stage feasible directions algorithm for nonlinear constrained optimization. Math. Program. **36**, 19–38 (1986)
118. J. Herskovits, Feasible direction interior-point technique for nonlinear optimization. J. Optim. Theory Appl. **99**, 121–146 (1998)
119. J.-B. Hiriart-Urruty, Mean-value theorems in nonsmooth analysis. Numer. Func. Anal. Optim. **2**, 1–30 (1980)
120. J.-B. Hiriart-Urruty, J.-J. Strodiot, V.H. Nguyen, Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data. Appl. Math. Optim. **11**, 43–56 (1984)
121. T. Hoheisel, C. Kanzow, First- and second-order optimality conditions for mathematical programs with vanishing constraints. Appl. Math. **52**, 495–514 (2007)
122. T. Hoheisel, C. Kanzow, Stationarity conditions for mathematical programs with vanishing constraints using weak constraint qualifications. J. Math. Anal. Appl. **337**, 292–310 (2008)
123. T. Hoheisel, C. Kanzow, On the Abadie and Guignard constraint qualifications for mathematical programs with vanishing constraints. Optimization **58**, 431–448 (2009)
124. Z.-H. Huang, D. Sun, G. Zhao, A smoothing Newton-type algorithm of stronger convergence for the quadratically constrained convex quadratic programming. Comput. Optim. Appl. **35**, 199–237 (2006)
125. C.-M. Ip, J. Kyparisis, Local convergence of quasi-Newton methods for B-differentiable equations. Math. Program. **56**, 71–89 (1992)
126. A.F. Izmailov, On the Lagrange methods for finding degenerate solutions of constrained optimization problems. Comput. Math. Math. Phys. **36**, 423–429 (1996)
127. A.F. Izmailov, Mathematical programs with complementarity constraints: regularity, optimality conditions, and sensitivity. Comput. Math. Math. Phys. **44**, 1145–1164 (2004)
128. A.F. Izmailov, On the analytical and numerical stability of critical Lagrange multipliers. Comput. Math. Math. Phys. **45**, 930–946 (2005)
129. A.F. Izmailov, Sensitivity of solutions to systems of optimality conditions under the violation of constraint qualifications. Comput. Math. Math. Phys. **47**, 533–554 (2007)

130. A.F. Izmailov, Solution sensitivity for Karush–Kuhn–Tucker systems with nonunique Lagrange multipliers. Optimization **59**, 747–775 (2010)
131. A.F. Izmailov, Strongly regular nonsmooth generalized equations. Math. Program. (2013). doi:10.1007/s10107-013-0717-1
132. A.F. Izmailov, M.M. Golishnikov, Newton-type methods for constrained optimization with nonregular constraints. Comput. Math. Math. Phys. **46**, 1299–1319 (2006)
133. A.F. Izmailov, A.S. Kurennoy, Partial Clarke generalized Jacobian and other generalized differentials, in *Theoretical and Applied Problems of Nonlinear Analysis*, ed. by V.A. Bereznyov (Computing Center RAS, Moscow, 2010, in Russian), pp. 77–90
134. A.F. Izmailov, A.S. Kurennoy, On regularity conditions for complementarity problems. Comput. Optim. Appl. (2013). doi:10.1007/s10589-013-9604-1
135. A.F. Izmailov, A.L. Pogosyan, Optimality conditions and Newton-type methods for mathematical programs with vanishing constraints. Comput. Math. Math. Phys. **49**, 1128–1140 (2009)
136. A.F. Izmailov, A.L. Pogosyan, A semismooth sequential quadratic programming method for lifted mathematical programs with vanishing constraints. Comput. Math. Math. Phys. **51**, 919–941 (2011)
137. A.F. Izmailov, A.L. Pogosyan, Active-set Newton methods for mathematical programs with vanishing constraints. Comput. Optim. Appl. **53**, 425–452 (2012)
138. A.F. Izmailov, M.V. Solodov, Superlinearly convergent algorithms for solving singular equations and smooth reformulations of complementarity problems. SIAM J. Optim. **13**, 386–405 (2002)
139. A.F. Izmailov, M.V. Solodov, The theory of 2-regularity for mappings with Lipschitzian derivatives and its applications to optimality conditions. Math. Oper. Res. **27**, 614–635 (2002)
140. A.F. Izmailov, M.V. Solodov, Karush–Kuhn–Tucker systems: regularity conditions, error bounds and a class of Newton-type methods. Math. Program. **95**, 631–650 (2003)
141. A.F. Izmailov, M.V. Solodov, Newton-type methods for optimization problems without constraint qualifications. SIAM J. Optim. **15**, 210–228 (2004)
142. A.F. Izmailov, M.V. Solodov, A note on solution sensitivity for Karush–Kuhn–Tucker systems. Math. Meth. Oper. Res. **61**, 347–363 (2005)
143. A.F. Izmailov, M.V. Solodov, An active-set Newton method for mathematical programs with complementarity constraints. SIAM J. Optim. **19**, 1003–1027 (2008)
144. A.F. Izmailov, M.V. Solodov, Examples of dual behaviour of Newton-type methods on optimization problems with degenerate constraints. Comput. Optim. Appl. **42**, 231–264 (2009)

145. A.F. Izmailov, M.V. Solodov, Mathematical programs with vanishing constraints: optimality conditions, sensitivity, and a relaxation method. J. Optim. Theory Appl. **142**, 501–532 (2009)
146. A.F. Izmailov, M.V. Solodov, On attraction of Newton-type iterates to multipliers violating second-order sufficiency conditions. Math. Program. **117**, 271–304 (2009)
147. A.F. Izmailov, M.V. Solodov, A truncated SQP method based on inexact interior-point solutions of subproblems. SIAM J. Optim. **20**, 2584–2613 (2010)
148. A.F. Izmailov, M.V. Solodov, Inexact Josephy–Newton framework for generalized equations and its applications to local analysis of Newtonian methods for constrained optimization. Comput. Optim. Appl. **46**, 347–368 (2010)
149. A.F. Izmailov, M.V. Solodov, On attraction of linearly constrained Lagrangian methods and of stabilized and quasi-Newton SQP methods to critical multipliers. Math. Program. **126**, 231–257 (2011)
150. A.F. Izmailov, M.V. Solodov, Stabilized SQP revisited. Math. Program. **133**, 93–120 (2012)
151. A.F. Izmailov, A.A. Tret'yakov, *Factor-Analysis of Nonlinear Mappings* (Nauka, Moscow, 1994) (in Russian)
152. A.F. Izmailov, A.A. Tret'yakov, *2-Regular Solutions of Nonlinear Problems* (Nauka, Moscow, 1999) (in Russian)
153. A.F. Izmailov, E.I. Uskov, Attraction of Newton method to critical Lagrange multipliers: fully quadratic case (2013). Available at http://www.optimization-online.org/DB_HTML/2013/02/3761.html
154. A.F. Izmailov, A.L. Pogosyan, M.V. Solodov, Semismooth SQP method for equality-constrained optimization problems with an application to the lifted reformulation of mathematical programs with complementarity constraints. Optim. Method. Softw. **26**, 847–872 (2011)
155. A.F. Izmailov, A.S. Kurennoy, M.V. Solodov, A note on upper Lipschitz stability, error bounds, and critical multipliers for Lipschitz-continuous KKT systems. Math. Program. **142**, 591–604 (2013)
156. A.F. Izmailov, A.L. Pogosyan, M.V. Solodov, Semismooth Newton method for the lifted reformulation of mathematical programs with complementarity constraints. Comput. Optim. Appl. **51**, 199–221 (2012)
157. A.F. Izmailov, M.V.Solodov, E.I. Uskov, Augmented Lagrangian methods applied to optimization problems with degenerate constraints, including problems with complementarity constraints. SIAM J. Optim. **22**, 1579–1606 (2012)
158. A.F. Izmailov, A.S. Kurennoy, M.V. Solodov, The Josephy–Newton method for semismooth generalized equations and semismooth SQP for optimization. Set-Valued Variational Anal. **21**, 17–45 (2013)

159. A.F. Izmailov, A.S. Kurennoy, M.V. Solodov, Composite-step constrained optimization methods interpreted via the perturbed sequential quadratic programming framework (2013). IMPA preprint A2405
160. H. Jäger, E.W. Sachs, Global convergence of inexact reduced SQP methods. Optim. Method. Softw. **7**, 83–110 (1997)
161. N.H. Josephy, Newton's method for generalized equations. Technical Summary Report No. 1965. Mathematics Research Center, University of Wisconsin, Madison, 1979
162. N.H. Josephy, Quasi-Newton methods for generalized equations. Technical Summary Report No. 1966. Mathematics Research Center, University of Wisconsin, Madison, 1979
163. C. Kanzow, Strictly feasible equation-based methods for mixed complementarity problems. Numer. Math. **89**, 135–160 (2001)
164. C. Kanzow, M. Fukushima, Solving box constrained variational inequalities by using the natural residual with D-gap function globalization. Oper. Res. Lett. **23**, 45–51 (1998)
165. E. Karas, A. Ribeiro, C. Sagastizábal, M. Solodov, A bundle-filter method for nonsmooth convex constrained optimization. Math. Program. **116**, 297–320 (2009)
166. W. Karush, Minima of functions of several variables with inequalities as side constraints. Technical Report (master's thesis), Department of Mathematics, University of Chicago, 1939
167. T. Kato, *Perturbation Theory for Linear Operators* (Springer, Berlin, 1984)
168. D. Klatte, Upper Lipschitz behaviour of solutions to perturbed $C^{1,1}$ programs. Math. Program. **88**, 285–311 (2000)
169. D. Klatte, B. Kummer, *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications* (Kluwer Academic, Dordrecht, 2002)
170. D. Klatte, K. Tammer, On the second order sufficient conditions to perturbed $C^{1,1}$ optimization problems. Optimization **19**, 169–180 (1988)
171. M. Kojima, Strongly stable stationary solutions in nonlinear programs, in *Analysis and Computation of Fixed Points*, ed. by S.M. Robinson (Academic, NewYork, 1980), pp. 93–138
172. M. Kojima, S. Shindo, Extensions of Newton and quasi-Newton methods to systems of PC^1 equations. J. Oper. Res. Soc. Japan **29**, 352–374 (1986)
173. S. Kruk, H. Wolkowicz, Sequential, quadratically constrained, quadratic programming for general nonlinear programming, in *Handbook of Semidefinite Programming*, ed. by H. Wolkowicz, R. Saigal, L. Vandenberghe (Kluwer Academic, Dordrecht, 2000), pp. 563–575
174. H.W. Kuhn, A.W. Tucker, Non-linear programming. in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, ed. by J. Neyman (University of California Press, Berkeley, 1951), pp. 481–493

175. B. Kummer, Newton's method for nondifferentiable functions, in *Advances in Mathematical Optimization*, vol 45, ed. by J. Guddat, B. Bank, H. Hollatz, P. Kall, D. Klatte, B. Kummer, K. Lommalzsch, L. Tammer, M. Vlach, K. Zimmerman (Akademie-Verlag, Berlin, 1988), pp. 114–125
176. B. Kummer, Newton's method based on generalized derivatives for nonsmooth functions, in *Advances in Optimization*, ed. by W. Oettli, D. Pallaschke (Springer, Berlin, 1992), pp. 171–194
177. F. Leibfritz, E.W. Sachs, Inexact SQP interior point methods and large scale optimal control problems. SIAM J. Control Optim. **38**, 272–293 (1999)
178. A. Levy, Solution sensitivity from general principles. SIAM J. Control Optim. **40**, 1–38 (2001)
179. A.S. Lewis, S.J. Wright, Identifying activity. SIAM J. Optim. **21**, 597–614 (2011)
180. D.-H. Li, L. Qi, Stabilized SQP method via linear equations. Applied Mathematics Technical Reptort AMR00/5. University of New South Wales, Sydney, 2000
181. G.H. Lin, M. Fukushima, Hybrid approach with active set identification for mathematical programs with complementarity constraints. J. Optim. Theory Appl. **128**, 1–28 (2006)
182. M.S. Lobo, L. Vandenberghe, S. Boyd, H. Lebret, Applications of second-order cone programming. Linear Algebra Appl. **284**, 193–228 (1998)
183. P.A. Lotito, L.A. Parente, M.V. Solodov, A class of variable metric decomposition methods for monotone variational inclusions. J. Convex Anal. **16**, 857–880 (2009)
184. Z.-Q. Luo, J.-S. Pang, D. Ralph, *Mathematical Programs with Equilibrium Constraints* (Cambridge University Press, Cambridge, 1996)
185. MacMPEC. (2012) http://wiki.mcs.anl.gov/leyffer/index.php/MacMPEC
186. O.L. Mangasarian, *Nonlinear Programming* (SIAM, Philadelphia, 1994)
187. O.L. Mangasarian, S. Fromovitz, The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. J. Math. Anal. Appl. **17**, 37–47 (1967)
188. N. Maratos, Exact penalty function algorithms for finite dimensional and control optimization problems. Ph.D. thesis, University of London, UK, 1978
189. P. Marcotte, J.-P. Dussault, A note on a globally convergent Newton method for solving monotone variational inequalities. Oper. Res. Lett. **6**, 35–42 (1987)
190. B. Martinet, Regularisation d'inequations variationelles par approximations successives. Revue Française d'Informatique et de Recherche Opérationelle **4**, 154–159 (1970)
191. J.M. Martínez, Practical quasi-Newton methods for solving nonlinear systems. J. Comput. Appl. Math. **124**, 97–121 (2000)

192. J.M. Martínez, Inexact restoration method with Lagrangian tangent decrease and new merit function for nonlinear programming. J. Optim. Theory Appl. **111**, 39–58 (2001)
193. J.M. Martínez, E.A. Pilotta, Inexact restoration algorithms for constrained optimization. J. Optim. Theory Appl. **104**, 135–163 (2000)
194. J.M. Martínez, E.A. Pilotta, Inexact restoration methods for nonlinear programming: advances and perspectives, in *Optimization and Control with Applications*, ed. by L.Q. Qi, K.L. Teo, X.Q. Yang (Springer, New York, 2005), pp. 271–292
195. J.M. Martínez, L. Qi, Inexact Newton methods for solving nonsmooth equations. J. Comput. Appl. Math. **60**, 127–145 (1995)
196. W.F. Mascarenhas, Newton's iterates can converge to non-stationary points. Math. Program. **112**, 327–334 (2008)
197. D.Q. Mayne, E. Polak, Feasible direction algorithms for optimization problems with equality and inequality constraints. Math. Program. **11**, 67–80 (1976)
198. R.M. McLeod, Mean value theorems for vector valued functions. Proc. Edinburgh Math. Soc. **14**, 197–209 (1965)
199. R. Mifflin, Semismooth and semiconvex functions in constrained optimization. SIAM J. Control Optim. **15**, 959–972 (1977)
200. J.L. Morales, J. Nocedal, Y. Wu, A sequential quadratic programming algorithm with an additional equality constrained phase. IMA J. Numer. Anal. **32**, 553–579 (2012)
201. B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation* (Springer, Berlin, 2006)
202. J.J. Moré, D.C. Sorensen, Computing a trust region step. SIAM J. Sci. Stat. Comput. **4**, 553–572 (1983)
203. E.M.E. Mostafa, L.N. Vicente, S.J. Wright, Numerical behavior of a stabilized SQP method for degenerate NLP problems, in *Global Optimization and Constraint Satisfaction*, ed. by C. Bliek, C. Jermann, A. Neumaier. Lecture Notes in Computer Science, vol 2861 (Springer, Berlin, 2003), pp. 123–141
204. W. Murray, F.J. Prieto, A sequential quadratic programming algorithm using an incomplete solution of the subproblem. SIAM J. Optim. **5**, 590–640 (1995)
205. B.A. Murtagh, M.A. Saunders, A projected Lagrangian algorithm and its implementation for sparse nonlinear constraints. Math. Program. Study. **16**, 84–117 (1982)
206. B.A. Murtagh, M.A. Saunders, MINOS 5.0 user's guide. Technical Report SOL 83.20. Stanford University, 1983
207. Y.E. Nesterov, A.S. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming: Theory and Applications* (SIAM Publications, Philadelphia, 1993)
208. J. Nocedal, S.J. Wright, *Numerical Optimization*, 2nd edn. (Springer, New York, 2006)

209. C. Oberlin, S.J. Wright, An accelerated Newton method for equations with semismooth Jacobians and nonlinear complementarity problems. Math. Program. **117**, 355–386 (2009)
210. E.O. Omojokun, Trust region algorithms for optimization with nonlinear equality and inequality constraints. Ph.D. thesis, Department of Computer Science, University of Colorado at Boulder, U.S.A., 1989
211. J.V. Outrata, M. Kocvara, J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results* (Kluwer Academic, Dordrecht, 1998)
212. J.-S. Pang, S.A. Gabriel, NE/SQP: a robust algorithm for the nonlinear complementarity problem. Math. Program. **60**, 295–338 (1993)
213. J.-S. Pang, L. Qi, Nonsmooth equations: motivation and algorithms. SIAM J. Optim. **3**, 443–465 (1993)
214. E.R. Panier, A.L. Tits, J. Herskovits, A QP-free, globally convergent, locally superlinearly convergent algorithm for inequality constrained optimization. SIAM J. Control Optim. **26**, 788–811 (1988)
215. V.M. Panin, A second-order method for discrete min-max problem. USSR Comput. Math. Math. Phys. **19**, 90–100 (1979)
216. L.A. Parente, P.A. Lotito, M.V. Solodov, A class of inexact variable metric proximal point algorithms. SIAM J. Optim. **19**, 240–260 (2008)
217. J.-M. Peng, M. Fukushima, A hybrid Newton method for solving the variational inequality problem via the D-gap function. Math. Program. **86**, 367–386 (1999)
218. J.-M. Peng, C. Kanzow, M. Fukushima, A hybrid Josephy-Newton method for solving box constrained variational inequality problem via the D-gap function. Optim. Method. Softw. **10**, 687–710 (1999)
219. M.J.D. Powell, A fast algorithm for nonlinearly constrained optimization calculations, in *Numerical Analysis Dundee 1977*, ed. by G.A. Watson. Lecture Notes in Mathematics, vol 630 (Springer, Berlin, 1978), pp. 144–157
220. M.J.D. Powell, Y. Yuan, A recursive quadratic programming algorithm that uses differentiable exact penalty function. Math. Program. **35**, 265–278 (1986)
221. L. Qi, LC^1 functions and LC^1 optimization problems. Technical Report AMR 91/21. School of Mathematics, The University of New South Wales, Sydney, 1991
222. L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res. **18**, 227–244 (1993)
223. L. Qi, Superlinearly convergent approximate Newton methods for LC^1 optimization problems. Math. Program. **64**, 277–294 (1994)
224. L. Qi, H. Jiang, Semismooth Karush–Kuhn–Tucker equations and convergence analysis of Newton and quasi-Newton methods for solving these equations. Math. Oper. Res. **22**, 301–325 (1997)
225. L. Qi, J. Sun, A nonsmooth version of Newton's method. Math. Program. **58**, 353–367 (1993)

226. H.-D. Qi, L. Qi, A new QP-free, globally convergent, locally superlinearlly convergent algorithm for inequality constrained optimization. SIAM J. Optim. **11**, 113–132 (2000)
227. H. Rademacher, Über partielle und totale differenzierbarkeit I. Math. Ann. **89**, 340–359 (1919)
228. D. Ralph, Global convergence of damped Newtons method for nonsmooth equations, via the path search. Math. Oper. Res. **19**, 352–389 (1994)
229. D. Ralph, Sequential quadratic programming for mathematical programs with linear complementarity constraints, in *Computational Techniques and Applications: CTAC95*, ed. by R.L. May, A.K. Easton (World Scientific, Singapore, 1996), pp. 663–668
230. A.R. Ribeiro, E.W. Karas, C.C. Gonzaga, Global convergence of filter methods for nonlinear programming. SIAM J. Optim. **14**, 646–669 (2008)
231. S.M. Robinson, A quadratically convergent algorithm for general nonlinear programming problems. Math. Program. **3**, 145–156 (1972)
232. S.M. Robinson, Perturbed Kuhn–Tucker points and rates of convergence for a class of nonlinear-programming algorithms. Math. Program. **7**, 1–16 (1974)
233. S.M. Robinson, Stability theory for systems of inequalities, Part II. Differentiable nonlinear systems, SIAM J. Numer. Anal. **13**, 497–513 (1976)
234. S.M. Robinson, Strongly regular generalized equations. Math. Oper. Res. **5**, 43–62 (1980)
235. R.T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, 1970)
236. R.T. Rockafellar, Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**, 877–898 (1976)
237. R.T. Rockafellar, Computational schemes for solving large-scale problems in extended linear-quadratic programming. Math. Program. **48**, 447–474 (1990)
238. R.T. Rockafellar, J.B. Wets, Generalized linear-quadratic problems of deterministic and stochastic optimal control in discrete time. SIAM J. Control Optim. **28**, 810–922 (1990)
239. R.T. Rockafellar, R.J.-B. Wets, *Variational Analysis* (Springer, Berlin, 1998)
240. H. Scheel, S. Scholtes, Mathematical programs with complementarity constraints: stationarity, optimality and sensitivity. Math. Oper. Res. **25**, 1–22 (2000)
241. S. Scholtes, M. Stöhr, How stringent is the linear independence assumption for mathematical programs with complementarity constraints? Math. Oper. Res. **26**, 851–863 (2001)
242. E. Schröder, Über unendlich viele algorithmen zur auflösung der gleichungen. Math. Annalen. **2**, 317–365 (1870)

243. D. Serre, *Matrices: Theory and Applications*, 2nd edn. (Springer, New York, 2010)
244. A. Shapiro, On concepts of directional differentiability. J. Optim. Theory Appl. **66**, 477–487 (1990)
245. A. Shapiro, Sensitivity analysis of generalized equations. J. Math. Sci. **115**, 2554–2565 (2003)
246. M.V. Solodov, A class of globally convergent algorithms for pseudomonotone variational inequalities, in *Complementarity: Applications, Algorithms and Extensions*, ed. by M.C. Ferris, O.L. Mangasarian, J.-S. Pang (Kluwer Academic, Dordrecht, 2001), pp. 297–315
247. M.V. Solodov, A class of decomposition methods for convex optimization and monotone variational inclusions via the hybrid inexact proximal point framework. Optim. Method. Softw. **19**, 557–575 (2004)
248. M.V. Solodov, On the sequential quadratically constrained quadratic programming methods. Math. Oper. Res. **29**, 64–79 (2004)
249. M.V. Solodov, Global convergence of an SQP method without boundedness assumptions on any of the iterative sequences. Math. Program. **118**, 1–12 (2009)
250. M.V. Solodov, Constraint qualifications, in *Wiley Encyclopedia of Operations Research and Management Science*, ed. by J.J. Cochran (Wiley, London, 2010)
251. M.V. Solodov, B.F. Svaiter, A globally convergent inexact Newton method for systems of monotone equations, in *Reformulation – Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, ed. by M. Fukushima, L. Qi (Kluwer Academic, Dordrecht, 1999), pp. 355–369
252. M.V. Solodov, B.F. Svaiter, A hybrid approximate extragradient–proximal point algorithm using the enlargement of a maximal monotone operator. Set-Valued Anal. **7**, 323–345 (1999)
253. M.V. Solodov, B.F. Svaiter, A hybrid projection–proximal point algorithm. J. Convex Anal. **6**, 59–70 (1999)
254. M.V. Solodov, B.F. Svaiter, A truly globally convergent Newton-type method for the monotone nonlinear complementarity problem SIAM J. Optim. **10**, 605–625 (2000)
255. M.V. Solodov, B.F. Svaiter, Error bounds for proximal point subproblems and associated inexact proximal point algorithms. Math. Program. **88**, 371–389 (2000)
256. M.V. Solodov, B.F. Svaiter, A unified framework for some inexact proximal point algorithms. Numer. Func. Anal. Optim. **22**, 1013–1035 (2001)
257. M.V. Solodov, B.F. Svaiter, A new proximal-based globalization strategy for the Josephy-Newton method for variational inequalities. Optim. Method. Softw. **17**, 965–983 (2002)
258. O. Stein, Lifting mathematical programs with complementarity constraints. Math. Program. **131**, 71–94 (2012)

259. K. Taji, M. Fukushima, T. Ibaraki, A globally convergent Newton method for solving strongly monotone variational inequalities. Math. Program. **58**, 369–383 (1993)
260. A.L. Tits, A. Wächter, S. Bakhtiari, T.J. Urban, C.T. Lawrence, A primal-dual interior-point method for nonlinear programming with strong global and local convergence properties. SIAM J. Optim. **14**, 173–199 (2003)
261. L. Trefethen, D. Bau, *Numerical Linear Algebra* (SIAM, Philadelphia, 1997)
262. K. Ueda, N. Yamashita, Convergence properties of the regularized Newton method for the unconstrained nonconvex optimization. Appl. Math. Optim. **62**, 27–46 (2010)
263. A. Vardi, A trust region algorithm for equality constrained minimization: convergence properties and implementation. SIAM J. Numer. Anal. **22**, 575–591 (1985)
264. A. Wächter, L. Biegler, Line search filter methods for nonlinear programming: local convergence. SIAM J. Optim. **16**, 32–48 (2005)
265. A. Wächter, L. Biegler, Line search filter methods for nonlinear programming: motivation and global convergence. SIAM J. Optim. **16**, 1–31 (2005)
266. E.J. Wiest, E. Polak, A generalized quadratic programming-based phase-I–phase-II method for inequality-constrained optimization. Appl. Math. Optim. **26**, 223–252 (1992)
267. K.A. Williamson, A robust trust region algorithm for nonlinear programming. Ph.D. thesis, Department of Math. Sciences, Rice University, Houston, 1990
268. R.B. Wilson, A simplicial algorithm for concave programming. Ph.D. thesis, Graduate School of Business Administration, Harvard University, U.S.A., 1963
269. S.J. Wright, Superlinear convergence of a stabilized SQP method to a degenerate solution. Comput. Optim. Appl. **11**, 253–275 (1998)
270. S.J. Wright, Modifying SQP for degenerate problems. SIAM J. Optim. **13**, 470–497 (2002)
271. S.J. Wright, Constraint identification and algorithm stabilization for degenerate nonlinear programs. Math. Program. **95**, 137–160 (2003)
272. S.J. Wright, An algorithm for degenerate nonlinear programming with rapid local convergence. SIAM J. Optim. **15**, 673–696 (2005)
273. Y.-F. Yang , D.-H. Li, L. Qi, A feasible sequential linear equation method for inequality constrained optimization. SIAM J. Optim. **13**, 1222–1244 (2003)
274. Y. Yuan, On the convergence of a new trust region algorithm. Numer. Math. **70**, 515–539 (1995)

Index

A.F. Izmailov and M.V. Solodov, *Newton-Type Methods for Optimization and Variational Problems*, Springer Series in Operations Research and Financial Engineering, DOI 10.1007/978-3-319-04247-3,

H

L

M