5

Bound Constrained Minimization

We shall now be concerned with the *bound constrained problem* to find

$$\min_{\mathbf{x}\in\Omega_B} f(\mathbf{x}) \tag{5.1}$$

with $\Omega_B = \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{x} \geq \boldsymbol{\ell}\}$, $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathsf{A}\mathbf{x} - \mathbf{x}^T\mathbf{b}$, $\boldsymbol{\ell}$ and $\mathbf{b}$ given column n-vectors, and A an $n \times n$ symmetric positive definite matrix. To include the possibility that not all the components of $\mathbf{x}$ are constrained, we admit $\ell_i = -\infty$. Here we are again interested in large, sparse problems with a well-conditioned A, and in algorithms that can be used also for the solution of equality and inequality constrained problems. Such algorithms should be able to return an approximate solution at a cost proportional to the precision and to recognize an acceptable solution when it is found.

Our choice is the active set strategy with auxiliary problems solved approximately by the conjugate gradient method introduced in Sect. 3.5. It turns out that this type of algorithm can exploit effectively the specific structure of Ω_B, including the possibility to evaluate the projections in the Euclidean norm. We shall show that the resulting algorithm has an R-linear rate of convergence. If its parameters are chosen properly, the algorithm enjoys the finite termination property, even in the dual degenerate case with some active constraints corresponding to zero multipliers. We consider the finite termination property important, as it indicates that the algorithm does not suffer from undesirable oscillations and can exploit the superconvergence properties of the conjugate gradient method for linear problems.

As in the previous chapter, we first briefly review alternative algorithms for the solution of bound constrained problems. Then we introduce a basic active set algorithm and its modifications that are motivated by our effort to get the results on the rate of convergence in terms of bounds on the spectrum of the Hessian matrix A and on the finite termination. We restricted our attention to bound constrained problems because of their special structure which we exploit in the development of our algorithms. Let us recall that the problems with more general inequality constraints can be reduced to (5.1) by duality.

Zdeněk Dostál, *Optimal Quadratic Programming Algorithms*,
Springer Optimization and Its Applications, DOI 10.1007/978-0-387-84806-8_5,
© Springer Science+Business Media, LLC 2009

Overview of algorithms

The *exact working (active) set method* of Sect. 5.3 reduces the solution of (5.1) to a sequence of unconstrained problems that are defined by the bounds which are assumed to be active at the solution. See also Algorithm 5.1. The performance of the algorithm is explained by the combinatorial arguments.

The *Polyak algorithm* is a variant of the working set method which solves the auxiliary linear problems by the conjugate gradient method. The active set is expanded whenever the unfeasible iterate is generated, typically by one index, but it is reduced only after the exact solution of an auxiliary unconstrained problem is found. The algorithm is described in Sect. 5.4. See Algorithm 5.2 for the formal description.

The *looking ahead Polyak algorithm* is based on observation that it is possible to recognize the incorrect active set before reaching the solution of the auxiliary unconstrained problem. The algorithm accepts inexact solutions of auxiliary unconstrained problems and preserves the finite termination property of the original Polyak algorithm. The algorithm is described in Sect. 5.5.2. See also Algorithm 5.3.

Even more relaxed solutions of the auxiliary unconstrained problems are accepted by the *easy re-release Polyak algorithm* of Sect 5.5.3. The algorithm preserves the finite termination property of the Polyak-type algorithms.

Unlike the Polyak-type algorithms, the *gradient projection with a fixed steplength* can typically add several indices to the active set in each step and it has established linear convergence in the bounds on the spectrum of the Hessian matrix. The algorithm is described in Sect. 5.6.3.

The *MPGP* (modified proportioning with gradient projections) algorithm of Sect. 5.7 uses the conjugate gradients to solve the auxiliary unconstrained problems with the precision controlled by the norm of violation of the Karush–Kuhn–Tucker conditions. The fixed steplength gradient projections are used to expand the active set. The basic scheme of MPGP is presented as Algorithm 5.6. The algorithm is proved to have an R-linear rate of convergence bounded in terms of the extreme eigenvalues of the Hessian matrix.

The *MPRGP* (modified proportioning with reduced gradient projections) algorithm of Sect. 5.8 is closely related to the MPGP algorithm, only the gradient projection step is replaced by the projection of the free gradient. The basic MPRGP scheme is presented as Algorithm 5.7. The R-linear rate of convergence is proved not only for the decrease of the cost function, but also for the norm of the projected gradient. The finite termination property is proved even for the problems with a dual degenerate solution.

The performance of MPGP and MPRGP can be improved by the *preconditioning* described in Sect. 5.10. The preconditioning in face improves the solution of the auxiliary unconstrained problems, while the preconditioning by the conjugate projector improves the convergence of the whole staff, including the nonlinear steps. The *monotonic MPRGP* and *semimonotonic MPRGP* algorithms which accept unfeasible iterations are described in Sect. 5.9.3.

5.1 Review of Alternative Methods

Before describing in detail the active set-based methods, let us briefly review alternative methods for the solution of the bound constrained problem (5.1).

Closely related to the active set strategy, various finite algorithms try to find $\mathbf{x} \in \mathbb{R}^n$ which solves the symmetric positive definite *LCP* (Linear Complementarity Problems)

$$\mathbf{g} = \mathsf{A}\mathbf{x} - \mathbf{b}, \quad \mathbf{x} \geq \mathbf{o}, \quad \mathbf{g} \geq \mathbf{o}, \quad \mathbf{x}^T\mathbf{g} = 0.$$

The LCP is equivalent to the minimization problem (5.1) with $\boldsymbol{\ell} = \mathbf{o}$. The algorithms are called finite as they find the solution in a finite number of steps; their analysis is based on the arguments of combinatorial nature. The most popular LCP solvers are probably *Lemke's algorithm* and *principal pivoting algorithm*, which reduce the LCP to the solution of a sequence of systems of linear equations in a way which is similar to the simplex method in linear programming. The solution of the auxiliary systems is typically implemented by LU-decompositions that are usually implemented by a rank one update. The result of the trial solve is used to improve a current approximation in order to reduce some characteristics of violation of the LCP conditions. These algorithms typically do not refer to the background minimization problems. The algorithms can be useful especially for more general LCP problems not considered here; see Cottle, Pang, and Stone [29].

Apart from the feasible active set methods presented in this chapter, it is possible to consider their unfeasible variants. For example, Kunisch and Rendl [139] proposed an iterative primal–dual algorithm which maintains the first-order optimality and complementarity conditions associated with (5.1) only; the feasibility is enforced by the update of the active set. The unfeasible methods are closely related to the *semismooth Newton method* applied to

$$\boldsymbol{\Phi}(\mathbf{x}) = \mathbf{o}, \quad \boldsymbol{\Phi}(\mathbf{x}) = \alpha^{-1}\big(\mathbf{x} - P_{\Omega_B}\big(\mathbf{x} - \alpha\nabla f(\mathbf{x})\big)\big), \quad \alpha > 0.$$

Hintermüller, Ito, and Kunisch [118] and Hintermüller, Kovtumenko, and Kunisch [119] describe the primal–dual semismooth Newton methods.

The bound constraints can be treated efficiently by the *interior point method*, which approximately minimizes the cost function modified by the parameterized barrier functions using Newton's method. The strong feature of the interior point methods is their capability to take into account all constraints, not only the active ones, at the cost of dealing with ill-conditioned problems. The performance of the interior point methods can exploit the sparsity pattern of the Hessian matrix A in the solution of auxiliary problems. There is a vast literature on this subject, see, e.g., the book by Wright [182] or the review paper by Forsgren, Gill, and Wright [90].

It is also possible to use the trust region-type methods that were developed to stabilize convergence of the Newton-type methods. We refer to Coleman and Lin [24, 25] for more details.

5.2 KKT Conditions and Related Inequalities

Since Ω_B is closed and convex and f is assumed to be strictly convex, the solution $\widehat{\mathbf{x}}$ of problem (5.1) exists and is necessarily unique by Proposition 2.5(i). Here we introduce some definitions and notations that enable us to exploit the special form of the KKT conditions in development of our algorithms. The KKT conditions fully determine the unique solution of (5.1).

By Proposition 2.18, the KKT conditions read

$$\mathsf{A}\widehat{\mathbf{x}} - \mathbf{b} \geq \mathbf{o} \quad \text{and} \quad (\mathsf{A}\widehat{\mathbf{x}} - \mathbf{b})^T(\widehat{\mathbf{x}} - \boldsymbol{\ell}) = 0,$$

or componentwise

$$\widehat{x}_i = \ell_i \;\Rightarrow\; \widehat{g}_i \geq 0 \quad \text{and} \quad \widehat{x}_i > \ell_i \;\Rightarrow\; \widehat{g}_i = 0, \quad i = 1, \ldots, n, \tag{5.2}$$

where $\widehat{g}_i = [\mathsf{A}\widehat{\mathbf{x}} - \mathbf{b}]_i$. It may be observed that $\widehat{g}_i$ are the components of the vector of Lagrange multipliers for the bound constraints.

The KKT conditions (5.2) determine three important subsets of the set $\mathcal{N} = \{1, 2, \ldots, n\}$ of all indices. The set of all indices for which $x_i = \ell_i$ is called an *active set* of $\mathbf{x}$. We denote it by $\mathcal{A}(\mathbf{x})$, so

$$\mathcal{A}(\mathbf{x}) = \{i \in \mathcal{N} : \; x_i = \ell_i\}.$$

Its complement

$$\mathcal{F}(\mathbf{x}) = \{i \in \mathcal{N} : \; x_i \neq \ell_i\}$$

and subsets

$$\mathcal{B}(\mathbf{x}) = \{i \in \mathcal{N} : \; x_i = \ell_i \text{ and } g_i > 0\}, \quad \mathcal{B}_0(\mathbf{x}) = \{i \in \mathcal{N} : \; x_i = \ell_i \text{ and } g_i \geq 0\}$$

are called a *free set*, a *binding set*, and a *weakly binding set*, respectively. Thus we can rewrite the KKT conditions in the form

$$\mathbf{g}_{\mathcal{A}}(\widehat{\mathbf{x}}) \geq \mathbf{o}_{\mathcal{A}} \quad \text{and} \quad \mathbf{g}_{\mathcal{F}}(\widehat{\mathbf{x}}) = \mathbf{o}_{\mathcal{F}}.$$

Using the subsets of $\mathcal{N}$, we can decompose the part of the gradient $\mathbf{g}(\mathbf{x}) = \mathsf{A}\mathbf{x} - \mathbf{b}$ which violates the KKT conditions into the *free gradient* $\boldsymbol{\varphi}$ and the *chopped gradient* $\boldsymbol{\beta}$ that are defined by

$$\begin{aligned} &\varphi_i(\mathbf{x}) = g_i(\mathbf{x}) \text{ for } i \in \mathcal{F}(\mathbf{x}), \quad && \varphi_i(\mathbf{x}) = 0 \text{ for } i \in \mathcal{A}(\mathbf{x}), \\ &\beta_i(\mathbf{x}) = 0 \text{ for } i \in \mathcal{F}(\mathbf{x}), && \beta_i(\mathbf{x}) = g_i^-(\mathbf{x}) \text{ for } i \in \mathcal{A}(\mathbf{x}), \end{aligned}$$

where we have used the notation $g_i^- = \min\{g_i, 0\}$. Introducing the *projected gradient*

$$\mathbf{g}^P(\mathbf{x}) = \boldsymbol{\varphi}(\mathbf{x}) + \boldsymbol{\beta}(\mathbf{x}),$$

we can write the Karush–Kuhn–Tucker conditions (5.2) conveniently as

$$\mathbf{g}^P(\mathbf{x}) = \mathbf{o}. \tag{5.3}$$

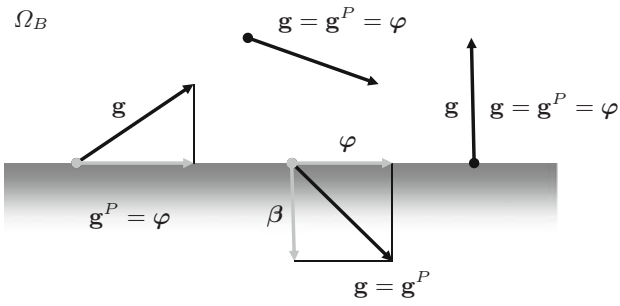

Fig. 5.1. Gradient splitting

Obviously $\boldsymbol{\beta}(\mathbf{x})$ and $\boldsymbol{\varphi}(\mathbf{x})$ are orthogonal and $-\boldsymbol{\beta}(\mathbf{x})$ and $-\boldsymbol{\varphi}(\mathbf{x})$ are feasible decrease directions of f at $\mathbf{x}$. See also Fig. 5.1.

If the dimension n of the bound constrained minimization problem (5.1) is large, it can be too ambitious to look for a solution which satisfies the gradient condition (5.3) exactly. A natural idea is to consider the weaker condition

$$\|\mathbf{g}^P(\mathbf{x})\| \le \varepsilon, \tag{5.4}$$

but to require that the feasibility condition $\mathbf{x} \in \Omega_B$ is satisfied exactly. Notice that we are not able to check directly that we are near the solution as we do not know it, but we can easily evaluate (5.4). Thus the typical "solution" returned by iterative solvers is just $\mathbf{x}$ that satisfies the condition (5.4) with a small ε. The following lemma guarantees that any feasible vector $\mathbf{x}$ which satisfies (5.4) is near the solution.

Lemma 5.1. *Let $\widehat{\mathbf{x}}$ be the solution of (5.1) with a positive definite A and let $\mathbf{g}^P = \mathbf{g}^P(\mathbf{x})$ denote the projected gradient at $\mathbf{x} \in \Omega_B$. Then*

$$\|\mathbf{x} - \widehat{\mathbf{x}}\|_{\mathsf{A}}^2 \le 2\big(f(\mathbf{x}) - f(\widehat{\mathbf{x}})\big) \le \|\mathbf{g}^P\|_{\mathsf{A}^{-1}} \le \lambda_{\min}^{-1} \|\mathbf{g}^P\|, \tag{5.5}$$

where $\lambda_{\min}$ denotes the smallest eigenvalue of A.

Proof. Let $\widehat{\mathcal{A}}$, $\widehat{\mathcal{F}}$, and $\widehat{\mathbf{g}}$ denote the active set, free set, and the gradient in the solution, respectively. Since $[\mathbf{x} - \widehat{\mathbf{x}}]_{\widehat{\mathcal{A}}} \ge \mathbf{o}_{\widehat{\mathcal{A}}}$, $\widehat{\mathbf{g}}_{\widehat{\mathcal{F}}} = \mathbf{o}_{\widehat{\mathcal{F}}}$, and $\widehat{\mathbf{g}} \ge \mathbf{o}$, we get

$$\begin{aligned} f(\mathbf{x}) - f(\widehat{\mathbf{x}}) &= \widehat{\mathbf{g}}^T(\mathbf{x} - \widehat{\mathbf{x}}) + \frac{1}{2}(\mathbf{x} - \widehat{\mathbf{x}})^T \mathsf{A}(\mathbf{x} - \widehat{\mathbf{x}}) \\ &= \widehat{\mathbf{g}}_{\widehat{\mathcal{A}}}^T [\mathbf{x} - \widehat{\mathbf{x}}]_{\widehat{\mathcal{A}}} + \frac{1}{2}\|\mathbf{x} - \widehat{\mathbf{x}}\|_{\mathsf{A}}^2 \ge \frac{1}{2}\|\mathbf{x} - \widehat{\mathbf{x}}\|_{\mathsf{A}}^2. \end{aligned}$$

This proves the left inequality of (5.5).

To prove the middle inequality, let $\mathcal{A} = \mathcal{A}(\mathbf{x})$ and $\mathcal{F} = \mathcal{F}(\mathbf{x})$ denote the active set and the free set of $\mathbf{x} \in \Omega_B$, respectively. Since

$$\mathbf{g}_{\mathcal{F}}^P = \mathbf{g}_{\mathcal{F}}, \quad [\widehat{\mathbf{x}} - \mathbf{x}]_{\mathcal{A}} \geq \mathbf{o}_{\mathcal{A}}, \quad \mathbf{g} = (\mathbf{g} - \mathbf{g}^P) + \mathbf{g}^P, \quad \text{and} \quad \mathbf{g} - \mathbf{g}^P \geq \mathbf{o},$$

we get

$$\begin{aligned}
0 &\geq 2\big(f(\widehat{\mathbf{x}}) - f(\mathbf{x})\big) = \|\widehat{\mathbf{x}} - \mathbf{x}\|_{\mathsf{A}}^2 + 2\mathbf{g}^T(\widehat{\mathbf{x}} - \mathbf{x}) \\
&= \|\widehat{\mathbf{x}} - \mathbf{x}\|_{\mathsf{A}}^2 + 2\left(\mathbf{g} - \mathbf{g}^P\right)^T(\widehat{\mathbf{x}} - \mathbf{x}) + 2\left(\mathbf{g}^P\right)^T(\widehat{\mathbf{x}} - \mathbf{x}) \\
&= \|\widehat{\mathbf{x}} - \mathbf{x}\|_{\mathsf{A}}^2 + 2\left[\mathbf{g} - \mathbf{g}^P\right]_{\mathcal{A}}^T[\widehat{\mathbf{x}} - \mathbf{x}]_{\mathcal{A}} + 2\left(\mathbf{g}^P\right)^T(\widehat{\mathbf{x}} - \mathbf{x}) \\
&\geq \|\widehat{\mathbf{x}} - \mathbf{x}\|_{\mathsf{A}}^2 + 2\left(\mathbf{g}^P\right)^T(\widehat{\mathbf{x}} - \mathbf{x}) \\
&\geq 2\left(\min_{\mathbf{y}\in\mathbb{R}^n} \frac{1}{2}\mathbf{y}^T\mathsf{A}\mathbf{y} + \left(\mathbf{g}^P\right)^T\mathbf{y}\right) = -(\mathbf{g}^P)^T\mathsf{A}^{-1}\mathbf{g}^P.
\end{aligned}$$

We used (2.11) in the last step. The middle inequality and the right inequality of (5.5) now follow by simple manipulations and (1.24), respectively. □

5.3 The Working Set Method with Exact Solutions

The basic idea of the *working set method*, or, as it is often called less correctly, the *active set method*, is to reduce the solution of an inequality constrained problem to the solution of a sequence of auxiliary equality constrained problems which are defined by a subset of the set $\mathcal{N} = \{1, \ldots, n\}$ of all indices of the constraints. This task would be very simple if we knew in advance which inequality constraints are active in the solution, as we could just replace the relevant inequalities by equalities, ignore the other inequalities, and solve the resulting equality constrained problem. As this is usually not the case, the working set method starts by making a guess which inequality constraints will be active in the solution, and if this guess turns out to be incorrect, it exploits the gradient and Lagrange multiplier information obtained by the trial minimization to define the next prediction.

5.3.1 Auxiliary Problems

If the working set method is applied to (5.1), it exploits the auxiliary equality constrained problems

$$\min_{\mathbf{y}\in\mathcal{W}_{\mathcal{I}}} f(\mathbf{y}), \tag{5.6}$$

where $\mathcal{I} \subseteq \mathcal{N}$ denotes the set of indices of bounds ℓ_i that are predicted to be active in the solution, and

$$\mathcal{W}_{\mathcal{I}} = \{\mathbf{y} : \ y_i = \ell_i, \ i \in \mathcal{I}\}.$$

The predicted set $\mathcal{I}$ of active bounds and $\mathcal{W}_{\mathcal{I}}$ are known as the *working set* and the *working face*, respectively. Since f is assumed to be strictly convex and $\mathcal{W}_{\mathcal{I}}$ is closed and convex, it follows by Proposition 2.5 that the auxiliary problem (5.6) has a unique solution $\widehat{\mathbf{y}}$.

Now observe that the equality constrained problem (5.6) can be reduced to an unconstrained problem in y_j, $j \notin \mathcal{I}$. To see its explicit form in the nontrivial cases $\mathcal{W}_{\mathcal{I}} \neq \{\boldsymbol{\ell}\}$ and $\mathcal{W}_{\mathcal{I}} \neq \mathbb{R}^n$, assume that $\emptyset \subsetneq \mathcal{J} \subsetneq \mathcal{N}$, and denote $\mathcal{J} = \mathcal{N} \setminus \mathcal{I}$, so that, after possibly rearranging the indices, we can write

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_{\mathcal{I}} \\ \mathbf{y}_{\mathcal{J}} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \mathbf{b}_{\mathcal{I}} \\ \mathbf{b}_{\mathcal{J}} \end{bmatrix}, \quad \text{and} \quad \mathsf{A} = \begin{bmatrix} \mathsf{A}_{\mathcal{I}\mathcal{I}} & \mathsf{A}_{\mathcal{I}\mathcal{J}} \\ \mathsf{A}_{\mathcal{J}\mathcal{I}} & \mathsf{A}_{\mathcal{J}\mathcal{J}} \end{bmatrix}. \tag{5.7}$$

Thus for any $\mathbf{y} \in \mathbb{R}^n$

$$f(\mathbf{y}) = \frac{1}{2}\mathbf{y}_{\mathcal{J}}^T \mathsf{A}_{\mathcal{J}\mathcal{J}} \mathbf{y}_{\mathcal{J}} + \mathbf{y}_{\mathcal{J}}^T \mathsf{A}_{\mathcal{J}\mathcal{I}} \mathbf{y}_{\mathcal{I}} + \frac{1}{2}\mathbf{y}_{\mathcal{I}}^T \mathsf{A}_{\mathcal{I}\mathcal{I}} \mathbf{y}_{\mathcal{I}} - \mathbf{y}_{\mathcal{J}}^T \mathbf{b}_{\mathcal{J}} - \mathbf{y}_{\mathcal{I}}^T \mathbf{b}_{\mathcal{I}}.$$

Since $\mathbf{y} \in \mathcal{W}_{\mathcal{I}}$ if and only if $\mathbf{y}_{\mathcal{I}} = \boldsymbol{\ell}_{\mathcal{I}}$, we have for any $\mathbf{y} \in \mathcal{W}_{\mathcal{I}}$

$$f(\mathbf{y}) = f_{\mathcal{J}}(\mathbf{y}_{\mathcal{J}}) = \frac{1}{2}\mathbf{y}_{\mathcal{J}}^T \mathsf{A}_{\mathcal{J}\mathcal{J}} \mathbf{y}_{\mathcal{J}} - \mathbf{y}_{\mathcal{J}}^T (\mathbf{b}_{\mathcal{J}} - \mathsf{A}_{\mathcal{J}\mathcal{I}} \boldsymbol{\ell}_{\mathcal{I}}) + \frac{1}{2}\boldsymbol{\ell}_{\mathcal{I}}^T \mathsf{A}_{\mathcal{I}\mathcal{I}} \boldsymbol{\ell}_{\mathcal{I}} - \mathbf{b}_{\mathcal{I}}^T \boldsymbol{\ell}_{\mathcal{I}}.$$

Thus the solution $\widehat{\mathbf{y}}$ of (5.6) has the components $\widehat{\mathbf{y}}_{\mathcal{I}} = \boldsymbol{\ell}_{\mathcal{I}}$ and

$$\widehat{\mathbf{y}}_{\mathcal{J}} = \arg \min_{\mathbf{y}_{\mathcal{J}} \in \mathbb{R}^m} f_{\mathcal{J}}(\mathbf{y}_{\mathcal{J}}). \tag{5.8}$$

Since

$$\nabla f_{\mathcal{J}}(\mathbf{y}_{\mathcal{J}}) = \mathsf{A}_{\mathcal{J}\mathcal{J}} \mathbf{y}_{\mathcal{J}} - (\mathbf{b}_{\mathcal{J}} - \mathsf{A}_{\mathcal{J}\mathcal{I}} \boldsymbol{\ell}_{\mathcal{I}})$$

and $\nabla f_{\mathcal{J}}(\widehat{\mathbf{y}}_{\mathcal{J}}) = \mathbf{o}$, we get that $\widehat{\mathbf{y}}_{\mathcal{J}}$ satisfies

$$\mathsf{A}_{\mathcal{J}\mathcal{J}} \widehat{\mathbf{y}}_{\mathcal{J}} = \mathbf{b}_{\mathcal{J}} - \mathsf{A}_{\mathcal{J}\mathcal{I}} \boldsymbol{\ell}_{\mathcal{I}}. \tag{5.9}$$

We can check easily that (5.9) has a unique solution. Indeed, since $\mathsf{A}_{\mathcal{J}\mathcal{J}}$ is a submatrix of a positive definite matrix A, we get by Cauchy's interlacing inequalities (1.21) that $\mathsf{A}_{\mathcal{J}\mathcal{J}}$ is also positive definite. Alternatively, we can verify directly that $\mathsf{A}_{\mathcal{J}\mathcal{J}}$ is positive definite by observing that if $\mathbf{y}$ has the components $\mathbf{y}_{\mathcal{I}} = \mathbf{o}$ and $\mathbf{y}_{\mathcal{J}} \neq \mathbf{o}$, then $\mathbf{y} \neq \mathbf{o}$ and

$$\mathbf{y}_{\mathcal{J}}^T \mathsf{A}_{\mathcal{J}\mathcal{J}} \mathbf{y}_{\mathcal{J}} = \mathbf{y}^T \mathsf{A} \mathbf{y} > 0.$$

5.3.2 Algorithm

The *working set method with exact solutions* of auxiliary problems starts from an arbitrary $\mathbf{x}^0 \in \Omega_B$ and $\mathcal{I}^0 = \mathcal{B}_0(\mathbf{x}^0)$. Assuming that $\mathbf{x}^k$ is known, we first check if $\mathbf{x}^k$ is the solution of (5.1) by evaluating the KKT conditions

$$\mathbf{g}^P(\mathbf{x}^k) = \boldsymbol{\beta}(\mathbf{x}^k) + \boldsymbol{\varphi}(\mathbf{x}^k) = \mathbf{o}.$$

If this is not the case, we find the solution $\widehat{\mathbf{y}}$ of the auxiliary problem (5.6) by solving (5.9). There are two possibilities.

If $\widehat{\mathbf{y}} \in \Omega_B$, then we define the next iteration by the *feasible step*

$$\mathbf{x}^{k+1} = \widehat{\mathbf{y}}$$

and set $\mathcal{I}^{k+1} = \mathcal{B}_0(\mathbf{x}^{k+1})$. Notice that $f(\mathbf{x}^{k+1}) < f(\mathbf{x}^k)$ as $-\mathbf{g}^P(\mathbf{x}^k)$ is a feasible decrease direction of f at $\mathbf{x}^k$ with respect to $\mathcal{W}_I$.

In the other case, we define $\mathbf{x}^{k+1}$ by an *expansion step* so that

$$f(\mathbf{x}^{k+1}) \le f(\mathbf{x}^k) \quad \text{and} \quad \mathcal{A}(\mathbf{x}^{k+1}) \supsetneq \mathcal{I}^k, \tag{5.10}$$

and then set $\mathcal{I}^{k+1} = \mathcal{A}(\mathbf{x}^{k+1})$. The basic working set algorithm in the form that is convenient for analysis reads as follows.

Algorithm 5.1. The working set method with exact solutions.

Given a symmetric positive definite matrix $\mathsf{A} \in \mathbb{R}^{n\times n}$ *and n-vectors* $\mathbf{b}$, $\boldsymbol{\ell}$.
Step 0. {*Initialization.*}
  Choose $\mathbf{x}^0 \in \Omega_B$, *set* $\mathcal{I}^0 = \mathcal{B}_0(\mathbf{x}^0)$, $k = 0$
  while $\|\mathbf{g}^P(\mathbf{x}^k)\| > 0$
Step 1. {*Minimization in face* $\mathcal{W}_{\mathcal{I}^k}$. }
    $\widehat{\mathbf{y}} = \arg\ \min_{\mathbf{y}\in\mathcal{W}_{\mathcal{I}^k}} f(\mathbf{y})$
    if $\widehat{\mathbf{y}} \in \Omega_B$
Step 2. {*Feasible step.*}
      $\mathbf{x}^{k+1} = \widehat{\mathbf{y}}$
      $\mathcal{I}^{k+1} = \mathcal{B}_0(\mathbf{x}^{k+1})$
    else
Step 3. {*Expansion step.*}
      Set $\mathbf{x}^{k+1}$ *so that* $f(\mathbf{x}^{k+1}) \le f(\mathbf{x}^k)$ *and* $\mathcal{A}(\mathbf{x}^{k+1}) \supsetneq \mathcal{I}^k$
      $\mathcal{I}^{k+1} = \mathcal{A}(\mathbf{x}^{k+1})$
    end *if*
    $k = k + 1$
  end *while*
Step 4. {*Return solution.*}
    $\widehat{\mathbf{x}} = \mathbf{x}^k$

To implement the algorithm, we should specify the expansion step in more detail. For example, if $\mathbf{x}^k \in \Omega_B$ and

$$\mathbf{d} = \mathbf{x}^k - \widehat{\mathbf{y}},$$

we can observe that $-\mathbf{d}$ is a feasible decrease direction and that $f(\mathbf{x}^k - \alpha\mathbf{d})$ is a decreasing function of α for $\alpha \in [0, 1]$. Thus we can look for $\mathbf{x}^{k+1}$ in the form $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha\mathbf{d}$, $\alpha \in (0, 1]$. A possible choice of α is given by

$$\alpha_f = \arg \min_{\alpha\in(0,1]} \{f(\mathbf{x}^k - \alpha\mathbf{d}) :\ \mathbf{x}^k - \alpha\mathbf{d} \in \Omega_B\}, \tag{5.11}$$

which can be evaluated by using

$$\alpha_f = \min\{\alpha_m, 1\}, \quad \alpha_m = \min\{(x_i^k - \ell_i)/d_i : \ d_i > 0, \ i \in \mathcal{N}\}. \tag{5.12}$$

See also Fig. 5.2. Notice that if $\widehat{\mathbf{y}} \notin \Omega_B$, then the steplength α_f necessarily results in the expansion of the working set, typically by one index.

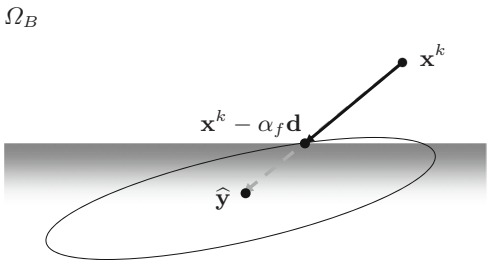

Fig. 5.2. Feasible steplength

This limitation may be overcome if we set $\mathbf{y} = \mathbf{x}^k - \alpha_f \mathbf{d}$ and define

$$\mathbf{x}^{k+1} = P_{\Omega_B}(\mathbf{y} - \alpha_p \mathbf{g}), \quad \alpha_p = \arg\min_{\alpha \geq 0} f\left(P_{\Omega_B}(\mathbf{y} - \alpha \mathbf{g})\right), \quad \mathbf{g} = \nabla f(\mathbf{y}),$$

where P_{Ω_B} is the Euclidean projection of Sect. 2.3.4. We prefer to use the gradient path, as the gradient defines a better local model of f than $\mathbf{d}$, though $-\mathbf{d}$ is the best global direction for minimization in the current working set. Figure 5.3 shows that α_f may be the best steplength for $\mathbf{d}$!

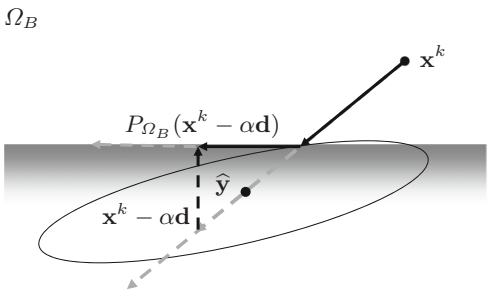

Fig. 5.3. Projected best unconstrained decrease path

To approximate α_p effectively, it is useful to notice that $f\left(P_{\Omega_B}(\mathbf{y}-\alpha\mathbf{g})\right)$ is a piecewise quadratic function because $P_{\Omega_B}(\mathbf{y}-\alpha\mathbf{g})$ is a linear mapping on any interval on which the active set of $P_{\Omega_B}(\mathbf{y}-\alpha\mathbf{g})$ is unchanged. We refer interested readers to Moré and Toraldo [153], Nocedal and Wright [155, Sect. 16.4], or to the discussion of the projected-gradient path in Conn, Gould, and Toint [28, Sect. 12.1.3]. We can also apply the fixed steplength reduced gradient projection which is described in Sect. 5.6.

The algorithm assumes by default that Step 1 is carried out by a direct method such as a matrix factorization, in which economies are possible by updating rather than recomputing the factorizations to account for gradual changes in the working set.

5.3.3 Finite Termination

The analysis of the working set method can be based on the following finite termination property.

Theorem 5.2. *Let Algorithm 5.1 be applied to find the solution* $\widehat{\mathbf{x}}$ *of (5.1) starting from* $\mathbf{x}^0 \in \Omega_B$. *Then there is* k *such that* $\mathbf{x}^k = \widehat{\mathbf{x}}$.

Proof. Since each expansion step adds at least one index into the working set, and the number of indices in the working set cannot exceed n, it follows that there are at most n consecutive expansion steps. Thus after each consecutive series of expansion steps, the algorithm either finds the solution of (5.1) and we are finished, or generates the next iterate, a feasible minimizer on the current face, by a feasible step. However, since $f(\mathbf{x}^k)$ is a nonincreasing sequence such that

$$f(\mathbf{x}^{k+1}) < f(\mathbf{x}^k)$$

whenever $\mathbf{x}^{k+1}$ is generated by a feasible step, it follows that no working set corresponding to an iterate generated by the feasible step can reappear. The number of different working sets being finite, we conclude that the working set method exploiting the exact solutions of auxiliary problems finds the solution of (5.1) in a finite number of steps. □

Since the number of different working sets is 2^n and there can be at most n expanding steps for each feasible step, the proof of Theorem 5.2 gives that the number N of iterations of the working set method with exact solution is bounded by

$$\overline{N} = n2^n. \tag{5.13}$$

This bound is very pessimistic and gives a poor theoretical support for practical computations, especially if we take into account the high cost of the iterations. The bound can be essentially improved for special problems. For example, if $\mathbf{x}^0 = \boldsymbol{\ell} = \mathbf{o}$ and the Hessian A of f is an M-matrix, then it is possible to show that Algorithm 5.1 generates only feasible steps and finds the solution in a number of iterations that does not exceed $n-p$, where p is the number of positive entries in $\mathbf{b}$. For more details see Diamond [32].

5.4 Polyak's Algorithm

If the auxiliary problems (5.6) are solved by the conjugate gradient method, it seems reasonable not to wait with the test of feasibility until their solution is found, but to modify the working set whenever unfeasible CG iteration is generated. This observation was enhanced in the *Polyak algorithm* [159], the starting point of our development of in a sense optimal algorithms.

5.4.1 Basic Algorithm

The new ingredient of the Polyak algorithm is that the minimization in face is replaced by a sequence of the *conjugate gradient steps* defined by

$$\mathbf{x}^k = \mathbf{x}^{k-1} - \alpha_{cg}\mathbf{p}^k, \tag{5.14}$$

where $\mathbf{p}^k$ denotes the recurrently constructed conjugate direction introduced in Sect. 3.2, and α_{cg} is the minimizer of $f(\mathbf{x}^{k-1} - \xi\mathbf{p}^k)$. The recurrence starts (or restarts) from $\mathbf{p}^{s+1} = \boldsymbol{\varphi}(\mathbf{x}^s)$ whenever $\mathbf{x}^s$ is generated by the expansion step or $s = 0$. If $\mathbf{p}^k$ is known, then $\mathbf{p}^{k+1}$ is given by the formulae

$$\mathbf{p}^{k+1} = \boldsymbol{\varphi}(\mathbf{x}^k) - \beta\mathbf{p}^k \quad \text{and} \quad \beta = \boldsymbol{\varphi}(\mathbf{x}^k)^T\mathsf{A}\mathbf{p}^k/(\mathbf{p}^k)^T\mathsf{A}\mathbf{p}^k, \tag{5.15}$$

obtained by specialization of those introduced in Sect. 3.2. Let us recall that the conjugate directions $\mathbf{p}^{s+1}, \ldots, \mathbf{p}^k$ that are generated by the recurrence (5.15) from the restart $\mathbf{x}^s$ are A-orthogonal, i.e., $(\mathbf{p}^i)^T\mathsf{A}\mathbf{p}^j = 0$ for any $i, j \in \{s+1, \ldots, k\}$, $i \neq j$. Using the arguments of Sect. 3.1, it follows that

$$f(\mathbf{x}^k) = \min\left\{f(\mathbf{x}^s + \mathbf{y}) : \ \mathbf{y} \in \mathrm{Span}\{\mathbf{p}^{s+1}, \ldots, \mathbf{p}^k\}\right\}. \tag{5.16}$$

The Polyak algorithm starts from an arbitrary feasible $\mathbf{x}^0$ by assigning $\mathcal{I}^0 = \mathcal{B}_0(\mathbf{x}^0)$ and initializing of the conjugate gradient loop (for details see Algorithm 5.2 or Sect. 3.2) for the minimization in $\mathcal{W}_{\mathcal{I}^0}$. Assuming that $\mathbf{x}^k$ is known, we first check if $\mathbf{x}^k$ solves either (5.1) or the auxiliary problem (5.6) by testing $\mathbf{g}^P(\mathbf{x}^k) = \mathbf{o}$ and $\boldsymbol{\varphi}(\mathbf{x}^k) = \mathbf{o}$, respectively. If $\mathbf{g}^P(\mathbf{x}^k) = \mathbf{o}$, we are finished; if $\boldsymbol{\varphi}(\mathbf{x}^k) = \mathbf{o}$, we reduce the working set to $\mathcal{I}^k = \mathcal{B}_0(\mathbf{x}^k)$ and initialize the conjugate gradient loop.

If the tests fail, we use the conjugate gradient step to define the trial iteration $\mathbf{y} = \mathbf{x}^k - \alpha_{cg}\mathbf{p}^{k+1}$. There are two possibilities. If $\mathbf{y}$ is feasible, then we set $\mathbf{x}^{k+1} = \mathbf{y}$. Otherwise we evaluate the feasible steplength by

$$\alpha_f = \arg\min_{\alpha\in(0,\alpha_{cg}]}\{f(\mathbf{x}^k - \alpha\mathbf{p}^{k+1}) : \ \mathbf{x}^k - \alpha\mathbf{p}^{k+1} \in \Omega_B\}, \tag{5.17}$$

set $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_f\mathbf{p}^{k+1}$, expand the working set by $\mathcal{I}^{k+1} = \mathcal{A}(\mathbf{x}^{k+1})$, and finally initialize the new conjugate gradient loop.

The basic Polyak algorithm for the solution of strictly convex bound constrained quadratic programming problems takes the form shown by the following algorithm, where we omitted the indices of the vectors that are not referred to in what follows.

Algorithm 5.2. Polyak's algorithm.

Given a symmetric positive definite matrix $\mathsf{A} \in \mathbb{R}^{n\times n}$ *and n-vectors* $\mathbf{b}$, $\boldsymbol{\ell}$.
Step 0. {*Initialization.*}
  Choose $\mathbf{x}^0 \in \Omega_B$, *set* $\mathbf{g} = \mathsf{A}\mathbf{x}^0 - \mathbf{b}$, $\mathbf{p} = \mathbf{g}^P(\mathbf{x}^0)$, $k = 0$
  while $\|\mathbf{g}^P(\mathbf{x}^k)\| > 0$
    if $\|\boldsymbol{\varphi}(\mathbf{x}^k)\| > 0$
Step 1. {*Trial conjugate gradient step.*}
      $\alpha_{cg} = \mathbf{g}^T\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathbf{p}$, $\mathbf{y} = \mathbf{x}^k - \alpha_{cg}\mathbf{p}$
      $\alpha_f = \max\{\alpha : \mathbf{x}^k - \alpha\mathbf{p} \in \Omega_B\} = \min\{(x_i^k - \ell_i)/p_i : p_i > 0\}$
      if $\alpha_{cg} \le \alpha_f$
Step 2. {*Conjugate gradient step.*}
        $\mathbf{x}^{k+1} = \mathbf{y}$, $\mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathbf{p}$,
        $\beta = \boldsymbol{\varphi}(\mathbf{y})^T\mathsf{A}\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathbf{p}$, $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{y}) - \beta\mathbf{p}$
      else
Step 3. {*Expansion step.*}
        $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_f\mathbf{p}$, $\mathbf{g} = \mathbf{g} - \alpha_f\mathsf{A}\mathbf{p}$, $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{x}^{k+1})$
      end *if*
    else
Step 4. {*Leaving the face after finding the minimizer.*}
      $\mathbf{d} = \boldsymbol{\beta}(\mathbf{x}^k)$, $\alpha_{cg} = \mathbf{g}^T\mathbf{d}/\mathbf{d}^T\mathsf{A}\mathbf{d}$,
      $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_{cg}\mathbf{d}$, $\mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathbf{d}$, $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{x}^{k+1})$
    end *if*
    $k = k + 1$
  end *while*
Step 5. {*Return solution.*}
  $\widehat{\mathbf{x}} = \mathbf{x}^k$

Our description of Algorithm 5.2 does not use explicitly the working sets; the information about the current working set is enhanced in the iterates $\mathbf{x}^k$ and the conjugate directions $\mathbf{p}^k$. Let us recall that the properties of the unconstrained conjugate gradient method are summarized in Theorem 3.1.

5.4.2 Finite Termination

Theorem 5.3. *Let Polyak's Algorithm 5.2 be applied to find the solution* $\widehat{\mathbf{x}}$ *of (5.1) starting from* $\mathbf{x}^0 \in \Omega_B$. *Then there is k such that* $\mathbf{x}^k = \widehat{\mathbf{x}}$.

Proof. First notice that by Theorem 3.1, there can be at most n consecutive conjugate gradient iterations before the minimizer in a face is found. If we remove all the iterates that are generated by Step 2 except the minimizers in the faces examined by the algorithm, which are used in Step 4 to generate the next iteration in the expanded face, we are left with the iterates that can be generated also by an implementation of Algorithm 5.1. The statement then follows by Theorem 5.2. □

The arguments of Sect. 5.3.3 can be used to show that the number of iterations of Polyak's algorithm is bounded by

$$\overline{N} = n^2 2^n. \tag{5.18}$$

Let us emphasize here that this bound is very pessimistic and can be improved, at least for special problems.

5.4.3 Characteristics of Polyak's Algorithm

The Polyak algorithm suffers from several drawbacks. The first one is related to an unpleasant consequence of application of the reduced conjugate gradient step with the steplength α_f defined by (5.17). Since the working set is typically expanded by one index only, there is a little chance that the number of iterations will be small when many indices of the binding set of the solution do not belong to $\mathcal{B}(\mathbf{x}^0)$. Another drawback concerns the basic approach combining the conjugate gradient method, which is now understood as an efficient iterative method for approximate solution of linear systems [4, 106, 163], and the finite termination strategy, which is based on combinatorial reasoning that requires exact solution of the auxiliary problems. Finally, as we have seen above, the combinatorial arguments give extremely poor bound on the number of iterations that are necessary to find the solution of (5.1). Though the bound (5.18) does not depend on the conditioning of A, it is rather poor and does not indicate why the algorithm should be efficient for the solution of well-conditioned problems.

5.5 Inexact Polyak's Algorithm

In this section we consider the variants of Polyak's algorithm which accept inexact solutions of auxiliary problems, but preserve the finite termination property.

5.5.1 Looking Ahead and Estimate

Let us first show that it is not necessary to solve the auxiliary problems (5.6) exactly in order to preserve the finite termination property of the Polyak algorithm. The key observation is that if $\mathbf{x}^{k+1} \in \Omega_B$ satisfies

$$f(\mathbf{x}^{k+1}) < \min\{f(\mathbf{x}) : \ \mathbf{x} \in \mathcal{W}_{\mathcal{I}}\}, \tag{5.19}$$

then the working set $\mathcal{I}$ cannot appear again as long as $\{f(\mathbf{x}^k)\}$ is nonincreasing. We shall use this simple observation to define both the precision control test and reduction of the active set.

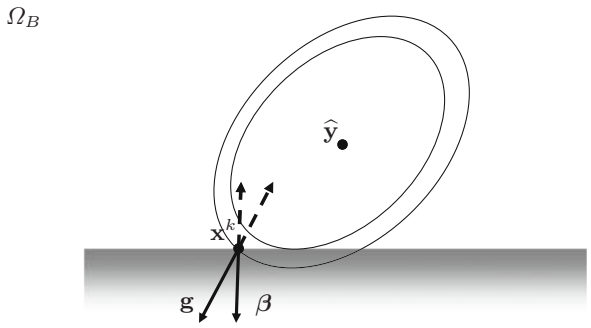

Fig. 5.4. Release directions at $\mathbf{x}^k$

Given $\mathbf{x}^k \in \mathcal{W}_{\mathcal{I}}$, we can try to find $\mathbf{x}^{k+1}$ which satisfies (5.19) in the form $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha\mathbf{d}$ with a given $\mathbf{d}$; if we are successful, we call $\mathbf{d}$ the *release direction of* $\mathcal{W}_{\mathcal{I}}$ *at* $\mathbf{x}^k$. The following lemma gives the conditions for $\mathbf{d}$, typically obtained from $\nabla f(\mathbf{x}^k)$ by reducing its components, to be a release direction. Such situation is depicted in Fig. 5.4 with $\mathbf{d} = \mathbf{g}(\mathbf{x}^k)$ and $\mathbf{d} = \boldsymbol{\beta}(\mathbf{x}^k)$.

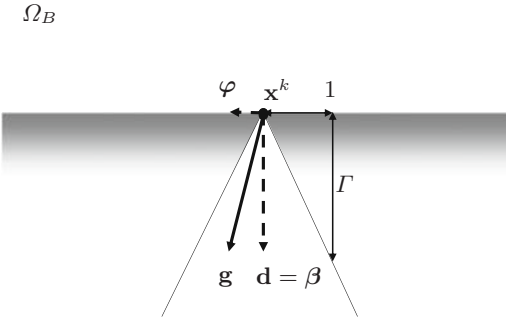

Fig. 5.5. The gradient and $\mathbf{d} = \boldsymbol{\beta}(\mathbf{x})$ that satisfy the release condition (5.20)

Lemma 5.4. *Let* $\mathcal{I} = \mathcal{A}(\mathbf{x})$ *and* $\Gamma \geq \kappa(\mathsf{A})^{1/2}$, *where* $\kappa(\mathsf{A})^{1/2}$ *denotes the spectral condition number of* A. *Denote* $\mathbf{g} = \nabla f(\mathbf{x})$ *and suppose that* $\mathbf{d}$ *satisfies*

$$\mathbf{g}^T\mathbf{d} \geq \|\mathbf{d}\|^2 \quad \textit{and} \quad \|\mathbf{d}\| > \Gamma\|\boldsymbol{\varphi}(\mathbf{x})\|. \tag{5.20}$$

Then the vector $\mathbf{y} = \mathbf{x} - \|\mathsf{A}\|^{-1}\mathbf{d}$ *satisfies*

$$f(\mathbf{y}) < \min\{f(\mathbf{x}) : \ \mathbf{x} \in \mathcal{W}_{\mathcal{I}}\}. \tag{5.21}$$

Proof. Let $\mathbf{x}$, Γ, and $\mathbf{d}$ satisfy the assumptions of Lemma 5.4 and notice that $\mathbf{g}^T\mathbf{d} \geq \|\mathbf{d}\|^2$ implies

$$f(\mathbf{y}) - f(\mathbf{x}) = \frac{1}{2}\|\mathsf{A}\|^{-2}\mathbf{d}^T\mathsf{A}\mathbf{d} - \|\mathsf{A}\|^{-1}\mathbf{d}^T\mathbf{g} \leq -\frac{1}{2}\|\mathsf{A}\|^{-1}\|\mathbf{d}\|^2. \tag{5.22}$$

Denoting $\mathcal{J} = \mathcal{F}(\mathbf{x})$, we have that $\|\mathbf{g}_{\mathcal{J}}\| = \|\boldsymbol{\varphi}(\mathbf{x})\|$ and by the assumptions

$$\|\mathbf{d}\|^2 > \kappa(\mathsf{A})\|\mathbf{g}_{\mathcal{J}}\|^2. \tag{5.23}$$

Substituting (5.23) into (5.22) then yields

$$f(\mathbf{y}) - f(\mathbf{x}) < -\frac{1}{2}\|\mathsf{A}^{-1}\|\|\mathbf{g}_{\mathcal{J}}\|^2. \tag{5.24}$$

Now denote by $\bar{\mathbf{x}}$ and $\bar{\mathbf{g}}$ the minimizer of $f(\mathbf{x})$ on $\mathcal{W}_{\mathcal{I}}$ and the corresponding gradient vector, respectively. Direct computations yield

$$f(\mathbf{x}) - f(\bar{\mathbf{x}}) = \frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}})^T\mathsf{A}(\mathbf{x} - \bar{\mathbf{x}}) + \bar{\mathbf{g}}^T(\mathbf{x} - \bar{\mathbf{x}}). \tag{5.25}$$

If we now rearrange the indices and take into account that

$$\bar{\mathbf{g}}_{\mathcal{J}} = \mathbf{o} \quad \text{and} \quad \mathbf{x}_{\mathcal{I}} = \bar{\mathbf{x}}_{\mathcal{I}},$$

we can further simplify the right-hand side of (5.25) to get

$$f(\mathbf{x}) - f(\bar{\mathbf{x}}) = \frac{1}{2}(\mathbf{x}_{\mathcal{J}} - \bar{\mathbf{x}}_{\mathcal{J}})^T\mathsf{A}_{\mathcal{J}\mathcal{J}}(\mathbf{x}_{\mathcal{J}} - \bar{\mathbf{x}}_{\mathcal{J}}). \tag{5.26}$$

To express $\mathbf{x}_{\mathcal{J}} - \bar{\mathbf{x}}_{\mathcal{J}}$ in terms of $\mathbf{g}_{\mathcal{J}}$, we can use the rearrangement (5.7) to get

$$\begin{bmatrix} \mathbf{g}_{\mathcal{I}} - \bar{\mathbf{g}}_{\mathcal{I}} \\ \mathbf{g}_{\mathcal{J}} \end{bmatrix} = \begin{bmatrix} \mathsf{A}_{\mathcal{I}\mathcal{I}} & \mathsf{A}_{\mathcal{I}\mathcal{J}} \\ \mathsf{A}_{\mathcal{J}\mathcal{I}} & \mathsf{A}_{\mathcal{J}\mathcal{J}} \end{bmatrix} \begin{bmatrix} \mathbf{o} \\ \mathbf{x}_{\mathcal{J}} - \bar{\mathbf{x}}_{\mathcal{J}} \end{bmatrix}. \tag{5.27}$$

In particular, since $\mathsf{A}_{\mathcal{J}\mathcal{J}}$ is also positive definite, it follows that

$$\mathbf{x}_{\mathcal{J}} - \bar{\mathbf{x}}_{\mathcal{J}} = \mathsf{A}_{\mathcal{J}\mathcal{J}}^{-1}\mathbf{g}_{\mathcal{J}}$$

and by (5.26)

$$f(\mathbf{x}) - f(\bar{\mathbf{x}}) = \frac{1}{2}\mathbf{g}_{\mathcal{J}}^T\mathsf{A}_{\mathcal{J}\mathcal{J}}^{-1}\mathbf{g}_{\mathcal{J}}. \tag{5.28}$$

Taking into account the interlacing properties of the spectra of principal submatrices of symmetric matrices (1.21), we get

$$\frac{1}{2}\mathbf{g}_{\mathcal{J}}^T\mathsf{A}_{\mathcal{J}\mathcal{J}}^{-1}\mathbf{g}_{\mathcal{J}} \leq \frac{1}{2}\|\mathsf{A}_{\mathcal{J}\mathcal{J}}^{-1}\|\|\mathbf{g}_{\mathcal{J}}\|^2 \leq \frac{1}{2}\|\mathsf{A}^{-1}\|\|\mathbf{g}_{\mathcal{J}}\|^2, \tag{5.29}$$

so that by (5.24) and (5.29)

$$\begin{aligned} f(\mathbf{y}) - f(\bar{\mathbf{x}}) &= \big(f(\mathbf{y}) - f(\mathbf{x})\big) + \big(f(\mathbf{x}) - f(\bar{\mathbf{x}})\big) \\ &< -\frac{1}{2}\|\mathsf{A}^{-1}\|\|\mathbf{g}_{\mathcal{J}}\|^2 + \frac{1}{2}\|\mathsf{A}^{-1}\|\|\mathbf{g}_{\mathcal{J}}\|^2 = 0. \end{aligned} \tag{5.30}$$

□

5.5.2 Looking Ahead Polyak's Algorithm

Using Lemma 5.4, we can now modify Polyak's algorithm so that it accepts approximate solution of the auxiliary problems and preserves its finite termination property. We only need to change the precision control of auxiliary problems. The *looking ahead Polyak algorithm* reads as follows.

Algorithm 5.3. Looking ahead Polyak's algorithm.

Given a symmetric positive definite matrix $\mathsf{A} \in \mathbb{R}^{n\times n}$, *n-vectors* $\mathbf{b},\ \boldsymbol{\ell}$.
Step 0. {*Initialization.*}
  Choose $\mathbf{x}^0 \in \Omega_B$, $\Gamma \geq \kappa(\mathsf{A})^{1/2}$, *set* $\mathbf{g} = \mathsf{A}\mathbf{x}^0 - \mathbf{b}$, $\mathbf{p} = \mathbf{g}^P(\mathbf{x}^0)$, $k = 0$
  while $\|\mathbf{g}^P(\mathbf{x}^k)\| > 0$
    if $\Gamma\|\boldsymbol{\varphi}(\mathbf{x}^k)\| \geq \|\boldsymbol{\beta}(\mathbf{x}^k)\|$
Step 1. {*Trial conjugate gradient step.*}
      $\alpha_{cg} = \mathbf{g}^T\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathbf{p},\ \mathbf{y} = \mathbf{x}^k - \alpha_{cg}\mathbf{p}$
      $\alpha_f = \max\{\alpha :\ \mathbf{x}^k - \alpha\mathbf{p} \in \Omega_B\} = \min\{(x_i^k - \ell_i)/p_i : p_i > 0\}$
      if $\alpha_{cg} \leq \alpha_f$
Step 2. {*Conjugate gradient step.*}
        $\mathbf{x}^{k+1} = \mathbf{y},\ \mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathbf{p},$
        $\beta = \boldsymbol{\varphi}(\mathbf{y})^T\mathsf{A}\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathbf{p},\ \mathbf{p} = \boldsymbol{\varphi}(\mathbf{y}) - \beta\mathbf{p}$
      else
Step 3. {*Expansion step.*}
        $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_f\mathbf{p},\ \mathbf{g} = \mathbf{g} - \alpha_f\mathsf{A}\mathbf{p},\ \mathbf{p} = \boldsymbol{\varphi}(\mathbf{x}^{k+1})$
      end *if*
    else
Step 4. {*Leaving the face in the release direction.*}
      $\mathbf{d} = \boldsymbol{\beta}(\mathbf{x}^k),\ \alpha_{cg} = \mathbf{g}^T\mathbf{d}/\mathbf{d}^T\mathsf{A}\mathbf{d},$
      $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_{cg}\mathbf{d},\ \mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathbf{d},\ \mathbf{p} = \boldsymbol{\varphi}(\mathbf{x}^{k+1})$
    end *if*
    $k = k + 1$
  end *while*
Step 5. {*Return solution.*}
  $\widehat{\mathbf{x}} = \mathbf{x}^k$

To see that Algorithm 5.3 deserves its name, denote $\mathbf{d} = \boldsymbol{\beta}(\mathbf{x}^k)$ and assume that

$$\mathbf{x}^k \in \Omega_B, \quad \|\boldsymbol{\beta}(\mathbf{x}^k)\| > \Gamma\|\boldsymbol{\varphi}(\mathbf{x}^k)\|, \quad \text{and} \quad \Gamma \geq \kappa(\mathsf{A})^{1/2}, \tag{5.31}$$

so that $\mathbf{d}$ and Γ satisfy the assumptions of Lemma 5.4. Observing that α_{cg} minimizes $f(\mathbf{x}^k - \alpha\mathbf{d})$ with respect to α, we get for $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_{cg}\mathbf{d}$ that

$$f(\mathbf{x}^{k+1}) \leq f(\mathbf{x}^k - \|\mathsf{A}\|^{-1}\mathbf{d}) < \min\{f(\mathbf{x}) :\ \mathbf{x} \in \mathcal{W}_{\mathcal{A}(\mathbf{x}^k)}\}.$$

Moreover, since $\mathbf{x}^k - \alpha\mathbf{d} \in \Omega_B$ for any $\alpha \geq 0$, we have $\mathbf{x}^{k+1} \in \Omega_B$. Thus the algorithm is able to recognize the face without the global solution before having a solution of the auxiliary problem, i.e., it "looks ahead".

The same reasoning as above can be carried out with $\mathbf{d} = \mathbf{g}^-(\mathbf{x}^k)$ or with some other nonzero vector $\mathbf{d}$ which satisfies the assumptions of Lemma 5.4. However, we found no significant evidence that there is a better choice than $\mathbf{d} = \boldsymbol{\beta}(\mathbf{x}^k)$.

5.5.3 Easy Re-release Polyak's Algorithm

We can consider the relations like

$$\Gamma\|\boldsymbol{\varphi}(\mathbf{x}^k)\| \geq \|\boldsymbol{\beta}(\mathbf{x}^k)\|$$

for any $\Gamma > 0$. A reasonable choice is $\Gamma = 1$, as it seems natural to leave the current face when the norm of the chopped gradient starts to dominate the violation of the KKT conditions. The following *easy re-release Polyak's algorithm* enhances this observation by means of Lemma 5.4.

Algorithm 5.4. Easy re-release Polyak's algorithm.

Given a symmetric positive definite matrix $\mathsf{A} \in \mathbb{R}^{n\times n}$, *n-vectors* $\mathbf{b}$, $\boldsymbol{\ell}$.
Step 0. {*Initialization.*}
  Choose $\mathbf{x}^0 \in \Omega_B$, $\Gamma_M \geq \kappa(\mathsf{A})^{1/2}$, $0 \leq \Gamma_m \leq \Gamma_M$, *set* $\Gamma = \Gamma_M$, $k = 0$
  $\mathbf{g} = \mathsf{A}\mathbf{x}^0 - \mathbf{b}$, $\mathbf{p} = \mathbf{g}^P(\mathbf{x}^0)$
  while $\|\mathbf{g}^P(\mathbf{x}^k)\| > 0$
    if $\Gamma\|\boldsymbol{\varphi}(\mathbf{x}^k)\| \geq \|\boldsymbol{\beta}(\mathbf{x}^k)\|$
Step 1. {*Trial conjugate gradient step.*}
      $\alpha_{cg} = \mathbf{g}^T\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathbf{p}$, $\mathbf{y} = \mathbf{x}^k - \alpha_{cg}\mathbf{p}$
      $\alpha_f = \max\{\alpha : \mathbf{x}^k - \alpha\mathbf{p} \in \Omega_B\} = \min\{(x_i^k - \ell_i)/p_i : p_i > 0\}$
      if $\alpha_{cg} \leq \alpha_f$
Step 2. {*Conjugate gradient step.*}
        $\mathbf{x}^{k+1} = \mathbf{y}$, $\mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathbf{p}$,
        $\beta = \boldsymbol{\varphi}(\mathbf{y})^T\mathsf{A}\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathbf{p}$, $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{y}) - \beta\mathbf{p}$
      else
Step 3. {*Expansion step.*}
        $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_f\mathbf{p}$, $\mathbf{g} = \mathbf{g} - \alpha_f\mathsf{A}\mathbf{p}$, $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{x}^{k+1})$, $\Gamma = \Gamma_M$
      end *if*
    else
Step 4. {*Leaving the face in the release direction.*}
      $\mathbf{d} = \boldsymbol{\beta}(\mathbf{x}^k)$, $\alpha_{cg} = \mathbf{g}^T\mathbf{d}/\mathbf{d}^T\mathsf{A}\mathbf{d}$,
      $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_{cg}\mathbf{d}$, $\mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathbf{d}$, $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{x}^{k+1})$, $\Gamma = \Gamma_m$
    end *if*
    $k = k + 1$
  end *while*
Step 5. {*Return solution.*}
  $\widehat{\mathbf{x}} = \mathbf{x}^k$

Algorithm 5.4 uses the observations that we need not release the indices from the index set in one step and that the *release coefficient* Γ can change from iteration to iteration. The *easy re-release Polyak algorithm* starts with $\Gamma = \Gamma_M$, switches to $\Gamma = \Gamma_m$ when any index is released from the active set, and restores $\Gamma = \Gamma_M$ when the working set is expanded. Our experience [41] shows that Algorithm 5.4 is not very sensitive to the choice of Γ_m and works well with $\Gamma_m \approx 1$.

In what follows, we often use Step 4 of Algorithm 5.4 to release indices from the current active set. For any given $\Gamma > 0$, the iterates which satisfy $\|\boldsymbol{\beta}(\mathbf{x}^k)\| \leq \Gamma\|\boldsymbol{\varphi}(\mathbf{x}^k)\|$ are called *proportional*. The *proportioning step* sets $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_{cg}\boldsymbol{\beta}(\mathbf{x}^k)$ in hope that the new iterate $\mathbf{x}^{k+1}$ is proportional.

5.5.4 Properties of Modified Polyak's Algorithms

Theorem 5.5. *Let the looking ahead Polyak Algorithm 5.3 or the easy re-release Polyak Algorithm 5.4 be applied to find the solution* $\widehat{\mathbf{x}}$ *of (5.1) starting from* $\mathbf{x}^0 \in \Omega_B$. *Then there is* k *such that* $\mathbf{x}^k = \widehat{\mathbf{x}}$.

Proof. First notice that the looking ahead Polyak Algorithm 5.3 generates the same iterates as the easy re-release Polyak Algorithm 5.4 provided $\Gamma_m = \Gamma_M$, so that it is enough to prove the statement for the latter algorithm. Since by Theorem 3.1 there can be at most n consecutive conjugate gradient iterations before the unconstrained minimizer is found, it follows that there can be at most n consecutive proportional conjugate gradient iterations. Moreover, since each proportioning step releases at least one index from the working set, which has at most n elements, we have that there can be at most n^2 iterations without an expansion step.

Now observe that the iterations start with $\Gamma = \Gamma_M$, that this value is reset by any expansion step, and that $\{f(\mathbf{x}^k)\}$ is nonincreasing. Since the chain of iterations with $\Gamma = \Gamma_M$ can be interrupted only after finding the iteration $\mathbf{x}^k$ which either solves (5.1), i.e., $\boldsymbol{\beta}(\mathbf{x}^k) = \boldsymbol{\varphi}(\mathbf{x}^k) = \mathbf{o}$, or is not proportional, i.e., satisfies $\|\boldsymbol{\beta}(\mathbf{x}^k)\| > \Gamma\|\boldsymbol{\varphi}(\mathbf{x}^k)\|$ with $\Gamma \geq \kappa(\mathsf{A})^{1/2}$, it follows by Lemma 5.4 that the associated active set $\mathcal{A}(\mathbf{x}^k)$ cannot be generated again in the following iterations. Since the number of all subsets of $\mathcal{N} = \{1, \ldots, n\}$ is bounded, and by Lemma 5.4 every iteration in the face with the solution is proportional when $\Gamma \geq \kappa(\mathsf{A})^{1/2}$, we conclude that the algorithm must generate $\mathbf{x}^k = \widehat{\mathbf{x}}$ in a finite number of steps. □

Our experience [41] indicates that our modifications of the Polyak algorithm outperform the original Polyak algorithm, but a little analysis shows that they suffer from many drawbacks described in Sect. 5.4.3. Moreover, their implementation requires an estimate of the condition number of A. The easy re-release Polyak algorithm with $\Gamma_m \approx 1$ usually outperforms the looking ahead Polyak algorithm as it can better avoid an "oversolve" of the auxiliary problems defined by the faces which do not contain the solution.

5.6 Gradient Projection Method

We shall now turn our attention to the iterative algorithms whose performance is substantiated by the convergence arguments. Instead of trying to find the exact solution of (5.1), these algorithms generate the iterates that steadily approach the solution until the KKT conditions are approximately satisfied. We start with a modification of the gradient method of Sect. 3.4 that uses the Euclidean projection P_{Ω_B} onto Ω_B to generate feasible iterates. The action of P_{Ω_B} is easy to calculate. As illustrated by Fig. 5.6, the components of the projection $P_{\Omega_B}(\mathbf{x})$ of $\mathbf{x}$ onto Ω_B are given by

$$[P_{\Omega_B}(\mathbf{x})]_i = \max\{\ell_i, x_i\},\ i = 1, \ldots, n.$$

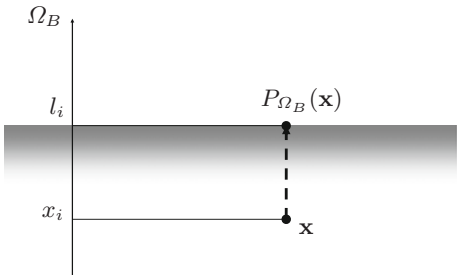

Fig. 5.6. Euclidean projection onto Ω_B

A typical step of the gradient projection method is in Fig. 5.7.

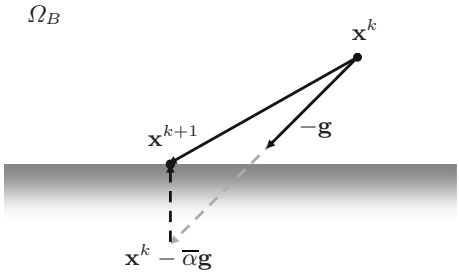

Fig. 5.7. Gradient projection step

5.6.1 Conjugate Gradient Versus Gradient Projections

Since the conjugate gradient is by Theorem 3.1 the best decrease direction which can be used to find the minimizer in the current Krylov space, probably the first idea how to plug the projection into the Polyak-type algorithms is to replace the reduced conjugate gradient step with the steplength α_f of (5.11) by the projected conjugate gradient step

$$\mathbf{x}^{k+1} = P_{\Omega_B}(\mathbf{x}^k - \alpha_{cg}\mathbf{p}^k).$$

However, if we examine Fig. 5.8, which depicts the 2D situation after the first conjugate gradient step, we can see that though the second conjugate gradient step finds the unconstrained minimizer $\mathbf{x}^k - \alpha_{cg}\mathbf{p}^k$, it can easily happen that

$$f(\mathbf{x}^k) < f(P_{\Omega_B}(\mathbf{x}^k - \alpha_{cg}\mathbf{p}^k)).$$

Figure 5.8 even suggests that it can happen for any $\alpha > \alpha_f$ that

$$f(P_{\Omega_B}(\mathbf{x}^k - \alpha\mathbf{p}^k)) > f\big(P_{\Omega_B}(\mathbf{x}^k - \alpha_f\mathbf{p}^k)\big).$$

Though Fig. 5.8 need not capture the typical situation when a small number of components of $\mathbf{x}^k - \alpha_f\mathbf{p}^k$ is affected by P_{Ω_B}, we conclude that the nice properties of the conjugate directions are guaranteed only in the feasible region. These observations comply with our discussion at the end of Sect. 3.5.

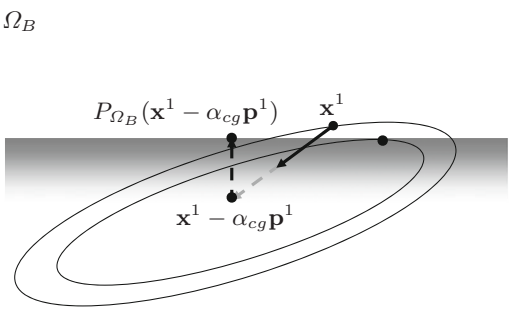

Fig. 5.8. Poor performance of the projected conjugate gradient step

On the other hand, since the gradient defines the direction of the steepest descent, it is natural to assume that for a small steplength the gradient perturbed by the projection P_{Ω_B} defines a decrease direction as in Fig. 5.9. We shall give a quantitative proof to this conjecture. In what follows, we restrict our attention to the analysis of the fixed steplength gradient iteration

$$\mathbf{x}^{k+1} = P_{\Omega_B}(\mathbf{x}^k - \overline{\alpha}\mathbf{g}^k), \tag{5.32}$$

where $\mathbf{g}^k = \nabla f(\mathbf{x}^k)$.

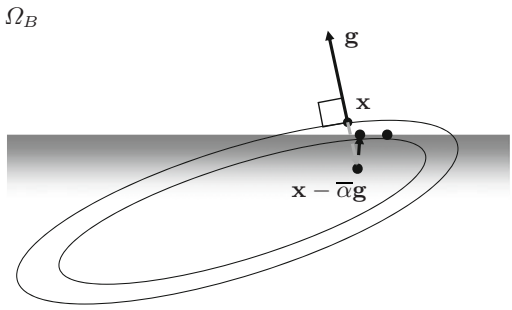

Fig. 5.9. Fixed steplength gradient step

5.6.2 Contraction in the Euclidean Norm

Which values of $\overline{\alpha}$ guarantee that the iterates defined by the fixed gradient projection step (5.32) approach the solution $\widehat{\mathbf{x}}$ in the Euclidean norm?

Proposition 5.6. *Let* $\mathbf{x} \in \Omega_B$ *and* $\mathbf{g} = \nabla f(\mathbf{x})$. *Then for any* $\overline{\alpha} > 0$

$$\|P_{\Omega_B}(\mathbf{x} - \overline{\alpha}\mathbf{g}) - \widehat{\mathbf{x}}\| \leq \eta_E \|\mathbf{x} - \widehat{\mathbf{x}}\|, \tag{5.33}$$

where $\lambda_{\min}$, $\lambda_{\max}$ *are the extreme eigenvalues of* A *and*

$$\eta_E = \max\{|1 - \overline{\alpha}\lambda_{\min}|, |1 - \overline{\alpha}\lambda_{\max}|\}. \tag{5.34}$$

Proof. Since $\widehat{\mathbf{x}} \in \Omega_B$ and the projected gradient at the solution satisfies $\widehat{\mathbf{g}}^P = \mathbf{o}$, it follows that

$$P_{\Omega_B}(\widehat{\mathbf{x}} - \overline{\alpha}\widehat{\mathbf{g}}) = \widehat{\mathbf{x}}.$$

Using that the projection P_Ω is nonexpansive by Corollary 2.7, the formula $\mathbf{g}(\mathbf{x}) = \mathsf{A}\mathbf{x} - \mathbf{b}$, the relations between the norm of a symmetric matrix and its spectrum (1.23), and the observation that if λ_i are the eigenvalues of A, then $1 - \overline{\alpha}\lambda_i$ are the eigenvalues of $\mathsf{I} - \overline{\alpha}\mathsf{A}$ (see also (1.26)), we get

$$\begin{aligned}
\|P_{\Omega_B}(\mathbf{x} - \overline{\alpha}\mathbf{g}) - \widehat{\mathbf{x}}\| &= \|P_{\Omega_B}(\mathbf{x} - \overline{\alpha}\mathbf{g}) - P_{\Omega_B}(\widehat{\mathbf{x}} - \overline{\alpha}\widehat{\mathbf{g}})\| \\
&\leq \|(\mathbf{x} - \overline{\alpha}\mathbf{g}) - (\widehat{\mathbf{x}} - \overline{\alpha}\widehat{\mathbf{g}})\| \\
&= \|\,(\mathbf{x} - \widehat{\mathbf{x}}) - \overline{\alpha}(\mathbf{g} - \widehat{\mathbf{g}})\| = \|\,(\mathbf{x} - \widehat{\mathbf{x}}) - \overline{\alpha}\mathsf{A}(\mathbf{x} - \widehat{\mathbf{x}})\| \\
&= \|(\mathsf{I} - \overline{\alpha}\mathsf{A})(\mathbf{x} - \widehat{\mathbf{x}})\| \\
&\leq \max\{|1 - \overline{\alpha}\lambda_{\min}|, |1 - \overline{\alpha}\lambda_{\max}|\}\|\mathbf{x} - \widehat{\mathbf{x}}\|.
\end{aligned}$$

□

We call η_E the *coefficient of Euclidean contraction*. If $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1})$, then $\eta_E < 1$.

Using elementary arguments of Sect. 3.5.3, we get that the coefficient η_E of Euclidean contraction (5.34) is minimized by

$$\overline{\alpha}_E^{opt} = \frac{2}{\lambda_{\min} + \lambda_{\max}} \tag{5.35}$$

and

$$\eta_E^{opt} = \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} = \frac{\kappa - 1}{\kappa + 1}, \tag{5.36}$$

where

$$\kappa = \lambda_{\max}/\lambda_{\min}$$

denotes the spectral condition number of A.

If we compare our new estimate (5.36) of the contraction of the projected gradient step with the optimal steplength $\overline{\alpha}$ in the A-norm with the estimate (3.26) of the unconstrained gradient step with the optimal steplength α_{cg} in the A-norm norm, we find, a bit surprisingly, that they are the same. This might suggest to use the A-norm optimal steplength α_{cg} also in the projected gradient step.

Unfortunately, this strategy does not work. The counterexample of Fig. 5.10 shows that if $\mathbf{g} = \mathbf{g}(\mathbf{x})$ is the eigenvector corresponding to the smallest eigenvalue $\lambda_{\min}$, then the gradient projection step with the optimal conjugate gradient steplength

$$\alpha_{cg} = \|\mathbf{g}\|^2/\mathbf{g}^T\mathsf{A}\mathbf{g} = 1/\lambda_{\min}$$

generates the iterate which is worse than $\mathbf{x}$.

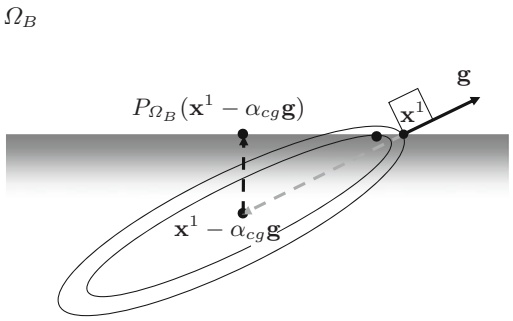

Fig. 5.10. Optimal unconstrained steplength may not be useful

Notice that the estimate (5.33) does not guarantee any bound on the decrease of the cost function. We study this topic in Sect. 5.6.5.

5.6.3 The Fixed Steplength Gradient Projection Method

Proposition 5.6 suggests that we can use the gradient projection with the fixed steplength to define an iterative algorithm with the rate of convergence in terms of bounds on the spectrum. To guarantee the convergence, the algorithm requires a computable upper bound on $\|\mathsf{A}\|$. Since A is assumed to be symmetric, it follows that $\|\mathsf{A}\|_1 = \|\mathsf{A}\|_\infty$ and, using (1.14), that $\|\mathsf{A}\| \le \|\mathsf{A}\|_\infty$. Thus we can use $\|\mathsf{A}\|_\infty$ as the upper bound. The latter inequality can be obtained also from (1.24). More hints concerning effective evaluation of an upper bound on $\|\mathsf{A}\|$ can be found in Sect. 5.9.4. The gradient projection algorithm with the fixed steplength takes the following form.

Algorithm 5.5. Gradient projection method with the fixed steplength.

Given a symmetric positive definite matrix $\mathsf{A} \in \mathbb{R}^{n\times n}$ *and n-vectors* $\mathbf{b}$, $\boldsymbol{\ell}$.
Step 0. {*Initialization.*}
Choose $\mathbf{x}^0 \in \Omega_B$, $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1})$, *set* $k = 0$
while $\|\mathbf{g}^P(\mathbf{x}^k)\|$ *is not small*
Step 1. {*The gradient projection step.*}
$\mathbf{x}^{k+1} = P_{\Omega_B}\left(\mathbf{x}^k - \overline{\alpha}\mathbf{g}(\mathbf{x}^k)\right)$
$k = k + 1$
end *while*
Step 2. {*Return (possibly inexact) solution.*}
$\widetilde{\mathbf{x}} = \mathbf{x}^k$

We can use recurrently the estimate (5.33) of Proposition 5.6 to get for $k \ge 1$ that

$$\|\mathbf{x}^k - \widehat{\mathbf{x}}\| \le \eta_E \|\mathbf{x}^{k-1} - \widehat{\mathbf{x}}\| \le \cdots \le \eta_E^k \|\mathbf{x}^0 - \widehat{\mathbf{x}}\|, \tag{5.37}$$

where $\eta_E < 1$ is the coefficient of Euclidean contraction defined by (5.34). It follows that Algorithm 5.5 generates the iterates $\mathbf{x}^k$ that *converge to the solution* $\widehat{\mathbf{x}}$ *of* (5.1) *in the Euclidean norm linearly* with the coefficient of contraction η_E. The iterates $\mathbf{x}^k$ converge in the A-norm only R-linearly with

$$\|\mathbf{x}^k - \widehat{\mathbf{x}}\|_{\mathsf{A}} \le \eta_E^k \|\mathsf{A}\| \|\mathbf{x}^0 - \widehat{\mathbf{x}}\|. \tag{5.38}$$

Though the cost of a step of Algorithm 5.5 is comparable to that of the Polyak-type algorithms, the performance of these algorithms essentially differs. A nice feature of the gradient projection algorithm is the rate of convergence in terms of bounds on the spectrum. This can hardly be proved for the Polyak algorithm; when a component of the current iterate is near the bound and the corresponding component of the conjugate direction is large, then the feasible steplength α_f and the relative decrease of the cost function can be arbitrarily small. On the other hand, unlike the Polyak algorithm, Algorithm 5.5 is not able to exploit information from the previous steps in one face.

5.6.4 Quadratic Functions with Identity Hessian

Which values of $\overline{\alpha}$ guarantee that the cost function f decreases in each iterate defined by the fixed gradient projection step (5.32)? How much does f decrease when the answer is positive? To answer these questions, it is useful to carry out some analysis of a special quadratic function

$$F(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{x} - \mathbf{c}^T\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n, \tag{5.39}$$

which is defined by a fixed $\mathbf{c} \in \mathbb{R}^n,\ \mathbf{c} = [c_i]$. We shall also use

$$F(\mathbf{x}) = \sum_{i=1}^{n} F_i(x_i), \quad F_i(x_i) = \frac{1}{2}x_i^2 - c_i x_i, \quad \mathbf{x} = [x_i]. \tag{5.40}$$

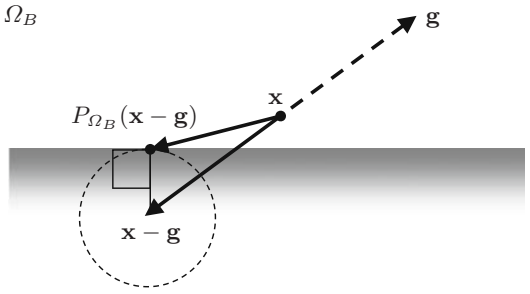

Fig. 5.11. Minimizer of F in Ω_B

The Hessian and the gradient of F are expressed, respectively, by

$$\nabla^2 F(\mathbf{x}) = \mathsf{I} \quad \text{and} \quad \mathbf{g} = \nabla F(\mathbf{x}) = \mathbf{x} - \mathbf{c}, \quad \mathbf{g} = [g_i]. \tag{5.41}$$

Thus $\mathbf{c} = \mathbf{x} - \mathbf{g}$ and for any $\mathbf{z} \in \mathbb{R}^n$

$$\|\mathbf{z} - \mathbf{c}\|^2 = \|\mathbf{z}\|^2 - 2\mathbf{c}^T\mathbf{z} + \|\mathbf{c}\|^2 = 2F(\mathbf{z}) + \|\mathbf{c}\|^2.$$

Since by Proposition 2.6 for any $\mathbf{z} \in \Omega_B$

$$\|\mathbf{z} - \mathbf{c}\| \geq \|P_{\Omega_B}(\mathbf{c}) - \mathbf{c}\|,$$

we get that for any $\mathbf{z} \in \Omega_B$

$$\begin{aligned} 2F(\mathbf{z}) &= \|\mathbf{z} - \mathbf{c}\|^2 - \|\mathbf{c}\|^2 \geq \|P_{\Omega_B}(\mathbf{c}) - \mathbf{c}\|^2 - \|\mathbf{c}\|^2 \\ &= 2F\left(P_{\Omega_B}(\mathbf{c})\right) = 2F\left(P_{\Omega_B}(\mathbf{x} - \mathbf{g})\right). \end{aligned} \tag{5.42}$$

We have thus proved that if $\mathbf{y} \in \Omega_B$, then, as illustrated in Fig. 5.11,

$$F\big(P_{\Omega_B}(\mathbf{x} - \mathbf{g})\big) \leq F(\mathbf{y}). \tag{5.43}$$

We are especially interested in the analysis of F along the projected-gradient path

$$p(\mathbf{x},\alpha) = P_{\Omega_B}\big(\mathbf{x} - \alpha\nabla F(\mathbf{x})\big) = \max\{\mathbf{x} - \alpha\mathbf{g}, \boldsymbol{\ell}\},$$

where the maximum is assumed to be carried out componentwise, $\alpha \geq 0$, and $\mathbf{x} \in \Omega_B$ is fixed. We shall often use that the projected-gradient path can be described by

$$p(\mathbf{x},\alpha) = P_{\Omega_B}\left(\mathbf{x} - \alpha\mathbf{g}\right) = \mathbf{x} - \alpha\widetilde{\mathbf{g}}(\alpha), \tag{5.44}$$

where $\widetilde{\mathbf{g}}(\alpha)$ denotes the *reduced gradient* whose components are defined by

$$\widetilde{g}_i(0) = 0 \quad \text{and} \quad \widetilde{g}_i(\alpha) = \min\{(x_i - \ell_i)/\alpha, g_i\} \quad \text{for} \quad \alpha > 0.$$

A geometric illustration of the projected-gradient path is in Fig. 5.12.

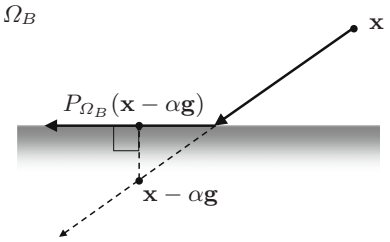

Fig. 5.12. Projected-gradient path

Due to the separability of F, the following analysis of a special case with F defined on $\mathbb{R}$ is important also in the general case.

Lemma 5.7. *Let $x, \ell, c \in \mathbb{R}$, $x \geq \ell$. Let F and g be defined by*

$$F(x) = \frac{1}{2}x^2 - cx \quad \text{and} \quad g = x - c.$$

Then for any $\delta \in [0,1]$

$$F\big(P_{\Omega_B}(x - (2-\delta)g)\big) \leq F\big(P_{\Omega_B}(x - \delta g)\big). \tag{5.45}$$

Proof. First assume that $x \geq l$ is fixed and denote

$$g = F'(x) = x - c, \quad \widetilde{g}(0) = 0, \quad \widetilde{g}(\alpha) = \min\{(x-\ell)/\alpha, g\}, \quad \alpha \neq 0.$$

For convenience, let us define

$$F\big(P_{\Omega_B}(x - \alpha g)\big) = F(x) + \Phi(\alpha), \quad \Phi(\alpha) = -\alpha\widetilde{g}(\alpha)g + \frac{\alpha^2}{2}\left(\widetilde{g}(\alpha)\right)^2, \quad \alpha \geq 0.$$

Moreover, using these definitions, it can be checked directly that Φ is defined explicitly by

$$\Phi(\alpha) = \begin{cases} \Phi_F(\alpha) & \text{for} \quad \alpha \in (-\infty, \overline{\xi}] \cap [0, \infty) \quad \text{or} \quad g \le 0, \\ \Phi_A(\alpha) & \text{for} \quad \alpha \in [\ \overline{\xi}\ , \infty) \cap [0, \infty) \quad \text{and} \quad g > 0, \end{cases}$$

where $\overline{\xi} = \infty$ if $g = 0$, $\overline{\xi} = (x - \ell)/g$ if $g \neq 0$,

$$\Phi_F(\alpha) = \left(-\alpha + \frac{\alpha^2}{2}\right) g^2, \quad \text{and} \quad \Phi_A(\alpha) = -g(x - \ell) + \frac{1}{2}(x - \ell)^2.$$

See also Fig. 5.13.

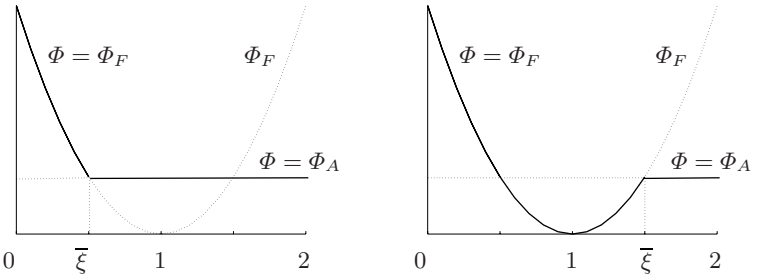

Fig. 5.13. Graphs of Φ for $\overline{\xi} < 1$ (left) and $\overline{\xi} > 1$ (right) when $g > 0$

It follows that for any α

$$\Phi_F(2 - \alpha) = \left(-(2 - \alpha) + \frac{(2 - \alpha)^2}{2}\right) g^2 = \Phi_F(\alpha), \tag{5.46}$$

and if $g \le 0$, then

$$\Phi(\alpha) = \Phi_F(\alpha) = \Phi_F(2 - \alpha) = \Phi(2 - \alpha).$$

Let us now assume that $g > 0$ and denote $\overline{\xi} = (x - \ell)/g$. Simple analysis shows that if $\overline{\xi} \in [0, 1]$, then Φ is nonincreasing on $[0, 2]$ and (5.45) is satisfied for $\alpha \in [0, 1]$. To finish the proof of (5.45), notice that if $1 < \overline{\xi}$, then

$$\Phi(\alpha) = \Phi_F(\alpha), \quad \alpha \in [0, 1], \quad \Phi(\alpha) \le \Phi_F(\alpha), \quad \alpha \in [1, 2],$$

so that we can use (5.46) to get that for $\alpha \in [0, 1]$

$$\Phi(2 - \alpha) \le \Phi_F(2 - \alpha) = \Phi_F(\alpha) = \Phi(\alpha).$$

□

The following property of F is essential in the analysis of the decrease of f along the projected-gradient path in the next subsection.

Corollary 5.8. *Let* $\mathbf{x}, \boldsymbol{\ell}, \mathbf{c} \in \mathbb{R}^n$, $\mathbf{x} \geq \boldsymbol{\ell}$. *Let* F *be defined by (5.39). Then for any* $\delta \in [0, 1]$

$$F\big(P_{\Omega_B}(\mathbf{x} - (2-\delta)\mathbf{g})\big) \leq F\big(P_{\Omega_B}(\mathbf{x} - \delta\mathbf{g})\big). \tag{5.47}$$

Proof. If $n = 1$, then the statement reduces to Lemma 5.7.

To prove the statement for $n > 1$, first observe that for any $\mathbf{y} \in \mathbb{R}$

$$[P_{\Omega_B}(\mathbf{y})]_i = \max\{y_i, \ell_i\},\ i = 1, \ldots, n.$$

It follows that P_{Ω_B} is separable and can be defined componentwise by the real functions

$$P_i(y) = \max\{y, \ell_i\},\ i = 1, \ldots, n.$$

Using the separable representation of F given by (5.40) and Lemma 5.7, we get

$$\begin{aligned} F\big(P_{\Omega_B}(\mathbf{x} - (2-\delta)\mathbf{g})\big) &= \sum_{i=1}^{n} F_i\big([P_{\Omega_B}(\mathbf{x} - (2-\delta)\mathbf{g})]_i\big) \\ &= \sum_{i=1}^{n} F_i\big(P_i(x_i - (2-\delta)g_i)\big) \\ &\leq \sum_{i=1}^{n} F_i\big(P_i(x_i - \delta g_i)\big) \\ &= F\big(P_{\Omega_B}(\mathbf{x} - \delta\mathbf{g})\big). \end{aligned}$$

□

5.6.5 Dominating Function and Decrease of the Cost Function

Now we are ready to give an estimate of the decrease of the cost function f in the iterates defined by the gradient projection step (5.32). The idea of the proof is to replace f by a suitable quadratic function F which dominates f and whose Hessian is the identity matrix.

Let us assume that $0 < \delta\|\mathsf{A}\| \leq 1$ and let $\mathbf{x} \in \Omega_B$ be arbitrary but fixed, so that we can define a quadratic function

$$F_\delta(\mathbf{y}) = \delta f(\mathbf{y}) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T(\mathsf{I} - \delta\mathsf{A})(\mathbf{y} - \mathbf{x}), \quad \mathbf{y} \in \mathbb{R}^n.$$

It is defined so that

$$F_\delta(\mathbf{x}) = \delta f(\mathbf{x}), \quad \nabla F_\delta(\mathbf{x}) = \delta\nabla f(\mathbf{x}) = \delta\mathbf{g}, \quad \text{and} \quad \nabla^2 F_\delta(\mathbf{y}) = \mathsf{I}. \tag{5.48}$$

Moreover, for any $\mathbf{y} \in \mathbb{R}^n$

$$\delta f(\mathbf{y}) \le F_\delta(\mathbf{y}). \tag{5.49}$$

It follows that

$$\delta f\left(P_{\Omega_B}(\mathbf{x}-\delta\mathbf{g})\right) - \delta f(\widehat{\mathbf{x}}) \le F_\delta\left(P_{\Omega_B}(\mathbf{x}-\delta\mathbf{g})\right) - \delta f(\widehat{\mathbf{x}}) \tag{5.50}$$

and

$$\nabla F_\delta(\mathbf{y}) = \delta\nabla f(\mathbf{y}) + (\mathsf{I}-\delta\mathsf{A})(\mathbf{y}-\mathbf{x}) = \mathbf{y} - (\mathbf{x}-\delta\mathbf{g}). \tag{5.51}$$

Using (5.43) and (5.48), we get that for any $\mathbf{z} \in \Omega_B$

$$F_\delta\left(P_{\Omega_B}(\mathbf{x}-\delta\mathbf{g})\right) \le F_\delta(\mathbf{z}). \tag{5.52}$$

The following lemma is due to Schöberl [165, 74].

Lemma 5.9. *Let $\widehat{\mathbf{x}}$ denote the unique solution of (5.1), let $\lambda_{\min}$ denote the smallest eigenvalue of A, $\mathbf{g} = \nabla f(\mathbf{x})$, $\mathbf{x} \in \Omega_B$, and $\delta \in (0, \|\mathsf{A}\|^{-1}]$. Then*

$$F_\delta\left(P_{\Omega_B}(\mathbf{x}-\delta\mathbf{g})\right) - \delta f(\widehat{\mathbf{x}}) \le \delta(1-\delta\lambda_{\min})\left(f(\mathbf{x}) - f(\widehat{\mathbf{x}})\right). \tag{5.53}$$

Proof. Let us denote

$$[\widehat{\mathbf{x}}, \mathbf{x}] = \mathrm{Conv}\{\widehat{\mathbf{x}}, \mathbf{x}\} \quad \text{and} \quad \mathbf{d} = \widehat{\mathbf{x}} - \mathbf{x}.$$

Using (5.52),

$$[\widehat{\mathbf{x}}, \mathbf{x}] = \{\mathbf{x} + t\mathbf{d} : t \in [0,1]\} \subseteq \Omega_B,$$

$0 < \lambda_{\min}\delta \le \|\mathsf{A}\|\delta \le 1$, and $\lambda_{\min}\|\mathbf{d}\|^2 \le \mathbf{d}^T\mathsf{A}\mathbf{d}$, we get

$$\begin{aligned}
F_\delta\left(P_{\Omega_B}(\mathbf{x}-\delta\mathbf{g})\right) - \delta f(\widehat{\mathbf{x}}) &= \min\{F_\delta(\mathbf{y}) - \delta f(\widehat{\mathbf{x}}) :\ \mathbf{y} \in \Omega_B\} \\
&\le \min\{F_\delta(\mathbf{y}) - \delta f(\widehat{\mathbf{x}}) :\ \mathbf{y} \in [\widehat{\mathbf{x}}, \mathbf{x}]\} \\
&= \min\{F_\delta(\mathbf{x}+t\mathbf{d}) - \delta f(\mathbf{x}+\mathbf{d}) :\ t \in [0,1]\} \\
&= \min\{\delta t\mathbf{d}^T\mathbf{g} + \frac{t^2}{2}\|\mathbf{d}\|^2 - \delta\mathbf{d}^T\mathbf{g} - \frac{\delta}{2}\mathbf{d}^T\mathsf{A}\mathbf{d} :\ t \in [0,1]\} \\
&\le \delta^2\lambda_{\min}\mathbf{d}^T\mathbf{g} + \frac{1}{2}\delta^2\lambda_{\min}^2\|\mathbf{d}\|^2 - \delta\mathbf{d}^T\mathbf{g} - \frac{\delta}{2}\mathbf{d}^T\mathsf{A}\mathbf{d} \\
&\le \delta^2\lambda_{\min}\mathbf{d}^T\mathbf{g} + \frac{1}{2}\delta^2\lambda_{\min}\mathbf{d}^T\mathsf{A}\mathbf{d} - \delta\mathbf{d}^T\mathbf{g} - \frac{\delta}{2}\mathbf{d}^T\mathsf{A}\mathbf{d} \\
&= \delta(\delta\lambda_{\min} - 1)(\mathbf{d}^T\mathbf{g} + \frac{1}{2}\mathbf{d}^T\mathsf{A}\mathbf{d}) \\
&= \delta(\delta\lambda_{\min} - 1)\left(f(\mathbf{x}+\mathbf{d}) - f(\mathbf{x})\right) \\
&= \delta(1-\delta\lambda_{\min})\left(f(\mathbf{x}) - f(\widehat{\mathbf{x}})\right).
\end{aligned}$$

This proves (5.53). □

Proposition 5.10. *Let* $\widehat{\mathbf{x}}$ *denote the unique solution of* (5.1), $\mathbf{g} = \nabla f(\mathbf{x})$, $\mathbf{x} \in \Omega_B$, *and let* $\lambda_{\min}$ *denote the smallest eigenvalue of* A.

If $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$, *then*

$$f\left(P_{\Omega_B}(\mathbf{x} - \overline{\alpha}\mathbf{g})\right) - f(\widehat{\mathbf{x}}) \leq \eta_f \left(f(\mathbf{x}) - f(\widehat{\mathbf{x}})\right), \tag{5.54}$$

where

$$\eta_f = 1 - \widehat{\alpha}\lambda_{\min} \tag{5.55}$$

is the cost function reduction coefficient and $\widehat{\alpha} = \min\{\overline{\alpha}, 2\|\mathsf{A}\|^{-1} - \overline{\alpha}\}$.

Proof. Let us first assume that $0 < \overline{\alpha}\|\mathsf{A}\| \leq 1$ and let $\mathbf{x} \in \Omega_B$ be arbitrary but fixed, so that we can use Lemma 5.9 with $\delta = \overline{\alpha}$ to get

$$F_{\overline{\alpha}}\left(P_{\Omega_B}(\mathbf{x} - \overline{\alpha}\mathbf{g})\right) - \overline{\alpha} f(\widehat{\mathbf{x}}) \leq \overline{\alpha}(1 - \overline{\alpha}\lambda_{\min})\left(f(\mathbf{x}) - f(\widehat{\mathbf{x}})\right). \tag{5.56}$$

In combination with (5.50), this proves (5.54) for $0 < \overline{\alpha} \leq \|\mathsf{A}\|^{-1}$.

To prove the statement for $\overline{\alpha} \in (\|A\|^{-1}, 2\|A\|^{-1}]$, let us first assume that $\|A\| = 1$ and let $\overline{\alpha} = 2 - \delta$, $\delta \in (0, 1)$. Then F_1 dominates f and

$$\delta F_1(\mathbf{y}) \leq \delta F_1(\mathbf{y}) + \frac{1-\delta}{2}\|\mathbf{y} - \mathbf{x}\|^2 = F_\delta(\mathbf{y}). \tag{5.57}$$

Thus we can apply (5.49), Corollary 5.8, and the latter inequality to get

$$\begin{aligned}\delta f\big(P_\Omega(\mathbf{x} - \overline{\alpha}\mathbf{g})\big) &\leq \delta F_1\big(P_\Omega(\mathbf{x} - \overline{\alpha}\mathbf{g})\big) \leq \delta F_1\big(P_\Omega(\mathbf{x} - \delta\mathbf{g})\big) \\ &\leq F_\delta\big(P_\Omega(\mathbf{x} - \delta\mathbf{g})\big).\end{aligned}$$

Combining the latter inequalities with (5.56) for $\overline{\alpha} = \delta$, we get

$$\delta f\big(P_\Omega(\mathbf{x} - \overline{\alpha}\mathbf{g})\big) - \delta f(\widehat{\mathbf{x}}) \leq \delta(1 - \delta\lambda_{\min})\big(f(\mathbf{x}) - f(\widehat{\mathbf{x}})\big).$$

This proves the statement for $\overline{\alpha} \in (\|A\|^{-1}, 2\|A\|^{-1})$ and $\|A\| = 1$. To finish the proof, apply the last inequality divided by η to the function $\|A\|^{-1} f$ and recall that f and P_Ω are continuous. □

The estimate (5.54) gives the best value

$$\eta_f^{opt} = 1 - \kappa(\mathsf{A})^{-1}$$

for $\overline{\alpha} = \|\mathsf{A}\|^{-1}$ with $\kappa(\mathsf{A}) = \|\mathsf{A}\|\|\mathsf{A}^{-1}\|$. If $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1})$ and the iterates $\{\mathbf{x}^i\}$ are generated by Algorithm 5.5, we can use (5.54) to get for $k \geq 1$

$$f(\mathbf{x}^k) - f(\widehat{\mathbf{x}}) \leq \eta_f \left(f(\mathbf{x}^{k-1}) - f(\widehat{\mathbf{x}})\right) \leq \cdots \leq \eta_f^k \left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\right), \tag{5.58}$$

where $\eta_f < 1$ is given by (5.55). It follows by Lemma 5.1 that

$$\|\mathbf{x}^k - \widehat{\mathbf{x}}\|_A^2 \leq 2\left(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\right) \leq 2\eta_f^k \left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\right) \leq 2\lambda_{\min}^{-1}\eta_f^k \|\mathbf{g}^P\|, \tag{5.59}$$

where $\mathbf{g}^P = \mathbf{g}^P(\mathbf{x}^0)$. The latter bound on the R-linear convergence in the energy norm is asymptotically worse than (5.38), but its right-hand side does not enhance the solution and can be effectively evaluated.

5.7 Modified Proportioning with Gradient Projections

In the previous sections, we learned that the solution of auxiliary problems in the active set algorithm for solving (5.1) can be implemented by the conjugate gradient method and we got the estimate (5.54) for the decrease of the cost function f in the gradient projection step with the fixed steplength $\overline{\alpha} \in \big(0, 2\|\mathsf{A}\|^{-1}\big)$. Now we are ready to combine these observations in order to develop an *effective algorithm with the R-linear rate of convergence of f* that can be expressed in terms of bounds on the spectrum of the Hessian of f. The only difficulty which we must overcome is to ensure that the free gradient is always sufficiently large in the conjugate gradient iterations, since the conjugate gradient method reduces efficiently only the free gradient and is inefficient when the norm of the chopped gradient dominates the error of the KKT conditions. Using the methods of the next section, it is possible to prove for our new algorithm the finite termination for regular solution and the convergence, but not the R-linear convergence, of the projected gradient to zero in the general case. Here we restrict our attention to the R-linear convergence of the iterates in the energy norm.

5.7.1 MPGP Schema

The algorithm that we propose here exploits a user-defined constant $\Gamma > 0$, a test which is used to decide when to leave the face, and three types of steps.

The *conjugate gradient step*, defined as in Polyak's algorithm on page 165 by

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_{cg}\mathbf{p}^{k+1}, \tag{5.60}$$

is used to carry out efficiently the minimization in the face $\mathcal{W}_{\mathcal{I}}$ given by $\mathcal{I} = \mathcal{A}(\mathbf{x}^s)$. We shall use in our proofs that by Theorem 3.1

$$f(\mathbf{x}^{k+1}) = \min\{f(\mathbf{x}^s + \mathbf{y}) : \ \mathbf{y} \in \mathrm{Span}\{\boldsymbol{\varphi}(\mathbf{x}^s), \ldots, \boldsymbol{\varphi}(\mathbf{x}^k)\}\}. \tag{5.61}$$

The *gradient projection step* is defined by the gradient projection

$$\mathbf{x}^{k+1} = P_{\Omega_B}\big(\mathbf{x}^k - \overline{\alpha}\mathbf{g}(\mathbf{x}^k)\big) = \max\{\boldsymbol{\ell}, \mathbf{x}^k - \overline{\alpha}\mathbf{g}(\mathbf{x}^k)\} \tag{5.62}$$

with the fixed steplength. This step can both add and remove indices from the current working set. To describe the gradient projection step in the form suitable for our analysis, let us introduce, for any $\mathbf{x} \in \Omega_B$ and $\alpha > 0$, the *reduced free gradient* $\widetilde{\boldsymbol{\varphi}}_\alpha(\mathbf{x})$ with the entries

$$\widetilde{\varphi}_i = \widetilde{\varphi}_i(\mathbf{x}, \alpha) = \min\{(x_i - \ell_i)/\alpha, \varphi_i\}, \quad i \in \mathcal{N} = \{1, \ldots, n\}. \tag{5.63}$$

Thus

$$P_{\Omega_B}\big(\mathbf{x} - \alpha\mathbf{g}(\mathbf{x})\big) = \mathbf{x} - \alpha\big(\widetilde{\boldsymbol{\varphi}}_\alpha(\mathbf{x}) + \boldsymbol{\beta}(\mathbf{x})\big). \tag{5.64}$$

If the steplength is equal to $\overline{\alpha}$ and the inequality

$$||\boldsymbol{\beta}(\mathbf{x}^k)||^2 \leq \Gamma^2 \widetilde{\boldsymbol{\varphi}}_{\overline{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k) \tag{5.65}$$

holds, then we call the iterate $\mathbf{x}^k$ *strictly proportional*. The test (5.65) is used to decide which components of the projected gradient $\mathbf{g}^P(\mathbf{x}^k)$ should be reduced in the next step. Notice that the right-hand side of (5.65) blends the information about the free gradient and its part that can be used in the gradient projection step.

The *proportioning step* is defined by

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_{cg} \boldsymbol{\beta}(\mathbf{x}^k) \tag{5.66}$$

with the steplength α_{cg} that minimizes $f\left(\mathbf{x}^k - \alpha \boldsymbol{\beta}(\mathbf{x}^k)\right)$. It has been shown in Sect. 3.1 that the CG steplength α_{cg} that minimizes $f(\mathbf{x} - \alpha \mathbf{d})$ for a given $\mathbf{d}$ and $\mathbf{x}$ can be evaluated using the gradient $\mathbf{g} = \mathbf{g}(\mathbf{x}) = \nabla f(\mathbf{x})$ at $\mathbf{x}$ by

$$\alpha_{cg} = \alpha_{cg}(\mathbf{d}) = \frac{\mathbf{d}^T \mathbf{g}}{\mathbf{d}^T \mathsf{A} \mathbf{d}}. \tag{5.67}$$

The purpose of the proportioning step is to remove the indices of the components of the gradient $\mathbf{g}$ that violate the KKT conditions from the working set. Note that if $\mathbf{x}^k \in \Omega_B$, then

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_{cg} \boldsymbol{\beta}(\mathbf{x}^k) \in \Omega_B.$$

Now we are ready to define the algorithm in the form that is convenient for analysis. For its implementation, see Sect. 5.9.

Algorithm 5.6. Modified proportioning with gradient projections (MPGP schema).

Given a symmetric positive definite matrix $\mathsf{A} \in \mathbb{R}^{n \times n}$ *and* n*-vectors* $\mathbf{b}$, $\boldsymbol{\ell}$. *Choose* $\mathbf{x}^0 \in \Omega_B$, $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1})$, *and* $\Gamma > 0$. *Set* $k = 0$. *For* $k \geq 0$ *and* $\mathbf{x}^k$ *known, choose* $\mathbf{x}^{k+1}$ *by the following rules:*

(i) If $\mathbf{g}^P(\mathbf{x}^k) = \mathbf{o}$, *set* $\mathbf{x}^{k+1} = \mathbf{x}^k$.

(ii) If $\mathbf{x}^k$ *is strictly proportional and* $\mathbf{g}^P(\mathbf{x}^k) \neq \mathbf{o}$, *try to generate* $\mathbf{x}^{k+1}$ *by the conjugate gradient step. If* $\mathbf{x}^{k+1} \in \Omega_B$, *then accept it, else generate* $\mathbf{x}^{k+1}$ *by the gradient projection step.*

(iii) If $\mathbf{x}^k$ *is not strictly proportional, define* $\mathbf{x}^{k+1}$ *by proportioning.*

We call our algorithm modified proportioning to distinguish it from earlier algorithms introduced independently by Friedlander and Martínez with their collaborators [94, 95, 96, 14, 33] and Dostál [41, 42]. These earlier algorithms applied the proportioning step when

$$\|\boldsymbol{\beta}(\mathbf{x}^k)\| \leq \Gamma^2 \|\boldsymbol{\varphi}\left(\mathbf{x}^k\right)\|.$$

5.7.2 Rate of Convergence

Now we are ready to prove the R-linear rate of convergence of MPGP in terms of bounds on the spectrum of the Hessian A for $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1})$.

Theorem 5.11. *Let $\{\mathbf{x}^k\}$ be generated by Algorithm 5.6 with $\mathbf{x}^0 \in \Omega_B$, $\Gamma > 0$, and $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$. Then*

$$f(\mathbf{x}^{k+1}) - f(\widehat{\mathbf{x}}) \le \eta_\Gamma \left(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\right), \tag{5.68}$$

where $\widehat{\mathbf{x}}$ denotes the unique solution of (5.1),

$$\eta_\Gamma = 1 - \frac{\widehat{\alpha}\lambda_{\min}}{\vartheta + \vartheta\widehat{\Gamma}^2}, \qquad \widehat{\Gamma} = \max\{\Gamma, \Gamma^{-1}\}, \tag{5.69}$$

$$\vartheta = 2\max\{\overline{\alpha}\|\mathsf{A}\|, 1\}, \quad \widehat{\alpha} = \min\{\overline{\alpha}, 2\|\mathsf{A}\|^{-1} - \overline{\alpha}\}, \tag{5.70}$$

and $\lambda_{\min}$ denotes the smallest eigenvalue of A.

The error in the A-norm is bounded by

$$\|\mathbf{x}^k - \widehat{\mathbf{x}}\|_{\mathsf{A}}^2 \le 2\eta_\Gamma^k \left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\right). \tag{5.71}$$

Proof. Since we have the estimate (5.54) for the gradient projection step with $\eta_f \le \eta_\Gamma$, it is enough to estimate the decrease of the cost function for the other two steps. Our main tools are (5.54) and the inequality

$$f\Big(P_{\Omega_B}\left(\mathbf{x}^k - \alpha\mathbf{g}(\mathbf{x}^k)\right)\Big) \ge f(\mathbf{x}^k) - \alpha\left(\widetilde{\boldsymbol{\varphi}}_\alpha(\mathbf{x}^k)^T\boldsymbol{\varphi}(\mathbf{x}^k) + \|\boldsymbol{\beta}(\mathbf{x}^k)\|^2\right), \tag{5.72}$$

which is valid for any $\alpha \ge 0$ and can be obtained from the Taylor expansion

$$f(\mathbf{x} + \mathbf{d}) = f(\mathbf{x}) + \mathbf{d}^T\mathbf{g}(\mathbf{x}) + \frac{1}{2}\mathbf{d}^T\mathsf{A}\mathbf{d} \ge f(\mathbf{x}) + \mathbf{d}^T\mathbf{g}(\mathbf{x}) \tag{5.73}$$

by substituting

$$\mathbf{x} = \mathbf{x}^k, \quad \mathbf{d} = -\alpha\big(\widetilde{\boldsymbol{\varphi}}_\alpha(\mathbf{x}^k) + \boldsymbol{\beta}(\mathbf{x}^k)\big), \quad \text{and} \quad \mathbf{g} = \boldsymbol{\varphi}(\mathbf{x}) + \boldsymbol{\beta}(\mathbf{x}).$$

If $\mathbf{x}^{k+1}$ is generated by the conjugate gradient step (5.60), then by (5.61) and (5.67)

$$\begin{aligned} f(\mathbf{x}^{k+1}) &\le f\left(\mathbf{x}^k - \alpha_{cg}\boldsymbol{\varphi}(\mathbf{x}^k)\right) = f(\mathbf{x}^k) - \frac{1}{2}\frac{\|\boldsymbol{\varphi}(\mathbf{x}^k)\|^4}{\boldsymbol{\varphi}(\mathbf{x}^k)^T\mathsf{A}\boldsymbol{\varphi}(\mathbf{x}^k)} \\ &\le f(\mathbf{x}^k) - \frac{1}{2}\|\mathsf{A}\|^{-1}\|\boldsymbol{\varphi}(\mathbf{x}^k)\|^2. \end{aligned}$$

Taking into account $\widehat{\alpha} \le \|\mathsf{A}\|^{-1}$ and $\widetilde{\varphi}_i\varphi_i \le \varphi_i^2$, $i = 1, \ldots, n$, we get

$$f(\mathbf{x}^{k+1}) \le f(\mathbf{x}^k) - \frac{1}{2}\|\mathsf{A}\|^{-1}\|\boldsymbol{\varphi}(\mathbf{x}^k)\|^2 \le f(\mathbf{x}^k) - \frac{\widehat{\alpha}}{2}\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T\boldsymbol{\varphi}(\mathbf{x}^k). \tag{5.74}$$

Now observe that the conjugate gradient step is used only when $\mathbf{x}^k$ is strictly proportional, i.e.,

$$\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2 \leq \Gamma^2 \widetilde{\boldsymbol{\varphi}}_{\overline{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k).$$

Since $\widehat{\alpha} \leq \overline{\alpha}$ implies

$$\widetilde{\boldsymbol{\varphi}}_{\overline{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k) \leq \widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k),$$

it follows that

$$\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2 \leq \Gamma^2 \widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k). \tag{5.75}$$

After substituting (5.75) into (5.72) with $\alpha = \widehat{\alpha}$, we get

$$f\big(P_{\Omega_B}\left(\mathbf{x}^k - \widehat{\alpha}\mathbf{g}(\mathbf{x}^k)\right)\big) \geq f(\mathbf{x}^k) - \widehat{\alpha}(1+\Gamma^2)\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k). \tag{5.76}$$

Thus for $\mathbf{x}^{k+1}$ generated by the conjugate gradient step, we get by elementary algebra and application of (5.76) that

$$\begin{aligned} f(\mathbf{x}^{k+1}) &\leq f(\mathbf{x}^k) - \frac{\widehat{\alpha}}{2}\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k) \\ &= \frac{1}{2+2\Gamma^2}\Big(f(\mathbf{x}^k) - \widehat{\alpha}(1+\Gamma^2)\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k) + (1+2\Gamma^2) f(\mathbf{x}^k)\Big) \\ &\leq \frac{1}{2+2\Gamma^2}\Big(f\big(P_{\Omega_B}\left(\mathbf{x}^k - \widehat{\alpha}\mathbf{g}(\mathbf{x}^k)\right)\big) + (1+2\Gamma^2) f(\mathbf{x}^k)\Big). \end{aligned}$$

After inserting $-f(\widehat{\mathbf{x}}) + f(\widehat{\mathbf{x}})$ into the last term and using (5.54) with simple manipulations, we get

$$\begin{aligned} f(\mathbf{x}^{k+1}) &\leq \frac{\eta_f + 1 + 2\Gamma^2}{2+2\Gamma^2} f(\mathbf{x}^k) + \frac{1-\eta_f}{2+2\Gamma^2} f(\widehat{\mathbf{x}}) \\ &= \frac{\eta_f + 1 + 2\Gamma^2}{2+2\Gamma^2}\left(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\right) + f(\widehat{\mathbf{x}}). \end{aligned} \tag{5.77}$$

Let us finally assume that $\mathbf{x}^{k+1}$ is generated by the proportioning step (5.66), so that

$$\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2 > \Gamma^2 \widetilde{\boldsymbol{\varphi}}_{\overline{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k) \tag{5.78}$$

and

$$\begin{aligned} f(\mathbf{x}^{k+1}) = f\left(\mathbf{x}^k - \alpha_{cg}\boldsymbol{\beta}(\mathbf{x}^k)\right) &= f(\mathbf{x}^k) - \frac{1}{2}\frac{\|\boldsymbol{\beta}(\mathbf{x}^k)\|^4}{\boldsymbol{\beta}(\mathbf{x}^k)^T \mathsf{A}\boldsymbol{\beta}(\mathbf{x}^k)} \\ &\leq f(\mathbf{x}^k) - \frac{1}{2}\|\mathsf{A}\|^{-1}\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2. \end{aligned}$$

Taking into account the definition of $\overline{\alpha}$ and ϑ, we get

$$\overline{\alpha}/\vartheta \leq \|\mathsf{A}\|^{-1}/2$$

and

$$f(\mathbf{x}^{k+1}) \le f(\mathbf{x}^k) - \frac{\overline{\alpha}}{\vartheta}\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2, \tag{5.79}$$

where the right-hand side may be rewritten in the form

$$f(\mathbf{x}^k) - \frac{\overline{\alpha}}{\vartheta}\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2 = \frac{1}{\vartheta(1+\Gamma^{-2})}\left(f(\mathbf{x}^k) - \overline{\alpha}(1+\Gamma^{-2})\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2\right) + \frac{\vartheta+\vartheta\Gamma^{-2}-1}{\vartheta(1+\Gamma^{-2})}f(\mathbf{x}^k). \tag{5.80}$$

We can also substitute (5.78) into (5.72) to get

$$f\left(P_{\Omega_B}\left(\mathbf{x}^k - \overline{\alpha}\mathbf{g}(\mathbf{x}^k)\right)\right) > f(\mathbf{x}^k) - \overline{\alpha}(1+\Gamma^{-2})\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2. \tag{5.81}$$

After substituting (5.81) into (5.80), using (5.79), (5.54) with $\mathbf{x} = \mathbf{x}^k$, and simple manipulations, we get

$$\begin{aligned} f(\mathbf{x}^{k+1}) &< \frac{1}{\vartheta+\vartheta\Gamma^{-2}} f\Big(P_{\Omega_B}\left(\mathbf{x}^k - \overline{\alpha}\mathbf{g}(\mathbf{x}^k)\right)\Big) + \frac{\vartheta+\vartheta\Gamma^{-2}-1}{\vartheta+\vartheta\Gamma^{-2}} f(\mathbf{x}^k) \\ &= \frac{1}{\vartheta+\vartheta\Gamma^{-2}}\Big(f\big(P_{\Omega_B}\left(\mathbf{x}^k - \overline{\alpha}\mathbf{g}(\mathbf{x}^k)\right)\big) - f(\widehat{\mathbf{x}})\Big) \\ &+ \frac{1}{\vartheta+\vartheta\Gamma^{-2}} f(\widehat{\mathbf{x}}) + \frac{\vartheta+\vartheta\Gamma^{-2}-1}{\vartheta+\vartheta\Gamma^{-2}} f(\mathbf{x}^k) \\ &\le \frac{\eta_f}{\vartheta+\vartheta\Gamma^{-2}}\Big(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\Big) + \frac{1}{\vartheta+\vartheta\Gamma^{-2}} f(\widehat{\mathbf{x}}) + \frac{\vartheta+\vartheta\Gamma^{-2}-1}{\vartheta+\vartheta\Gamma^{-2}} f(\mathbf{x}^k) \\ &= \frac{\eta_f+\vartheta+\vartheta\Gamma^{-2}-1}{\vartheta+\vartheta\Gamma^{-2}}\Big(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\Big) + f(\widehat{\mathbf{x}}). \end{aligned}$$

Comparing the last inequality with (5.77) and taking into account that by the definition $\Gamma \le \widehat{\Gamma}$, $\Gamma^{-1} \le \widehat{\Gamma}$, and $\vartheta \ge 2$, we obtain that the estimate

$$f(\mathbf{x}^{k+1}) - f(\widehat{\mathbf{x}}) \le \frac{\eta_f+\vartheta+\vartheta\Gamma^{-2}-1}{\vartheta+\vartheta\Gamma^{-2}}\Big(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\Big)$$

is valid for both the CG step and the proportioning step. The proof of (5.68) is completed by

$$\eta_\Gamma = \frac{\eta_f+\vartheta+\vartheta\Gamma^{-2}-1}{\vartheta+\vartheta\Gamma^{-2}} = 1 - \frac{1-\eta_f}{\vartheta+\vartheta\Gamma^{-2}} = 1 - \frac{\widehat{\alpha}\lambda_{\min}}{\vartheta+\vartheta\widehat{\Gamma}^2}.$$

To get the error bound (5.71), notice that by Lemma 5.1

$$\|\mathbf{x}^k - \widehat{\mathbf{x}}\|_A^2 \le 2\left(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\right) \le 2\eta_\Gamma^k\left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\right). \quad \square \tag{5.82}$$

Theorem 5.11 gives the best bound on the rate of convergence for $\Gamma = \widehat{\Gamma} = 1$ in agreement with the heuristics that we should leave the face when the chopped gradient dominates the violation of the Karush–Kuhn–Tucker conditions. The formula for the best bound η_Γ^{opt} which corresponds to $\Gamma = 1$ and $\overline{\alpha} = \|\mathsf{A}\|^{-1}$ reads

$$\eta_\Gamma^{opt} = 1 - \kappa(\mathsf{A})^{-1}/4, \tag{5.83}$$

where $\kappa(\mathsf{A})$ denotes the spectral condition number of A.

5.8 Modified Proportioning with Reduced Gradient Projections

Even though the MPGP algorithm of the previous section combines the conjugate gradient method with the gradient projections in a way which enables to prove its linear rate of convergence that can be expressed in terms of bounds on the spectrum of the Hessian of f, there is still room for improvements. The reason is that the gradient projection at the same time adds and removes the indices from the active set, so the algorithm releases the indices from the active set rather randomly. The result is that MPGP may not exploit fully the self-preconditioning effect of the conjugate gradient method [168] and can suffer from the oscillations often attributed to the iterative active set methods. In this section we show that these drawbacks can be relieved if we replace the gradient projection step by the *free gradient projection* with a fixed steplength $\overline{\alpha}$. We show that the modified algorithm not only preserves the linear rate of convergence of the cost function, but *it has the finite termination property even for dual degenerate QP problems* with zero Lagrange multipliers corresponding to the active constraints and the *R-linear rate of convergence in the norm of projected gradient.*

5.8.1 MPRGP Schema

The algorithm that we propose here exploits a constant $\Gamma > 0$ defined by a user, a test to decide when to leave the face, and three types of steps. The test and two of the three steps, the conjugate gradient step and the proportioning step, are exactly those introduced in Sect. 5.7.1.

The gradient projection step is replaced by the *expansion step* defined by the free gradient projection

$$\mathbf{x}^{k+1} = P_{\Omega_B}\big(\mathbf{x}^k - \overline{\alpha}\boldsymbol{\varphi}(\mathbf{x}^k)\big) = \max\{\boldsymbol{\ell}, \mathbf{x}^k - \overline{\alpha}\boldsymbol{\varphi}(\mathbf{x}^k)\} \tag{5.84}$$

with the fixed steplength. This step expands the current working set. To describe it in the form suitable for analysis, let us recall, for any $\mathbf{x} \in \Omega_B$ and $\alpha > 0$, that the *reduced free gradient* $\widetilde{\boldsymbol{\varphi}}_\alpha(\mathbf{x})$ is defined by the entries

$$\widetilde{\varphi}_i = \widetilde{\varphi}_i(\mathbf{x}, \alpha) = \min\{(x_i - \ell_i)/\alpha, \varphi_i\}, \quad i \in \mathcal{N} = \{1, \ldots, n\}, \tag{5.85}$$

so that

$$P_{\Omega_B}\big(\mathbf{x} - \alpha\boldsymbol{\varphi}(\mathbf{x})\big) = \mathbf{x} - \alpha\widetilde{\boldsymbol{\varphi}}_\alpha(\mathbf{x}). \tag{5.86}$$

Using the new notation, we can write also

$$P_{\Omega_B}\big(\mathbf{x} - \alpha\mathbf{g}(\mathbf{x})\big) = \mathbf{x} - \alpha\big(\widetilde{\boldsymbol{\varphi}}_\alpha(\mathbf{x}) + \boldsymbol{\beta}(\mathbf{x})\big). \tag{5.87}$$

Now we are ready to define the algorithm in the form that is convenient for analysis, postponing the discussion about implementation to the next section. Notice that we admit the fixed steplength $\overline{\alpha} = 2\|\mathsf{A}\|^{-1}$ which guarantees neither the contraction of the distance from the solution nor the decrease of the cost function in the expansion steps.

Algorithm 5.7. Modified proportioning with reduced gradient projections (MPRGP schema).

Given a symmetric positive definite matrix $\mathsf{A} \in \mathbb{R}^{n\times n}$ *and* n*-vectors* $\mathbf{b}$, $\boldsymbol{\ell}$. *Choose* $\mathbf{x}^0 \in \Omega_B$, $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$, *and* $\Gamma > 0$. *Set* $k = 0$. *For* $k \geq 0$ *and* $\mathbf{x}^k$ *known, choose* $\mathbf{x}^{k+1}$ *by the following rules:*

(i) If $\mathbf{g}^P(\mathbf{x}^k) = \mathbf{o}$, *set* $\mathbf{x}^{k+1} = \mathbf{x}^k$.

(ii) If $\mathbf{x}^k$ *is strictly proportional and* $\mathbf{g}^P(\mathbf{x}^k) \neq \mathbf{o}$, *try to generate* $\mathbf{x}^{k+1}$ *by the conjugate gradient step. If* $\mathbf{x}^{k+1} \in \Omega_B$, *then accept it, else generate* $\mathbf{x}^{k+1}$ *by the expansion step.*

(iii) If $\mathbf{x}^k$ *is not strictly proportional, define* $\mathbf{x}^{k+1}$ *by proportioning.*

Proposition 5.12. *Let* $\{\mathbf{x}^k\}$ *be generated by Algorithm 5.7 with* $\mathbf{x}^0 \in \Omega_B$, $\Gamma > 0$, *and* $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$. *Then* $\{\mathbf{x}^k\}$ *converges to the solution* $\{\widehat{\mathbf{x}}\}$ *and* $\{\mathbf{g}^P(\mathbf{x}^k)\}$ *converges to zero.*

Proof. MPRGP is a variant of the proportioning algorithm studied in [42]; it converges when each iterate $\mathbf{x}^{k+1}$ generated by the expansion step satisfies

$$f(\mathbf{x}^{k+1}) - f(\mathbf{x}^k) \leq 0.$$

This condition is satisfied by Proposition 5.10 for $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$; the convergence is driven by the proportioning step, which is a spacer iteration (see, e.g., Bertsekas [12]). The second statement is an easy corollary of the identification lemma 5.17 and of the continuity of $\mathbf{g}(\mathbf{x})$. □

5.8.2 Rate of Convergence

The main tool of our analysis is the quadratic function

$$F(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{x} - \mathbf{c}^T\mathbf{x} + d, \quad \mathbf{x}, \mathbf{c} \in \mathbb{R}^n, \quad \mathbf{c} = [c_i], \quad d \in \mathbb{R}, \tag{5.88}$$

and its properties similar to those developed in Sect. 5.6.5. In particular,

$$F(\mathbf{x}) = \sum_{i=1}^{n} F_i(x_i) + d, \quad F_i(x_i) = \frac{1}{2}x_i^2 - c_i x_i, \quad \mathbf{x} = [x_i]. \tag{5.89}$$

If $\mathbf{x} \in \mathbb{R}^n$ is arbitrary but fixed, we associate with f and $\delta \in (0, \|\mathsf{A}\|^{-1}]$ the quadratic function of the form (5.88)

$$F_\delta(\mathbf{y}) = \delta f(\mathbf{y}) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T(\mathsf{I} - \delta\mathsf{A})(\mathbf{y} - \mathbf{x}) \geq \delta f(\mathbf{y}). \tag{5.90}$$

It is defined so that

$$F_\delta(\mathbf{x}) = \delta f(\mathbf{x}), \quad \nabla F_\delta(\mathbf{x}) = \delta\nabla f(\mathbf{x}) = \delta\mathbf{g}, \quad \text{and} \quad \nabla^2 F_\delta(\mathbf{y}) = \mathsf{I}. \tag{5.91}$$

We need the following lemma which is analogous to Corollary 5.8.

Lemma 5.13. *Let* $\mathbf{x}, \boldsymbol{\ell}, \mathbf{c} \in \mathbb{R}^n$, $\mathbf{x} \geq \boldsymbol{\ell}$. *Let* F *be defined by* (5.88). *Then for any* $\delta \in [0,1]$

$$F\big(P_{\Omega_B}(\mathbf{x} - (2-\delta)\boldsymbol{\varphi}(\mathbf{x}))\big) \leq F\big(P_{\Omega_B}(\mathbf{x} - \delta\boldsymbol{\varphi}(\mathbf{x}))\big). \tag{5.92}$$

Proof. First recall that P_{Ω_B} is separable and can be defined componentwise by $P_i(y) = \max\{y, \ell_i\}$, $i = 1, \ldots, n$, $y \in \mathbb{R}$. Denoting $\mathcal{F}$, $\mathcal{A}$, and g_i the free set of $\mathbf{x}$, the active set of $\mathbf{x}$, and the components of the gradient $\mathbf{g}(\mathbf{x})$, respectively, we can use the representation of F given by (5.89) and Lemma 5.7 to get

$$\begin{aligned}
F\big(P_{\Omega_B}(\mathbf{x} - (2-\delta)\boldsymbol{\varphi}(\mathbf{x}))\big) &= \sum_{i=1}^{n} F_i\big([P_{\Omega_B}(\mathbf{x} - (2-\delta)\boldsymbol{\varphi}(\mathbf{x}))]_i\big) + d \\
&= \sum_{i\in\mathcal{F}} F_i\big(P_i(x_i - (2-\delta)g_i)\big) + \sum_{i\in\mathcal{A}} F_i\big(P_i(x_i)\big) + d \\
&\leq \sum_{i\in\mathcal{F}} F_i\big(P_i(x_i - \delta g_i)\big) + \sum_{i\in\mathcal{A}} F_i\big(P_i(x_i)\big) + d \\
&= F\big(P_{\Omega_B}(\mathbf{x} - \delta\boldsymbol{\varphi}(\mathbf{x}))\big). \qquad \square
\end{aligned}$$

Now we are ready to prove the R-linear rate of convergence of MPRGP.

Theorem 5.14. *Let* $\{\mathbf{x}^k\}$ *be generated by Algorithm 5.7 with* $\mathbf{x}^0 \in \Omega_B$, $\Gamma > 0$, *and* $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$. *Then*

$$f(\mathbf{x}^{k+1}) - f(\widehat{\mathbf{x}}) \leq \eta_\Gamma \left(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\right), \tag{5.93}$$

where $\widehat{\mathbf{x}}$ *denotes a unique solution of* (5.1),

$$\eta_\Gamma = 1 - \frac{\widehat{\alpha}\lambda_{\min}}{\vartheta + \vartheta\widehat{\Gamma}^2}, \qquad \widehat{\Gamma} = \max\{\Gamma, \Gamma^{-1}\}, \tag{5.94}$$

$$\vartheta = 2\max\{\overline{\alpha}\|\mathsf{A}\|, 1\}, \quad \widehat{\alpha} = \min\{\overline{\alpha}, 2\|\mathsf{A}\|^{-1} - \overline{\alpha}\}, \tag{5.95}$$

and $\lambda_{\min}$ *denotes the smallest eigenvalue of* A. *The error in the* A*-norm is bounded by*

$$\|\mathbf{x}^k - \widehat{\mathbf{x}}\|_{\mathsf{A}}^2 \leq 2\eta_\Gamma^k \left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\right). \tag{5.96}$$

Proof. First observe that the only new type of iteration, as compared with MPGP of Sect. 5.7, is the expansion step. Moreover, the estimate (5.68) with η_Γ defined by (5.69) of Theorem 5.11 is the same as our estimate (5.93) with η_Γ defined by (5.94). Thus we can reduce our analysis to the expansion step. Our main tools are again (5.54) and the inequality

$$f\Big(P_{\Omega_B}\left(\mathbf{x}^k - \widehat{\alpha}\mathbf{g}(\mathbf{x}^k)\right)\Big) \geq f(\mathbf{x}^k) - \widehat{\alpha}\big(\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T\boldsymbol{\varphi}(\mathbf{x}^k) + \|\boldsymbol{\beta}(\mathbf{x}^k)\|^2\big), \tag{5.97}$$

which can be obtained by the Taylor expansion and (5.87).

Let us first assume that $\|\mathsf{A}\| = 1$ and let $\mathbf{x}^{k+1}$ be generated by the expansion step (5.84). Using in sequence the definition of the dominating function (5.90) associated with $\mathbf{x} = \mathbf{x}^k$, Lemma 5.13, the assumption,

$\|\mathsf{A}\| = 1$ and $\widehat{\alpha} \leq 1$ with (5.57), the Taylor expansion with (5.86), (5.91), $\|\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)\|^2 \leq \widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k)$, and simple manipulations, we get

$$\begin{aligned}
\widehat{\alpha} f(\mathbf{x}^{k+1}) &\leq \widehat{\alpha} F_1(\mathbf{x}^{k+1}) = \widehat{\alpha} F_1\big(P_{\Omega_B}\left(\mathbf{x}^k - \overline{\alpha}\boldsymbol{\varphi}(\mathbf{x}^k)\right)\big) \\
&\leq \widehat{\alpha} F_1\big(P_{\Omega_B}\left(\mathbf{x}^k - \widehat{\alpha}\boldsymbol{\varphi}(\mathbf{x}^k)\right)\big) \leq F_{\widehat{\alpha}}\big(P_{\Omega_B}\left(\mathbf{x}^k - \widehat{\alpha}\boldsymbol{\varphi}(\mathbf{x}^k)\right)\big) \\
&= F_{\widehat{\alpha}}\left(\mathbf{x}^k\right) - \widehat{\alpha}^2 \widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k) + \frac{\widehat{\alpha}^2}{2}\|\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)\|^2 \\
&\leq F_{\widehat{\alpha}}\left(\mathbf{x}^k\right) - \frac{\widehat{\alpha}^2}{2}\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k) = \widehat{\alpha} f(\mathbf{x}^k) - \frac{\widehat{\alpha}^2}{2}\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k).
\end{aligned}$$

Thus

$$f(\mathbf{x}^{k+1}) \leq f(\mathbf{x}^k) - \frac{\widehat{\alpha}}{2}\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k). \tag{5.98}$$

The expansion step is used only when $\mathbf{x}^k$ is strictly proportional, i.e.,

$$\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2 \leq \Gamma^2 \widetilde{\boldsymbol{\varphi}}_{\overline{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k).$$

Since $\widehat{\alpha} \leq \overline{\alpha}$ by the definition, it follows that

$$\widetilde{\boldsymbol{\varphi}}_{\overline{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k) \leq \widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k)$$

and

$$\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2 \leq \Gamma^2 \widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k). \tag{5.99}$$

After substituting (5.99) into (5.97), we get

$$f\big(P_{\Omega_B}\left(\mathbf{x}^k - \widehat{\alpha}\mathbf{g}(\mathbf{x}^k)\right)\big) \geq f(\mathbf{x}^k) - \widehat{\alpha}(1+\Gamma^2)\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k). \tag{5.100}$$

Thus for $\mathbf{x}^{k+1}$ generated by the expansion step, we get by elementary algebra and application of (5.100) that

$$\begin{aligned}
f(\mathbf{x}^{k+1}) &\leq f(\mathbf{x}^k) - \frac{\widehat{\alpha}}{2}\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k) \\
&= \frac{1}{2+2\Gamma^2}\big(f(\mathbf{x}^k) - \widehat{\alpha}(1+\Gamma^2)\widetilde{\boldsymbol{\varphi}}_{\widehat{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k) + (1+2\Gamma^2) f(\mathbf{x}^k)\big) \\
&\leq \frac{1}{2+2\Gamma^2}\big(f\left(P_{\Omega_B}\left(\mathbf{x}^k - \widehat{\alpha}\mathbf{g}(\mathbf{x}^k)\right)\right) + (1+2\Gamma^2) f(\mathbf{x}^k)\big).
\end{aligned}$$

Inserting $-f(\widehat{\mathbf{x}}) + f(\widehat{\mathbf{x}})$ into the last term and substituting (5.54) with $\mathbf{x} = \mathbf{x}^k$ and $\overline{\alpha} = \widehat{\alpha}$ into the last expression, we get

$$\begin{aligned}
f(\mathbf{x}^{k+1}) &\leq \frac{\eta_f + 1 + 2\Gamma^2}{2+2\Gamma^2} f(\mathbf{x}^k) + \frac{1-\eta_f}{2+2\Gamma^2} f(\widehat{\mathbf{x}}) \\
&= \frac{\eta_f + 1 + 2\Gamma^2}{2+2\Gamma^2}\left(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\right) + f(\widehat{\mathbf{x}}).
\end{aligned} \tag{5.101}$$

The proof of (5.93) for $\|\mathsf{A}\| = 1$ is completed by

$$\frac{\eta_f + 1 + 2\Gamma^2}{2 + 2\Gamma^2} = \frac{\eta_f - 1 + 2 + 2\Gamma^2}{2 + 2\Gamma^2} = 1 - \frac{1 - \eta_f}{2 + 2\Gamma^2} = 1 - \frac{\widehat{\alpha}\lambda_{\min}}{2 + 2\Gamma^2} \leq \eta_\Gamma.$$

To prove the general case, it is enough to apply the theorem to $h = \|\mathsf{A}\|^{-1} f$.

To get the error bound (5.96), notice that by Lemma 5.1

$$\|\mathbf{x}^k - \widehat{\mathbf{x}}\|_{\mathsf{A}}^2 \leq 2\left(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\right) \leq 2\eta_\Gamma^k \left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\right). \tag{5.102}$$

□

The formula for the best bound η_Γ^{opt} is given by (5.83). Notice that the coefficient of the Euclidean contraction η_E defined by (5.34) is smaller than η_Γ and by (5.38) guarantees faster convergence in the energy norm. Does it follow that the gradient projection method is faster than MPRGP? The answer is *no*. We have got both estimates by the worst case analysis of just one step of each method. Such analysis at least partly enhances the improvement due to the long sequence of the same type of iterations of the projected gradient method, while this is not true in the case of MPRGP; the worst case assumes that the algorithm switches the types of iterations. The error in energy norm need not even decrease in one step of the gradient projection method.

5.8.3 Rate of Convergence of Projected Gradient

To use the MPRGP algorithm in the inner loops of other algorithms, we must be able to *recognize* when we are near the solution. There is a catch – though by Lemma 5.1 the latter can be tested by a norm of the projected gradient, Theorem 5.14 does not guarantee that such test is positive near the solution. The projected gradient is not a continuous function of the iterates! A large projected gradient near the solution is in Fig. 5.14. The R-linear convergence of the projected gradient is treated by the following theorem.

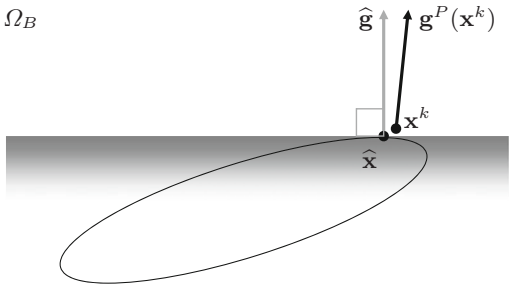

Fig. 5.14. Large projected gradient near the solution

Theorem 5.15. *Let* $\{\mathbf{x}^k\}$ *be generated by Algorithm 5.7 with* $\mathbf{x}^0 \in \Omega_B$, $\Gamma > 0$, *and* $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$. *Let* $\widehat{\mathbf{x}}$ *denote the unique solution of* (5.1) *and let* $\widehat{\Gamma}$, η_Γ, $\widehat{\alpha}$, *and* ϑ *be those of Theorem 5.14.*

Then for any $k \geq 1$

$$\|\mathbf{g}^P(\mathbf{x}^{k+1})\|^2 \leq a_1 \eta_\Gamma^k \left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\right) \tag{5.103}$$

with

$$a_1 = \frac{38}{\widehat{\alpha}(1-\eta_\Gamma)} = \frac{38\vartheta(1+\widehat{\Gamma}^2)}{\widehat{\alpha}^2\lambda_{\min}}. \tag{5.104}$$

Proof. First notice that it is enough to estimate separately $\boldsymbol{\beta}(\mathbf{x}^k)$ and $\boldsymbol{\varphi}(\mathbf{x}^k)$ as

$$\|\mathbf{g}^P(\mathbf{x}^k)\|^2 = \|\boldsymbol{\beta}(\mathbf{x}^k)\|^2 + \|\boldsymbol{\varphi}(\mathbf{x}^k)\|^2.$$

In particular, since $\widehat{\alpha} \leq \|\mathsf{A}^{-1}\|$, we have for any vector $\mathbf{d}$ which satisfies $\mathbf{d}^T\mathbf{g}(\mathbf{x}) \geq \|\mathbf{d}\|^2$

$$f(\mathbf{x}) - f(\mathbf{x} - \widehat{\alpha}\mathbf{d}) = \widehat{\alpha}\mathbf{d}^T\mathbf{g}(\mathbf{x}) - \frac{1}{2}\widehat{\alpha}^2\mathbf{d}^T\mathsf{A}\mathbf{d} \geq \frac{\widehat{\alpha}}{2}\|\mathbf{d}\|^2. \tag{5.105}$$

It follows that we can combine (5.105) with

$$\mathbf{x}^k - \widehat{\alpha}\boldsymbol{\beta}(\mathbf{x}^k) \geq \boldsymbol{\ell}$$

to estimate $\|\boldsymbol{\beta}(\mathbf{x}^k)\|$ by

$$\begin{aligned} f(\mathbf{x}^k) - f(\widehat{\mathbf{x}}) &= \left(f(\mathbf{x}^k) - f\left(\mathbf{x}^k - \widehat{\alpha}\boldsymbol{\beta}(\mathbf{x}^k)\right)\right) + \left(f\left(\mathbf{x}^k - \widehat{\alpha}\boldsymbol{\beta}(\mathbf{x}^k)\right) - f(\widehat{x})\right) \\ &\geq f(\mathbf{x}^k) - f\left(\mathbf{x}^k - \widehat{\alpha}\boldsymbol{\beta}(\mathbf{x}^k)\right) \geq \frac{\widehat{\alpha}}{2}\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2. \end{aligned} \tag{5.106}$$

Applying (5.93), we get

$$\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2 \leq \frac{2}{\widehat{\alpha}}\left(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\right) \leq \frac{2\eta_\Gamma^k}{\widehat{\alpha}}\left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\right). \tag{5.107}$$

To estimate $\|\boldsymbol{\varphi}(\mathbf{x}^k)\|$, notice that the algorithm "does not know" about the components of the constraint vector $\boldsymbol{\ell}$ when it generates $\mathbf{x}^{k+1}$ unless their indices belong to $\mathcal{A}(\mathbf{x}^k)$ or $\mathcal{A}(\mathbf{x}^{k+1})$. It follows that $\mathbf{x}^{k+1}$ may be considered also as an iterate generated by Algorithm 5.7 from $\mathbf{x}^k$ for the problem

$$\text{minimize} \quad f(\mathbf{x}) \quad \text{subject to} \quad x_i \geq \ell_i \text{ for } i \in \mathcal{A}(\mathbf{x}^k) \cup \mathcal{A}(\mathbf{x}^{k+1}). \tag{5.108}$$

If we denote

$$\overline{f}^k = \min\{f(\mathbf{x}) : \ x_i \geq \ell_i \text{ for } i \in \mathcal{A}(\mathbf{x}^k) \cup \mathcal{A}(\mathbf{x}^{k+1})\} \leq f(\widehat{\mathbf{x}})$$

and $\overline{\delta}_k = f(\widehat{\mathbf{x}}) - \overline{f}^k \geq 0$, we can use (5.93) to get

$$\begin{aligned}\overline{\delta}_k &= f(\widehat{\mathbf{x}}) - \overline{f}^k \le f(\mathbf{x}^{k+1}) - \overline{f}^k \le \eta_\Gamma \left(f(\mathbf{x}^k) - \overline{f}^k \right) \\ &= \eta_\Gamma \left(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}}) \right) + \eta_\Gamma \overline{\delta}_k,\end{aligned}$$

so that

$$\overline{\delta}_k \le \frac{\eta_\Gamma}{1-\eta_\Gamma} \left(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}}) \right) \le \frac{\eta_\Gamma^{k+1}}{1-\eta_\Gamma} \left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}}) \right). \tag{5.109}$$

Now observe that the indices of the unconstrained components of the minimization problem (5.108) are those belonging to $\mathcal{I}^k = \mathcal{F}(\mathbf{x}^k) \cap \mathcal{F}(\mathbf{x}^{k+1})$ as

$$\begin{aligned}\mathcal{I}^k = \mathcal{F}(\mathbf{x}^k) \cap \mathcal{F}(\mathbf{x}^{k+1}) &= \left(\mathcal{N} \setminus \mathcal{A}(\mathbf{x}^k) \right) \cap \left(\mathcal{N} \setminus \mathcal{A}(\mathbf{x}^{k+1}) \right) \\ &= \mathcal{N} \setminus \left(\mathcal{A}(\mathbf{x}^k) \cup \mathcal{A}(\mathbf{x}^{k+1}) \right).\end{aligned}$$

It follows that if $\mathcal{I}^k$ is nonempty, then by the definition of $\overline{\delta}_k$ and (5.105)

$$\overline{\delta}_k \ge f(\widehat{\mathbf{x}}) - f\left(\widehat{\mathbf{x}} - \widehat{\alpha} \mathbf{g}_{\mathcal{I}^k}(\widehat{\mathbf{x}}) \right) \ge \frac{\widehat{\alpha}}{2} \|\mathbf{g}_{\mathcal{I}^k}(\widehat{\mathbf{x}})\|^2. \tag{5.110}$$

For convenience, let us define $\mathbf{g}_{\mathcal{I}}(\mathbf{x}) = \mathbf{o}$ for any $\mathbf{x} \in \mathbb{R}^n$ and empty set $\mathcal{I} = \emptyset$. Then (5.110) remains valid for $\mathcal{I}^k = \emptyset$, so that we can combine it with (5.109) to get

$$\|\mathbf{g}_{\mathcal{I}^k}(\widehat{\mathbf{x}})\|^2 \le \frac{2}{\widehat{\alpha}} \overline{\delta}_k \le \frac{2\eta_\Gamma^{k+1}}{\widehat{\alpha}(1-\eta_\Gamma)} \big(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}}) \big). \tag{5.111}$$

Since our algorithm is defined so that either $\mathcal{I}^k = \mathcal{F}(\mathbf{x}^k) \subseteq \mathcal{F}(\mathbf{x}^{k+1})$ or $\mathcal{I}^k = \mathcal{F}(\mathbf{x}^{k+1}) \subseteq \mathcal{F}(\mathbf{x}^k)$, it follows that either

$$\begin{aligned}\|\mathbf{g}_{\mathcal{F}(\mathbf{x}^k)}(\widehat{\mathbf{x}})\|^2 = \|\mathbf{g}_{\mathcal{I}^k}(\widehat{\mathbf{x}})\|^2 &\le \frac{2\eta_\Gamma^{k+1}}{\widehat{\alpha}(1-\eta_\Gamma)} (f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})) \\ &\le \frac{2\eta_\Gamma^{k}}{\widehat{\alpha}(1-\eta_\Gamma)} (f(\mathbf{x}^0) - f(\widehat{\mathbf{x}}))\end{aligned} \tag{5.112}$$

or

$$\|\mathbf{g}_{\mathcal{F}(\mathbf{x}^{k+1})}(\widehat{\mathbf{x}})\|^2 = \|\mathbf{g}_{\mathcal{I}^k}(\widehat{\mathbf{x}})\|^2 \le \frac{2\eta_\Gamma^{k+1}}{\widehat{\alpha}(1-\eta_\Gamma)} (f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})).$$

Using the same reasoning for $\mathbf{x}^{k-1}$ and $\mathbf{x}^k$, we conclude that the estimate (5.112) is valid for any $\mathbf{x}^k$ such that

$$\mathcal{F}(\mathbf{x}^{k-1}) \supseteq \mathcal{F}(\mathbf{x}^k) \quad \text{or} \quad \mathcal{F}(\mathbf{x}^k) \subseteq \mathcal{F}(\mathbf{x}^{k+1}). \tag{5.113}$$

Let us now recall that by Lemma 5.1 and (5.96)

$$\begin{aligned}\|\mathbf{g}(\mathbf{x}^k) - \mathbf{g}(\widehat{\mathbf{x}})\|^2 = \|\mathsf{A}(\mathbf{x}^k - \widehat{\mathbf{x}})\|^2 &\le \|\mathsf{A}\| \|\mathbf{x}^k - \widehat{\mathbf{x}}\|_{\mathsf{A}}^2 \le 2\|\mathsf{A}\| \left(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}}) \right) \\ &\le \frac{2}{\widehat{\alpha}} \eta_\Gamma^k \left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}}) \right),\end{aligned} \tag{5.114}$$

so that for any k satisfying the relations (5.113), we get

$$\begin{aligned}\|\boldsymbol{\varphi}(\mathbf{x}^k)\| = \|\mathbf{g}_{\mathcal{F}(\mathbf{x}^k)}(\mathbf{x}^k)\| &\le \|\mathbf{g}_{\mathcal{F}(\mathbf{x}^k)}(\mathbf{x}^k) - \mathbf{g}_{\mathcal{F}(\mathbf{x}^k)}(\widehat{\mathbf{x}})\| + \|\mathbf{g}_{\mathcal{F}(\mathbf{x}^k)}(\widehat{\mathbf{x}})\| \\ &\le \sqrt{\frac{2}{\widehat{\alpha}}\eta_\Gamma^k\big(f(x^0) - f(\widehat{\mathbf{x}})\big)} + \sqrt{\frac{2}{\widehat{\alpha}(1-\eta_\Gamma)}\eta_\Gamma^k\big(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\big)} \\ &\le 2\sqrt{\frac{2}{\widehat{\alpha}(1-\eta_\Gamma)}\eta_\Gamma^k\big(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\big)}.\end{aligned}$$

Combining the last inequality with (5.107), we get for any k satisfying the relations (5.113) that

$$\|\mathbf{g}^P(\mathbf{x}^k)\|^2 = \|\boldsymbol{\beta}(\mathbf{x}^k)\|^2 + \|\boldsymbol{\varphi}(\mathbf{x}^k)\|^2 \le \frac{10}{\widehat{\alpha}(1-\eta_\Gamma)}\eta_\Gamma^k\big(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\big). \quad (5.115)$$

Now notice that the estimate (5.115) is valid for any iterate $\mathbf{x}^k$ which satisfies $\mathcal{F}(\mathbf{x}^{k-1}) \supseteq \mathcal{F}(\mathbf{x}^k)$, i.e., when $\mathbf{x}^k$ is generated by the conjugate gradient step or the expansion step. Thus it remains to estimate the projected gradient of the iterate $\mathbf{x}^k$ generated by the proportioning step. In this case

$$\mathcal{F}(\mathbf{x}^{k-1}) \subseteq \mathcal{F}(\mathbf{x}^k),$$

so that we can use the estimate (5.115) to get

$$\|\mathbf{g}^P(\mathbf{x}^{k-1})\| \le \sqrt{\frac{10}{\widehat{\alpha}(1-\eta_\Gamma)}\eta_\Gamma^{k-1}\big(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\big)}. \quad (5.116)$$

Since the proportioning step is defined by $\mathbf{x}^k = \mathbf{x}^{k-1} - \alpha_{cg}\boldsymbol{\beta}(\mathbf{x}^{k-1})$, it follows that

$$\|\mathbf{g}_{\mathcal{F}(\mathbf{x}^k)}(\mathbf{x}^{k-1})\| = \|\mathbf{g}^P(\mathbf{x}^{k-1})\|.$$

Moreover, using the basic properties of the norm, we get

$$\begin{aligned}\|\boldsymbol{\varphi}(\mathbf{x}^k)\| = \|\mathbf{g}_{\mathcal{F}(\mathbf{x}^k)}(\mathbf{x}^k)\| &\le \|\mathbf{g}_{\mathcal{F}(\mathbf{x}^k)}(\mathbf{x}^k) - \mathbf{g}_{\mathcal{F}(\mathbf{x}^k)}(\mathbf{x}^{k-1})\| + \|\mathbf{g}_{\mathcal{F}(\mathbf{x}^k)}(\mathbf{x}^{k-1})\| \\ &\le \|\mathbf{g}(\mathbf{x}^k) - \mathbf{g}(\widehat{\mathbf{x}})\| + \|\mathbf{g}(\widehat{\mathbf{x}}) - \mathbf{g}(\mathbf{x}^{k-1})\| + \|\mathbf{g}^P(\mathbf{x}^{k-1})\|,\end{aligned}$$

and by (5.114) and (5.116)

$$\begin{aligned}\|\boldsymbol{\varphi}(\mathbf{x}^k)\| &\le \sqrt{\frac{2}{\widehat{\alpha}}\eta_\Gamma^k\big(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\big)} + \sqrt{\frac{2}{\widehat{\alpha}}\eta_\Gamma^{k-1}\big(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\big)} \\ &\quad + \sqrt{\frac{10}{\widehat{\alpha}(1-\eta_\Gamma)}\eta_\Gamma^{k-1}\big(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\big)} \\ &\le (\sqrt{5}+2)\sqrt{\frac{2}{\widehat{\alpha}(1-\eta_\Gamma)}\eta_\Gamma^{k-1}\big(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\big)}.\end{aligned}$$

Combining the last inequality with (5.107), we get by simple computation that

$$\|\mathbf{g}^P(\mathbf{x}^k)\|^2 = \|\boldsymbol{\varphi}(\mathbf{x}^k)\|^2 + \|\boldsymbol{\beta}(\mathbf{x}^k)\|^2 \le \frac{38}{\widehat{\alpha}(1-\eta_\Gamma)}\eta_\Gamma^{k-1}\left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\right).$$

Since the last estimate is obviously weaker than (5.115), it follows that (5.103) is valid for all indices k. □

The bound on the rate of convergence as given by (5.103) is rather poor. The reason is that it has been obtained by the worst case analysis of a general couple of consecutive iterations and does not reflect the structure of a longer chain of the same type of iterations. Recall that Fig. 5.14 shows that no bound can be obtained by the analysis of a single iteration!

5.8.4 Optimality

Theorems 5.14 and 5.15 give the bounds on the rates of convergence of the iterates and corresponding projected gradients that depend only on the bounds on the spectrum, but do not depend on the constraint vector $\boldsymbol{\ell}$. It simply follows that if we have a class of bound constrained problems with the spectrum of the Hessian of the cost function in an a priori fixed interval, then the rate of convergence of the MPRGP algorithm can be bounded uniformly for the whole class. To present explicitly this feature of Algorithm 5.7, let $\mathcal{T}$ denote any set of indices and assume that for any $t \in \mathcal{T}$ there is defined a problem

$$\text{minimize} \quad f_t(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \in \Omega_{B_t} \tag{5.117}$$

with $\Omega_{B_t} = \{\mathbf{x} \in \mathbb{R}^{n_t} : \mathbf{x} \ge \boldsymbol{\ell}_t\}$, $f_t(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathsf{A}_t\mathbf{x} - \mathbf{b}_t^T\mathbf{x}$, $\mathsf{A}_t \in \mathbb{R}^{n_t \times n_t}$ symmetric positive definite, and $\boldsymbol{\ell}_t \in \mathbb{R}^{n_t}$. Our optimality result then reads as follows.

Theorem 5.16. *Let $a_{\max} > a_{\min} > 0$ denote given constants and let $\{\mathbf{x}_t^k\}$ be generated by Algorithm 5.7 for the solution of the bound constrained problem (5.117) with $0 < \overline{\alpha} \le 2a_{\max}^{-1}$ and $\Gamma > 0$ starting from $\mathbf{x}_t^0 = \max\{\mathbf{o}, \boldsymbol{\ell}_t\}$. Let the class of problems (5.117) satisfy*

$$a_{\min} \le \lambda_{\min}(\mathsf{A}_t) \le \lambda_{\max}(\mathsf{A}_t) \le a_{\max},$$

where $\lambda_{\min}(\mathsf{A}_t)$ and $\lambda_{\max}(\mathsf{A}_t)$ denote respectively the smallest and the largest eigenvalues of A_t.

Then there are integers $\overline{k}$ and $\overline{\ell}$ such that for any $t \in \mathcal{T}$ and $\varepsilon > 0$

$$\|\mathbf{g}_t^P(\mathbf{x}_t^{\overline{k}})\| \le \varepsilon\|\mathbf{g}_t^P(\mathbf{x}_t^0)\|$$

and

$$f_t(\mathbf{x}_t^{\overline{\ell}}) - f_t(\widehat{\mathbf{x}}_t) \le \varepsilon\left(f_t(\mathbf{x}_t^0) - f(\widehat{\mathbf{x}}_t)\right).$$

Proof. First denote

$$\eta_\Gamma^t = 1 - \frac{\widehat{\alpha}\lambda_{\min}^t}{\vartheta + \vartheta\widehat{\Gamma}^2}, \qquad \overline{\eta}_\Gamma = 1 - \frac{\widehat{\alpha}a_{\min}}{\vartheta + \vartheta\widehat{\Gamma}^2},$$
$$a_1^t = \frac{38\vartheta(1+\widehat{\Gamma}^2)}{\widehat{\alpha}^2\lambda_{\min}^t}, \qquad \overline{a}_1 = \frac{38\vartheta(1+\widehat{\Gamma}^2)}{\widehat{\alpha}^2 a_{\min}},$$

where $\widehat{\Gamma} = \max\{\Gamma, \Gamma^{-1}\}$, so that

$$\eta_\Gamma^t \le \overline{\eta}_\Gamma < 1 \qquad \text{and} \qquad a_1^t \le \overline{a}_1.$$

Combining these estimates with Theorem 5.15 and inequality (5.5), we get for any $k \ge 1$

$$\|\mathbf{g}_t^P(\mathbf{x}^{k+1})\|^2 \le \overline{a}_1\overline{\eta}_\Gamma^k \left(f_t(\mathbf{x}_t^0) - f_t(\widehat{\mathbf{x}}_t)\right) \le \frac{\overline{a}_1}{a_{\min}}\overline{\eta}_\Gamma^k\|\mathbf{g}_t^P(\mathbf{x}_t^0)\|^2.$$

Similarly, using Theorem 5.14, we get

$$f_t(\mathbf{x}_t^k) - f_t(\widehat{\mathbf{x}}_t) \le \overline{\eta}_\Gamma^k \left(f_t(\mathbf{x}_t^0) - f(\widehat{\mathbf{x}}_t)\right).$$

To finish the proof, it is enough to take $\overline{k}$ and $\overline{\ell}$ so that

$$\frac{\overline{a}_1}{a_{\min}}\overline{\eta}_\Gamma^{\overline{k}-1} \le \varepsilon \qquad \text{and} \qquad \overline{\eta}_\Gamma^{\overline{\ell}} \le \varepsilon. \qquad \square$$

5.8.5 Identification Lemma and Finite Termination

Let us consider the conditions which guarantee that the MPRGP algorithm finds the solution $\widehat{\mathbf{x}}$ of (5.1) in a finite number of steps. There are at least two reasons to consider such results important. First the algorithm with the finite termination property is less likely to suffer from the oscillations that are often attributed to the working set-based algorithms as it is less likely to reexamine the working sets; if any working set reappears, it can happen "only" finitely many times. The second reason is that such algorithm is more likely to generate longer sequences of the conjugate gradient iterations. Thus the reduction of the cost function value is bounded by the "global" estimate (3.21), and finally switches to the conjugate gradient method, so that it can exploit its nice self-acceleration property [168]. It is difficult to enhance these characteristics of the algorithm into the rate of convergence as they cannot be obtained by the analysis of just one step of the method.

We first examine the finite termination of Algorithm 5.7 in a simpler case when the solution $\widehat{\mathbf{x}}$ of (5.1) is *regular*, i.e., the vector of Lagrange multipliers $\widehat{\lambda}$ of the solution satisfies the *strict complementarity condition* $\widehat{\lambda}_i > 0$ for $i \in \mathcal{A}(\widehat{\mathbf{x}})$. The proof is based on simple geometrical observations. For example, examining Fig. 5.15, it is easy to see that the free sets of the iterates $\mathbf{x}^k$ soon contain the free set of the solution $\widehat{\mathbf{x}}$. The formal analysis of such observations is a subject of the following identification lemma.

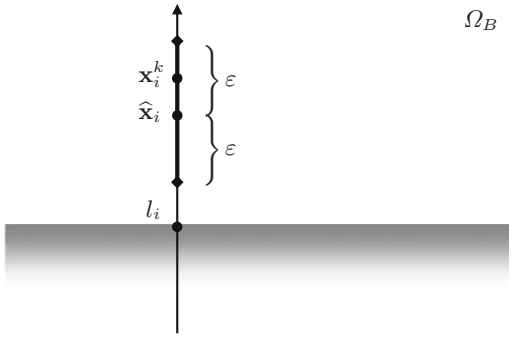

Fig. 5.15. Identification of the free set of the solution

Lemma 5.17. *Let* $\{\mathbf{x}^k\}$ *be generated by Algorithm 5.7 with* $\mathbf{x}^0 \in \Omega_B$, $\Gamma > 0$, *and* $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$. *Then there is* k_0 *such that for* $k \geq k_0$

$$\mathcal{F}(\widehat{\mathbf{x}}) \subseteq \mathcal{F}(\mathbf{x}^k), \quad \mathcal{F}(\widehat{\mathbf{x}}) \subseteq \mathcal{F}(\mathbf{x}^k - \overline{\alpha}\widetilde{\boldsymbol{\varphi}}(\mathbf{x}^k)), \quad \textit{and} \quad \mathcal{B}(\widehat{\mathbf{x}}) \subseteq \mathcal{B}(\mathbf{x}^k), \quad (5.118)$$

where $\widetilde{\boldsymbol{\varphi}}(\mathbf{x}^k) = \widetilde{\boldsymbol{\varphi}}_{\overline{\alpha}}(\mathbf{x}^k)$ *is defined by* (5.85).

Proof. Since (5.118) is trivially satisfied when there is $k = k_0$ such that $\mathbf{x}^k = \widehat{\mathbf{x}}$, we shall assume in what follows that $\mathbf{x}^k \neq \widehat{\mathbf{x}}$ for any $k \geq 0$. Let us denote $x_i^k = [\mathbf{x}^k]_i$ and $\widehat{x}_i = [\widehat{\mathbf{x}}]_i$, $i = 1, \ldots, n$.

Let us first assume that $\mathcal{F}(\widehat{\mathbf{x}}) \neq \emptyset$ and $\mathcal{B}(\widehat{\mathbf{x}}) \neq \emptyset$, so that we can define

$$\varepsilon = \min\{\widehat{x}_i - \ell_i : \ i \in \mathcal{F}(\widehat{\mathbf{x}})\} > 0 \ \text{ and } \ \delta = \min\{g_i(\widehat{\mathbf{x}}) : \ i \in \mathcal{B}(\widehat{\mathbf{x}})\} > 0.$$

Since by Proposition 5.12 $\{\mathbf{x}^k\}$ converges to $\widehat{\mathbf{x}}$, there is k_0 such that for any $k \geq k_0$

$$g_i(\mathbf{x}^k) \leq \frac{\varepsilon}{4\overline{\alpha}} \text{ for } i \in \mathcal{F}(\widehat{\mathbf{x}}) \quad (5.119)$$

$$x_i^k \geq \ell_i + \frac{\varepsilon}{2} \text{ for } i \in \mathcal{F}(\widehat{\mathbf{x}}) \quad (5.120)$$

$$x_i^k \leq \ell_i + \frac{\overline{\alpha}\delta}{8} \text{ for } i \in \mathcal{B}(\widehat{\mathbf{x}}) \quad (5.121)$$

$$g_i(\mathbf{x}^k) \geq \frac{\delta}{2} \text{ for } i \in \mathcal{B}(\widehat{\mathbf{x}}). \quad (5.122)$$

In particular, for $k \geq k_0$, the first inclusion of (5.118) follows from (5.120), while the second inclusion follows from (5.119) and (5.120), as for $i \in \mathcal{F}(\widehat{\mathbf{x}})$

$$x_i^k - \overline{\alpha}\varphi_i(\mathbf{x}^k) = x_i^k - \overline{\alpha}g_i(\mathbf{x}^k) \geq \ell_i + \frac{\varepsilon}{2} - \frac{\overline{\alpha}\varepsilon}{4\overline{\alpha}} > \ell_i.$$

Let $k \geq k_0$ and observe that, by (5.121) and (5.122), for any $i \in \mathcal{B}(\widehat{\mathbf{x}})$

$$x_i^k - \overline{\alpha} g_i(\mathbf{x}^k) \le \ell_i + \frac{\overline{\alpha}\delta}{8} - \frac{\overline{\alpha}\delta}{2} < \ell_i,$$

so that if some $\mathbf{x}^{k+1}$ is generated by the expansion step (5.84), $k \ge k_0$, and $i \in \mathcal{B}(\widehat{\mathbf{x}})$, then

$$x_i^{k+1} = \max\{\ell_i, x_i^k - \overline{\alpha} g_i(\mathbf{x}^k)\} = \ell_i.$$

It follows that if $k \ge k_0$ and $\mathbf{x}^{k+1}$ is generated by the expansion step, then $\mathcal{B}(\mathbf{x}^{k+1}) \supseteq \mathcal{B}(\widehat{\mathbf{x}})$. Moreover, using (5.122) and definition of Algorithm 5.7, we can directly verify that if $\mathcal{B}(\mathbf{x}^k) \supseteq \mathcal{B}(\widehat{\mathbf{x}})$ and $k \ge k_0$, then also $\mathcal{B}(\mathbf{x}^{k+1}) \supseteq \mathcal{B}(\widehat{\mathbf{x}})$. Thus it remains to prove that there is $s \ge k_0$ such that $\mathbf{x}^s$ is generated by the expansion step.

Let us examine what can happen for $k \ge k_0$. First observe that we can never take the full CG step in the direction $\mathbf{p}^k = \boldsymbol{\varphi}(\mathbf{x}^k)$. The reason is that

$$\alpha_{cg}(\mathbf{p}^k) = \frac{\boldsymbol{\varphi}(\mathbf{x}^k)^T \mathbf{g}(\mathbf{x}^k)}{\boldsymbol{\varphi}(\mathbf{x}^k)^T \mathsf{A} \boldsymbol{\varphi}(\mathbf{x}^k)} = \frac{\|\boldsymbol{\varphi}(\mathbf{x}^k)\|^2}{\boldsymbol{\varphi}(\mathbf{x}^k)^T \mathsf{A} \boldsymbol{\varphi}(\mathbf{x}^k)} \ge \|\mathsf{A}\|^{-1} \ge \frac{\overline{\alpha}}{2},$$

so that for $i \in \mathcal{F}(\mathbf{x}^k) \cap \mathcal{B}(\widehat{\mathbf{x}})$, by (5.121) and (5.122),

$$x_i^k - \alpha_{cg} p_i^k = x_i^k - \alpha_{cg} g_i(\mathbf{x}^k) \le x_i^k - \frac{\overline{\alpha}}{2} g_i(\mathbf{x}^k) \le \ell_i + \frac{\overline{\alpha}\delta}{8} - \frac{\overline{\alpha}\delta}{4} < \ell_i. \quad (5.123)$$

It follows by definition of Algorithm 5.7 that if $\mathbf{x}^k, k \ge k_0$, is generated by the proportioning step, then the following trial conjugate gradient step is not feasible, and $\mathbf{x}^{k+1}$ is necessarily generated by the expansion step.

To complete the proof, observe that Algorithm 5.7 can generate only a finite sequence of consecutive conjugate gradient iterates. Indeed, if there is neither proportioning step nor the expansion step for $k \ge k_0$, then it follows by the finite termination property of the conjugate gradient method that there is $l \le n$ such that $\boldsymbol{\varphi}(\mathbf{x}^{k_0+l}) = \mathbf{o}$. Thus either $\mathbf{x}^{k_0+l} = \widehat{\mathbf{x}}$ and $\mathcal{B}(\mathbf{x}^k) = \mathcal{B}(\widehat{\mathbf{x}})$ for $k \ge k_0 + l$ by rule (i), or $\mathbf{x}^{k_0+l}$ is not strictly proportional, $\mathbf{x}^{k_0+l+1}$ is generated by the proportioning step, and $\mathbf{x}^{k_0+l+2}$ is generated by the expansion step. This completes the proof, as the cases $\mathcal{F}(\widehat{\mathbf{x}}) = \emptyset$ and $\mathcal{B}(\widehat{\mathbf{x}}) = \emptyset$ can be proved by a direct analysis of the above arguments. □

Proposition 5.18. *Let $\{\mathbf{x}^k\}$ be generated by Algorithm 5.7 with $\mathbf{x}^0 \in \Omega_B$, $\Gamma > 0$, and $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$. Let the solution $\widehat{\mathbf{x}}$ satisfy the condition of strict complementarity, i.e., $\widehat{x}_i = \ell_i$ implies $g_i(\widehat{\mathbf{x}}) > 0$. Then there is $k \ge 0$ such that $\mathbf{x}^k = \widehat{\mathbf{x}}$.*

Proof. If $\widehat{\mathbf{x}}$ satisfies the condition of strict complementarity, then $\mathcal{A}(\widehat{\mathbf{x}}) = \mathcal{B}(\widehat{\mathbf{x}})$, and, by Lemma 5.17, there is $k_0 \ge 0$ such that for $k \ge k_0$ we have $\mathcal{F}(\mathbf{x}^k) = \mathcal{F}(\widehat{\mathbf{x}})$ and $\mathcal{B}(\mathbf{x}^k) = \mathcal{B}(\widehat{\mathbf{x}})$. Thus, for $k \ge k_0$, all $\mathbf{x}^k$ that satisfy $\widehat{\mathbf{x}} \ne \mathbf{x}^{k-1}$ are generated by the conjugate gradient steps and, by the finite termination property of the CG, there is $k \le k_0 + n$ such that $\mathbf{x}^k = \widehat{\mathbf{x}}$. □

5.8.6 Finite Termination for Dual Degenerate Solution

Our final goal is to prove the finite termination of Algorithm 5.7 when the solution of (5.1) does not satisfy the strict complementarity condition as in Fig. 5.16, where the iterations with different active sets are near the solution.

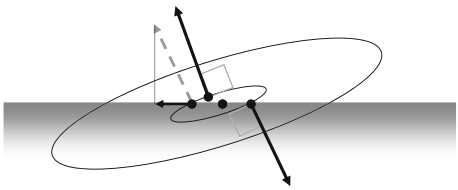

Fig. 5.16. Projected gradients near dual degenerate solution

Lemma 5.19. *Let* $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$, $\mathbf{x} \in \Omega_B$, *and* $\mathbf{y} = \mathbf{x} - \overline{\alpha}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})$. *Then*

$$\|\boldsymbol{\varphi}(\mathbf{y})\|^2 \le 9\widetilde{\boldsymbol{\varphi}}(\mathbf{x})^T\boldsymbol{\varphi}(\mathbf{x}) \quad \textit{and} \quad \|\boldsymbol{\beta}(\mathbf{y})\| \ge \|\boldsymbol{\beta}(\mathbf{x})\| - 4\|\widetilde{\boldsymbol{\varphi}}(\mathbf{x})\|, \tag{5.124}$$

where the reduced free gradient $\widetilde{\boldsymbol{\varphi}}(\mathbf{x}) = \widetilde{\boldsymbol{\varphi}}_{\overline{\alpha}}(\mathbf{x})$ *is defined by* (5.85).

Proof. First notice that $\mathcal{F}(\mathbf{y}) \subseteq \mathcal{F}(\mathbf{x})$. Since

$$\mathbf{g}(\mathbf{y}) = \mathbf{g}(\mathbf{x}) - \overline{\alpha}\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x}) \ \text{ and } \ \widetilde{\boldsymbol{\varphi}}_{\mathcal{F}(\mathbf{y})}(\mathbf{x}) = \boldsymbol{\varphi}_{\mathcal{F}(\mathbf{y})}(\mathbf{x}) = \mathbf{g}_{\mathcal{F}(\mathbf{y})}(\mathbf{x}), \tag{5.125}$$

we get

$$\begin{aligned}\|\boldsymbol{\varphi}(\mathbf{y})\| = \|\mathbf{g}_{\mathcal{F}(\mathbf{y})}(\mathbf{y})\| &= \|\mathbf{g}_{\mathcal{F}(\mathbf{y})}(\mathbf{x}) - \overline{\alpha}\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{F}(\mathbf{y})}\,\| \\ &\le \|\widetilde{\boldsymbol{\varphi}}_{\mathcal{F}(\mathbf{y})}(\mathbf{x})\| + \overline{\alpha}\,\|\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{F}(\mathbf{y})}\,\| \le 3\|\widetilde{\boldsymbol{\varphi}}(\mathbf{x})\|.\end{aligned}$$

Using the latter inequalities and the definition of $\widetilde{\boldsymbol{\varphi}}(\mathbf{x})$, we get

$$\|\boldsymbol{\varphi}(\mathbf{y})\|^2 \le 9\|\widetilde{\boldsymbol{\varphi}}(\mathbf{x})\|^2 \le 9\widetilde{\boldsymbol{\varphi}}(\mathbf{x})^T\boldsymbol{\varphi}(\mathbf{x}).$$

To prove the second inequality of (5.124), denote

$$\mathcal{C} = \{i \in \mathcal{A}(\mathbf{x}) : \ g_i(\mathbf{x}) \le 0\}$$

and notice that

$$\mathcal{A}(\mathbf{y}) \supseteq \mathcal{A}(\mathbf{x}) \supseteq \mathcal{C}.$$

Thus

$$\|\boldsymbol{\beta}(\mathbf{y})\| = \|\mathbf{g}_{\mathcal{A}(\mathbf{y})}(\mathbf{y})^-\| \geq \|\mathbf{g}_{\mathcal{C}}(\mathbf{y})^-\| = \|\big(\mathbf{g}_{\mathcal{C}}(\mathbf{x}) - \overline{\alpha}\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{C}}\big)^-\|$$
$$= \|\big(\boldsymbol{\beta}_{\mathcal{C}}(\mathbf{x}) - \overline{\alpha}\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{C}}\big)^-\|. \tag{5.126}$$

Using in sequence

$$\|\boldsymbol{\beta}_{\mathcal{C}}(\mathbf{x})\| = \|\boldsymbol{\beta}(\mathbf{x})\|, \quad \|\overline{\alpha}\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{C}}\,\| \leq 2\|\widetilde{\boldsymbol{\varphi}}(\mathbf{x})\|,$$

inequality (5.126), properties of the norm, $\boldsymbol{\beta}(\mathbf{x})^- = \boldsymbol{\beta}(\mathbf{x})$, and

$$\|\mathbf{z} - \mathbf{z}^-\| \leq \|\mathbf{z} - \mathbf{t}\|$$

for any $\mathbf{t}$ with nonpositive entries, we get

$$\|\boldsymbol{\beta}(\mathbf{x})\| - \|\widetilde{\boldsymbol{\varphi}}(\mathbf{x})\| - \|\boldsymbol{\beta}(\mathbf{y})\|$$
$$\leq \|\boldsymbol{\beta}_{\mathcal{C}}(\mathbf{x})\| - \frac{1}{2}\|\overline{\alpha}\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{C}}\,\| - \|\big(\boldsymbol{\beta}_{\mathcal{C}}(\mathbf{x}) - \overline{\alpha}\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{C}}\big)^-\|$$
$$\leq \|\boldsymbol{\beta}_{\mathcal{C}}(\mathbf{x}) - \frac{\overline{\alpha}}{2}\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{C}}\,\| - \|\big(\boldsymbol{\beta}_{\mathcal{C}}(\mathbf{x}) - \overline{\alpha}\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{C}}\big)^-\|$$
$$\leq \|\,(\boldsymbol{\beta}_{\mathcal{C}}(\mathbf{x}) - \overline{\alpha}\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{C}}) - \big(\boldsymbol{\beta}_{\mathcal{C}}(\mathbf{x}) - \overline{\alpha}\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{C}}\big)^-\| + \frac{\overline{\alpha}}{2}\|\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{C}}\,\|$$
$$\leq \|\big(\boldsymbol{\beta}_{\mathcal{C}}(\mathbf{x}) - \overline{\alpha}\,[\mathsf{A}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})]_{\mathcal{C}}\big) - \boldsymbol{\beta}_{\mathcal{C}}(\mathbf{x})\| + \|\widetilde{\boldsymbol{\varphi}}(\mathbf{x})\| \leq 3\|\widetilde{\boldsymbol{\varphi}}(\mathbf{x})\|.$$

This proves the second inequality of (5.124). □

Corollary 5.20. *Let $\Gamma \geq 4$, $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$, $\mathbf{x} \in \Omega_B$, and*

$$\Gamma^2\widetilde{\boldsymbol{\varphi}}(\mathbf{x})^T\boldsymbol{\varphi}(\mathbf{x}) < \|\boldsymbol{\beta}(\mathbf{x})\|^2, \tag{5.127}$$

where the reduced free gradient $\widetilde{\boldsymbol{\varphi}}(\mathbf{x}) = \widetilde{\boldsymbol{\varphi}}_{\overline{\alpha}}(\mathbf{x})$ is defined by (5.85).

Then the vector $\mathbf{y} = \mathbf{x} - \overline{\alpha}\widetilde{\boldsymbol{\varphi}}(\mathbf{x})$ satisfies

$$\frac{\Gamma - 4}{3}\|\boldsymbol{\varphi}(\mathbf{y})\| < \|\boldsymbol{\beta}(\mathbf{y})\|. \tag{5.128}$$

Proof. Inequality (5.128) holds trivially for $\Gamma = 4$. For $\Gamma > 4$, using in sequence (5.124), $\|\widetilde{\boldsymbol{\varphi}}(x)\|^2 \leq \widetilde{\boldsymbol{\varphi}}(\mathbf{x})^T\boldsymbol{\varphi}(\mathbf{x})$, twice (5.127), and (5.124), we get

$$\|\boldsymbol{\beta}(\mathbf{y})\| \geq \|\boldsymbol{\beta}(\mathbf{x})\| - 4\|\widetilde{\boldsymbol{\varphi}}(\mathbf{x})\| \geq \|\boldsymbol{\beta}(\mathbf{x})\| - 4\sqrt{\widetilde{\boldsymbol{\varphi}}^T(\mathbf{x})\boldsymbol{\varphi}(\mathbf{x})} > (1 - 4\Gamma^{-1})\|\boldsymbol{\beta}(\mathbf{x})\|$$
$$> (\Gamma - 4)\sqrt{\widetilde{\boldsymbol{\varphi}}^T(\mathbf{x})\boldsymbol{\varphi}(\mathbf{x})} \geq \frac{\Gamma - 4}{3}\|\boldsymbol{\varphi}(\mathbf{y})\|. \tag{5.129}$$

□

Theorem 5.21. *Let* $\{\mathbf{x}^k\}$ *denote the sequence generated by Algorithm 5.7 with*

$$\mathbf{x}^0 \in \Omega_B, \quad \Gamma \geq 3\left(\sqrt{\kappa(\mathsf{A})}+4\right), \quad \text{and} \quad \overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]. \tag{5.130}$$

Then there is $k \geq 0$ *such that* $\mathbf{x}^k = \widehat{\mathbf{x}}$.

Proof. Let $\mathbf{x}^k$ be generated by Algorithm 5.7 and let Γ satisfy (5.130). Let k_0 be that of Lemma 5.17 and let $k \geq k_0$ be such that $\mathbf{x}^k$ is not strictly proportional, i.e., $\Gamma^2 \widetilde{\boldsymbol{\varphi}}_{\overline{\alpha}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k) < \|\boldsymbol{\beta}(\mathbf{x}^k)\|^2$. Then by Corollary 5.20 the vector $\mathbf{y} = \mathbf{x}^k - \overline{\alpha}\widetilde{\boldsymbol{\varphi}}(\mathbf{x}^k)$ satisfies

$$\Gamma_1 \|\boldsymbol{\varphi}(\mathbf{y})\| < \|\boldsymbol{\beta}(\mathbf{y})\| \tag{5.131}$$

with

$$\Gamma_1 = (\Gamma - 4)/3 \geq \sqrt{\kappa(\mathsf{A})}.$$

Moreover, $\mathbf{y} \in \Omega_B$, and by Lemma 5.17 and definition of $\mathbf{y}$

$$\mathcal{A}(\widehat{\mathbf{x}}) \supseteq \mathcal{A}(\mathbf{y}) \supseteq \mathcal{A}(\mathbf{x}^k) \supseteq \mathcal{B}(\mathbf{x}^k) \supseteq \mathcal{B}(\widehat{\mathbf{x}}). \tag{5.132}$$

It follows by Lemma 5.4 that the vector $\mathbf{z} = \mathbf{y} - \|\mathsf{A}\|^{-1}\boldsymbol{\beta}(\mathbf{y})$ satisfies

$$f(\mathbf{z}) < \min\{f(\mathbf{x}) : \ \mathbf{x} \in \mathcal{W}_{\mathcal{I}}\} \tag{5.133}$$

with $\mathcal{I} = \mathcal{A}(\mathbf{y})$. Since $\mathcal{I}$ satisfies by (5.132) $\mathcal{A}(\widehat{\mathbf{x}}) \supseteq \mathcal{I} \supseteq \mathcal{B}(\widehat{\mathbf{x}})$, we have also

$$f(\widehat{\mathbf{x}}) = \min\{f(\mathbf{x}) : \ \mathbf{x} \in \Omega_B\} = \min\{f(\mathbf{x}) : \ \mathbf{x} \in \mathcal{W}_{\mathcal{I}}\}. \tag{5.134}$$

However, $\mathbf{z} \in \Omega_B$, so that (5.134) contradicts (5.133). Thus all $\mathbf{x}^k$ are strictly proportional for $k \geq k_0$, so that

$$\mathcal{A}(\mathbf{x}^{k_0}) \subseteq \mathcal{A}(\mathbf{x}^{k_0+1}) \subseteq \ldots.$$

Using the finite termination property of the conjugate gradient method, we conclude that there is $k \geq k_0$ such that $\widehat{\mathbf{x}} = \mathbf{x}^k$. □

Let us recall that the finite termination property of the MPRGP algorithm with a dual degenerate solution and

$$\overline{\alpha} \in (0, \|\mathsf{A}\|^{-1}]$$

has been proved for

$$\Gamma \geq 2\left(\sqrt{\kappa(\mathsf{A})}+1\right).$$

For the details see Dostál [74].

5.9 Implementation of MPRGP with Optional Modifications

In this section, we describe Algorithm 5.7 in the form that is convenient for implementation. We include also some modifications that may be used to improve its performance. Implementation of Algorithm 5.6 is similar.

5.9.1 Expansion Step with Feasible Half-Step

To improve the efficiency of the expansion step, we can use the trial conjugate gradient direction $\mathbf{p}^k$ which is generated before the expansion step is invoked. We propose to generate first

$$\mathbf{x}^{k+\frac{1}{2}} = \mathbf{x}^k - \alpha_f \mathbf{p}^k \quad \text{and} \quad \mathbf{g}^{k+\frac{1}{2}} = \mathbf{g}^k - \alpha_f \mathsf{A}\mathbf{p}^k,$$

where the feasible steplength α_f for $\mathbf{p}^k$ is defined by

$$\alpha_f = \max\{\alpha : \ \mathbf{x}^k - \alpha \mathbf{p}^k \in \Omega_B\} = \min_{i=1,\dots,n} \{(x_i^k - \ell_i)/p_i^k, \ p_i^k > 0\},$$

and then define

$$\mathbf{x}^{k+1} = P_{\Omega_B}\left(\mathbf{x}^{k+\frac{1}{2}} - \overline{\alpha}\boldsymbol{\varphi}(\mathbf{x}^{k+\frac{1}{2}})\right).$$

The half-step is illustrated in Fig. 5.17. Such modification does not require any additional matrix–vector multiplication and the estimate (5.93) remains valid as $f(\mathbf{x}^{k+\frac{1}{2}}) - f(\mathbf{x}^k) \le 0$ and

$$\begin{aligned} f(\mathbf{x}^{k+1}) - f(\widehat{\mathbf{x}}) &\le \eta_\Gamma \left((f(\mathbf{x}^{k+\frac{1}{2}}) - f(\mathbf{x}^k)) + f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\right) \\ &\le \eta_\Gamma \left(f(\mathbf{x}^k) - f(\widehat{\mathbf{x}})\right). \end{aligned}$$

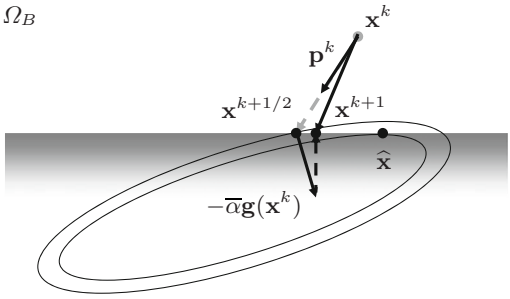

Fig. 5.17. Feasible half-step

5.9.2 MPRGP Algorithm

Now we are ready to give the details of the implementation of the MPRGP algorithm which was briefly described in the form suitable for analysis as Algorithm 5.7. To preserve readability, we do not distinguish generations of variables by indices unless it is convenient for further reference.

Algorithm 5.8. Modified proportioning with reduced gradient projections (MPRGP).

Given a symmetric positive definite matrix A *of the order* n*, n-vectors* $\mathbf{b}$*,* $\boldsymbol{\ell}$*,* $\Omega_B = \{\mathbf{x} :\ \mathbf{x} \geq \boldsymbol{\ell}\}$*,* $\mathbf{x}^0 \in \Omega_B$*.*
Step 0. {*Initialization.*}
Choose $\Gamma > 0$*,* $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$*, set* $k = 0,\ \mathbf{g} = \mathsf{A}\mathbf{x}^0 - \mathbf{b},\ \mathbf{p} = \boldsymbol{\varphi}(\mathbf{x}^0)$
while $\|\mathbf{g}^P(\mathbf{x}^k)\|$ *is not small*
if $\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2 \leq \Gamma^2 \widetilde{\boldsymbol{\varphi}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k)$
Step 1. {*Proportional* $\mathbf{x}^k$*. Trial conjugate gradient step.*}
$\alpha_{cg} = \mathbf{g}^T\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathbf{p},\ \mathbf{y} = \mathbf{x}^k - \alpha_{cg}\mathbf{p}$
$\alpha_f = \max\{\alpha :\ \mathbf{x}^k - \alpha\mathbf{p} \in \Omega_B\} = \min\{(x_i^k - \ell_i)/p_i :\ p_i > 0\}$
if $\alpha_{cg} \leq \alpha_f$
Step 2. {*Conjugate gradient step.*}
$\mathbf{x}^{k+1} = \mathbf{y},\ \mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathbf{p},$
$\beta = \boldsymbol{\varphi}(\mathbf{y})^T\mathsf{A}\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathbf{p},\ \mathbf{p} = \boldsymbol{\varphi}(\mathbf{y}) - \beta\mathbf{p}$
else
Step 3. {*Expansion step.*}
$\mathbf{x}^{k+\frac{1}{2}} = \mathbf{x}^k - \alpha_f\mathbf{p},\ \mathbf{g} = \mathbf{g} - \alpha_f\mathsf{A}\mathbf{p}$
$\mathbf{x}^{k+1} = P_{\Omega_B}(\mathbf{x}^{k+\frac{1}{2}} - \overline{\alpha}\boldsymbol{\varphi}(\mathbf{x}^{k+\frac{1}{2}}))$
$\mathbf{g} = \mathsf{A}\mathbf{x}^{k+1} - \mathbf{b},\ \mathbf{p} = \boldsymbol{\varphi}(\mathbf{x}^{k+1})$
end if
else
Step 4. {*Proportioning step.*}
$\mathbf{d} = \boldsymbol{\beta}(\mathbf{x}^k),\ \alpha_{cg} = \mathbf{g}^T\mathbf{d}/\mathbf{d}^T\mathsf{A}\mathbf{d}$
$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_{cg}\mathbf{d},\ \mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathbf{d},\ \mathbf{p} = \boldsymbol{\varphi}(\mathbf{x}^{k+1})$
end if
$k = k + 1$
end while
Step 5. {*Return (possibly inexact) solution.*}
$\widetilde{\mathbf{x}} = \mathbf{x}^k$

In our description, we denote by $\widetilde{\boldsymbol{\varphi}}(\mathbf{x}) = \widetilde{\boldsymbol{\varphi}}_{\overline{\alpha}}(\mathbf{x})$ the reduced free gradient defined by (5.85). Let us recall that by Proposition 5.12 the algorithm converges for any $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$ and by Theorem 5.14 its R-linear rate of convergence is guaranteed for $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1})$.

5.9.3 Unfeasible MPRGP

The "global" bound on the rate of convergence of the CG method guaranteed by Theorem 3.2 indicates that MPRGP converges fast when it generates long chains of CG iterations. Thus it may be advantageous to continue the CG iterations when the trial CG step is unfeasible. The modification of MPRGP proposed here is based on the observation that the convergence of MPRGP is preserved when we insert between the last feasible iteration and the expansion step a finite number of unfeasible iterates as long as $\{f\left(P_{\Omega_B}(\mathbf{x}^k)\right)\}$ decreases. Thus if $f(P_{\Omega_B}(\mathbf{y})) \leq f(P_{\Omega_B}(\mathbf{x}^k))$, we can define $\mathbf{x}^{k+1} = \mathbf{y}$ and continue the CG iterations; otherwise we generate $\mathbf{x}^{k+1}$ by the modified expansion step. The resulting *monotonic MPRGP* algorithm reads as follows.

Algorithm 5.9. Monotonic MPRGP.

Given a symmetric positive definite matrix A *of the order* n*, n-vectors* $\mathbf{b}$*,* $\boldsymbol{\ell}$*,* $\Omega_B = \{\mathbf{x} : \mathbf{x} \geq \boldsymbol{\ell}\}$*,* $\mathbf{x}^0 \in \Omega_B$*.*
Step 0. {*Initialization.*}
Choose $\Gamma > 0$*,* $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$*, set* $k = 0$*,* $\mathbf{g} = \mathsf{A}\mathbf{x}^0 - \mathbf{b}$*,* $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{x}^0)$
while $\|\mathbf{g}^P\left(\mathbf{x}^k\right)\|$ *is not small*
if $\|\boldsymbol{\beta}\left(\mathbf{x}^k\right)\|^2 \leq \Gamma^2 \widetilde{\boldsymbol{\varphi}}\left(\mathbf{x}^k\right)^T \boldsymbol{\varphi}\left(\mathbf{x}^k\right)$
$\alpha_{cg} = \mathbf{g}^T\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathbf{p}$*,* $\mathbf{y} = \mathbf{x}^k - \alpha_{cg}\mathbf{p}$
while $f\left(P_{\Omega_B}(\mathbf{y})\right) \leq f(P_{\Omega_B}\left(\mathbf{x}^k\right))$ **and**
$\|\boldsymbol{\beta}\left(\mathbf{x}^k\right)\|^2 \leq \Gamma^2 \widetilde{\boldsymbol{\varphi}}\left(\mathbf{x}^k\right)^T \boldsymbol{\varphi}\left(\mathbf{x}^k\right)$ **and** $\|\mathbf{g}^P\left(\mathbf{x}^k\right)\|$ *not small*
Step 1. {*Conjugate gradient step.*}
$\mathbf{x}^{k+1} = \mathbf{y}$*,* $\mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathbf{p}$*,*
$\beta = \boldsymbol{\varphi}(\mathbf{y})^T\mathsf{A}\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathbf{p}$*,* $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{y}) - \beta\mathbf{p}$*,* $k = k + 1$
Step 2. {*Trial CG step for the next iteration of the CG loop.*}
$\alpha_{cg} = \mathbf{g}^T\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathbf{p}$*,* $\mathbf{y} = \mathbf{x}^k - \alpha_{cg}\mathbf{p}$
end *while for CG loop*
end *if*
if $\mathbf{y} \notin \Omega_B$ **and** $\|\mathbf{g}^P\left(\mathbf{x}^k\right)\|$ *not small*
Step 3. {*Expansion step.*}
$\mathbf{y} = P_{\Omega_B}\left(\mathbf{x}^k\right)$*,* $\mathbf{x}^{k+1} = P_{\Omega_B}(\mathbf{y} - \overline{\alpha}\boldsymbol{\varphi}(\mathbf{y}))$
$\mathbf{g} = \mathsf{A}\mathbf{x}^{k+1} - \mathbf{b}$*,* $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{x}^{k+1})$*,* $k = k + 1$
else
if $\|\boldsymbol{\beta}\left(\mathbf{x}^k\right)\|^2 > \Gamma^2 \widetilde{\boldsymbol{\varphi}}\left(\mathbf{x}^k\right)^T \boldsymbol{\varphi}\left(\mathbf{x}^k\right)$ **and** $\|\mathbf{g}^P\left(\mathbf{x}^k\right)\|$ *not small*
Step 4. {*Proportioning step.*}
$\mathbf{d} = \boldsymbol{\beta}\left(\mathbf{x}^k\right)$*,* $\mathbf{g} = \mathsf{A}\mathbf{x}^k - \mathbf{b}$*,* $\alpha_{cg} = \mathbf{g}^T\mathbf{d}/\mathbf{d}^T\mathsf{A}\mathbf{d}$
$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_{cg}\mathbf{d}$*,* $\mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathbf{d}$*,* $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{x}^{k+1})$
$k = k + 1$
end if
end if
end while
Step 5. {*Return (possibly inexact) solution.*}
$\widetilde{\mathbf{x}} = P_{\Omega_B}\left(\mathbf{x}^k\right)$

To see that the algorithm is well defined, namely, that $\mathbf{p} \neq \mathbf{o}$ in Step 1, it is enough to notice that this step is carried out when

$$\|\mathbf{g}^P\left(\mathbf{x}^k\right)\| > 0 \quad \text{and} \quad \|\boldsymbol{\beta}\left(\mathbf{x}^k\right)\|^2 \leq \Gamma^2 \widetilde{\boldsymbol{\varphi}}\left(\mathbf{x}^k\right)^T \boldsymbol{\varphi}\left(\mathbf{x}^k\right),$$

where $\widetilde{\boldsymbol{\varphi}}(\mathbf{x}) = \widetilde{\boldsymbol{\varphi}}_{\overline{\alpha}}(\mathbf{x})$ denotes the reduced free gradient defined by (5.85). Thus

$$\|\boldsymbol{\beta}\left(\mathbf{x}^k\right)\| + \|\boldsymbol{\varphi}\left(\mathbf{x}^k\right)\| > 0 \quad \text{and} \quad \|\boldsymbol{\beta}\left(\mathbf{x}^k\right)\| \leq \Gamma \|\boldsymbol{\varphi}\left(\mathbf{x}^k\right)\|.$$

It follows easily that $\boldsymbol{\varphi}\left(\mathbf{x}^k\right) \neq \mathbf{o}$. Since

$$\|\mathbf{p}\| \geq \|\boldsymbol{\varphi}\left(\mathbf{x}^k\right)\|,$$

we have $\mathbf{p} \neq \mathbf{o}$. If $\mathbf{x}^k$ is feasible, we can optionally implement the expansion step with the feasible half-step of Sect. 5.9.1. Notice that $\mathbf{x}^k$ is always feasible at the beginning of the outer loop.

Each unfeasible CG step of our implementation of the monotonic MPRGP algorithm requires two matrix–vector multiplications; the additional multiplication is necessary for evaluation of the test associated with the inner CG loop. To carry out the unfeasible CG step in one matrix–vector multiplication, we can use that for any $\mathbf{x}, \mathbf{d} \in \mathbb{R}^n$

$$f(\mathbf{x}+\mathbf{d}) = f(\mathbf{x}) + \mathbf{g}^T\mathbf{d} + \frac{1}{2}\mathbf{d}^T\mathsf{A}\mathbf{d} \leq f(\mathbf{x}) + \mathbf{g}^T\mathbf{d} + \frac{1}{2}\|\mathsf{A}\|\|\mathbf{d}\|^2$$

and

$$f(\mathbf{x}+\mathbf{d}) = f(\mathbf{x}) + \mathbf{g}^T\mathbf{d} + \frac{1}{2}\mathbf{d}^T\mathsf{A}\mathbf{d} \geq f(\mathbf{x}) + \mathbf{g}^T\mathbf{d}.$$

For example, if $\mathbf{x}^i,\ i = k, k+1, \ldots,$ are generated in the inner CG loop of the monotonic MPRGP algorithm, $\mathbf{x}^k$ is feasible, and

$$\mathbf{d}^i = P_{\Omega_B}(\mathbf{y}^i) - \mathbf{x}^i, \quad \mathbf{g}^i = \mathbf{g}(\mathbf{y}^i), \quad i = k, k+1, \ldots,$$

where $\mathbf{y}^i$ is the trial CG iteration entering into the ith step, then we can use (3.9) to evaluate $f(\mathbf{y}^i)$ without additional matrix–vector multiplication,

$$f(\mathbf{y}^i) + (\mathbf{g}^i)^T\mathbf{d}^i + \|\mathsf{A}\|\|\mathbf{d}^i\|^2 \leq f(\mathbf{x}^k) \tag{5.135}$$

implies

$$f\left(P_{\Omega_B}(\mathbf{y}^i)\right) \leq f\left(\mathbf{x}^k\right),$$

and the unfeasible iterates $\mathbf{x}^{i+1} = \mathbf{y}^i$ which satisfy (5.135) can be accepted. Thus we can use (5.135) to modify the test at the beginning of the CG loop of Algorithm 5.9 so that the resulting *semimonotonic MPRGP* algorithm generates a *converging sequence of iterates that are evaluated at one matrix–vector multiplication.*

Using the lower bound on $f(\mathbf{x}^i)$, it is possible to develop a test applicable to unfeasible $\mathbf{x}^k$. The modifications presented in this section are closely related to the semismooth Newton methods.

5.9.4 Choice of Parameters

Our experience indicates that MPRGP is not sensitive to Γ as long as $\Gamma \approx 1$. Since $\Gamma = 1$ minimizes the upper bound on the rate of convergence and guarantees that the CG steps reduce directly the larger of the two components of the projected gradient, we can expect good efficiency with this value.

The choice of $\overline{\alpha}$ requires an estimate of $\|\mathsf{A}\|$. If the entries of A are available, we can use $\|\mathsf{A}\| \le \|\mathsf{A}\|_\infty$ to define $\overline{\alpha} = 2\|\mathsf{A}\|_\infty^{-1}$ which guarantees convergence. If this is not the case, or if $\|\mathsf{A}\|_\infty$ gives a poor upper bound on $\|\mathsf{A}\|$, then we can carry out a few, e.g., five, iterations of the following power method.

Algorithm 5.10. Power method for the estimate of $\|\mathsf{A}\|$.

Given a symmetric positive definite matrix $\mathsf{A} \in \mathbb{R}^{n\times n}$*, returns* $A \approx \|\mathbf{A}\|$.
Choose $\mathbf{x} \in \mathbb{R}^n$ *such that* $\mathbf{x} \neq \mathbf{o}$, $n_{\rm it} \ge 1$
for $i = 1, 2, \ldots, n_{\rm it}$
$\mathbf{y} = \mathsf{A}\mathbf{x}$, $\mathbf{x} = \|\mathbf{y}\|^{-1}\mathbf{y}$
end for
$A = \|\mathsf{A}\mathbf{x}\|$

Alternatively, we can use the Lanczos method (see, e.g., Golub and van Loan [103]). We can conveniently enhance the Lanczos method into the conjugate gradient loop of the MPRGP algorithm by defining

$$\mathbf{q}_i = \|\boldsymbol{\varphi}(\mathbf{x}^{s+i})\|^{-1}\boldsymbol{\varphi}(\mathbf{x}^{s+i}), \quad i = 0, \ldots, p,$$

where $\boldsymbol{\varphi}(\mathbf{x}^s)$ and $\boldsymbol{\varphi}(\mathbf{x}^{s+i})$ are the free gradients at respectively the initial and the ith iterate in one CG loop. Then we can estimate $\|\mathsf{A}\|$ by evaluation of the ℓ_∞-norm of the tridiagonal matrix

$$\mathsf{T} = \mathsf{Q}^T\mathsf{A}\mathsf{Q}, \quad \mathsf{Q} = [\mathbf{q}_0, \ldots, \mathbf{q}_p].$$

Though these methods typically give only a lower bound A on the norm of $\|\mathsf{A}\|$, the choice like $\overline{\alpha} = 1.8A^{-1}$ is often sufficient in practice. The decrease of f can be achieved more reliably by initializing $\overline{\alpha} = 2(\mathbf{b}^T\mathsf{A}\mathbf{b})^{-1}\|\mathbf{b}\|^2$ and by inserting the following piece of code into the expansion step:

Algorithm 5.11. Modification of the steplength of the expansion step.

A piece of code to be inserted at the end of the expansion step of Algorithm 5.8.
if $f\left(P_{\Omega_B}(\mathbf{x}^{k+1})\right) > f(\mathbf{x}^k)$
$\overline{\alpha} = \overline{\alpha}/2$ *and repeat the expansion step*
end if

The modified algorithm can outperform that with $\overline{\alpha} = \|\mathsf{A}\|^{-1}$; the longer steps in an early stage of computations can be effective for identification of the solution. We observed a good performance with $\overline{\alpha}$ close to, but not greater than $2\|\mathsf{A}\|^{-1}$, near $\overline{\alpha}_E^{opt}$ which minimizes the coefficient η_E of the Euclidean contraction (5.36). Notice that Theorem 5.14 guarantees that the inserted loop of Algorithm 5.11 reduces the steplength in a small number of steps.

5.9.5 Dynamic Release Coefficient

The estimates given by Lemma 5.4 and Theorem 5.21 indicate that the value of $\Gamma = 1$, which gives the best upper bound on the rate of convergence of the MPRGP algorithm, may be too small to exclude repeated exploitation of any face. On the other hand, while discussing the original Polyak algorithm in Sect. 5.4, we have already expressed doubts that it is efficient to carry out the minimization in face to a high precision, especially in the early stage of computations, when we are far from the solution.

To accommodate these contradicting requirements, let us return to the description of the MPRGP algorithm in Sect. 5.8 and replace in its kth step the release coefficient Γ by Γ_k, so that it can change from iteration to iteration. For example, we shall now say that the iterate $\mathbf{x}^k$ is *strictly proportional* if

$$||\boldsymbol{\beta}(\mathbf{x}^k)||^2 \le \Gamma_k^2 \widetilde{\boldsymbol{\varphi}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k). \tag{5.136}$$

Repeating the arguments of the proof of Theorem 5.14, we can prove its following modification:

Theorem 5.22. *Let $\Gamma_{\max} \ge \Gamma_{\min}$ denote given positive numbers, let $\{\Gamma_i\}$ denote a given sequence such that $\Gamma_{\max} \ge \Gamma_k \ge \Gamma_{\min}$, let $\lambda_{\min}$ denote the smallest eigenvalue of A, and let $\{\mathbf{x}^k\}$ denote the sequence generated by Algorithm 5.7 with $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$ and Γ replaced in the kth step by Γ_k.*

Then the error in the A-norm is bounded by

$$\|\mathbf{x}^k - \widehat{\mathbf{x}}\|_{\mathsf{A}}^2 \le 2\eta_{\Gamma_1} \dots \eta_{\Gamma_k} \left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\right) \le 2\eta_{\Gamma}^k \left(f(\mathbf{x}^0) - f(\widehat{\mathbf{x}})\right), \tag{5.137}$$

where $\widehat{\mathbf{x}}$ denotes the unique solution of (5.1),

$$\eta_{\Gamma} = 1 - \frac{\overline{\alpha}\lambda_{\min}}{\vartheta + \vartheta\widehat{\Gamma}^2}, \quad \eta_{\Gamma_k} = 1 - \frac{\overline{\alpha}\lambda_{\min}}{\vartheta + \vartheta\widehat{\Gamma}_k^2}, \tag{5.138}$$

$$\vartheta = 2\max\{\overline{\alpha}/2, 1\}, \quad \widehat{\Gamma} = \max\{\Gamma_{\max}, \Gamma_{\min}^{-1}\}, \quad \widehat{\Gamma}_k = \max\{\Gamma_k, \Gamma_k^{-1}\}.$$

This definition opens room for implementation of heuristics that can be useful in some specific cases. Typically, the series of release coefficients $\{\Gamma_k\}$ is defined by a suitable function of $\|\mathbf{g}^P(\mathbf{x}^k)\|$. For example, specification

$$\overline{\alpha} = \|\mathsf{A}\|^{-1} \quad \text{and} \quad \Gamma_k = \begin{cases} 1 & \text{for} \quad \|\mathbf{g}^P(\mathbf{x}^k)\| \ge 10\varepsilon, \\ 2(\sqrt{\kappa(\mathsf{A})} + 1) & \text{for} \quad \|\mathbf{g}^P(\mathbf{x}^k)\| < 10\varepsilon \end{cases}$$

guarantees both favorable bound on the rate of convergence in the early stage of computation and the finite termination property.

5.10 Preconditioning

A natural way to improve the performance of the conjugate gradient-based methods is to apply the preconditioning described in Sect. 3.6. However, the application of preconditioning requires some care, as the *preconditioning transforms the variables, turning the bound constraints into more general inequality constraints*. In this section we present two strategies which preserve the bound constraints.

5.10.1 Preconditioning in Face

Probably the most straightforward preconditioning strategy which preserves the bound constraints is the preconditioning applied to the diagonal block $\mathsf{A}_{\mathcal{FF}}$ of the Hessian matrix A in the conjugate gradient loop which minimizes the cost function f in the face defined by a free set $\mathcal{F}$. Such preconditioning requires that we are able to define for each diagonal block $\mathsf{A}_{\mathcal{FF}}$ a regular matrix $\mathsf{M}(\mathcal{F})$ which satisfies the following two conditions. First, we require that $\mathsf{M}(\mathcal{F})$ approximates $\mathsf{A}_{\mathcal{FF}}$ so that the convergence of the conjugate gradients method is significantly accelerated. The second condition requires that the solution of the system

$$\mathsf{M}(\mathcal{F})\mathbf{x} = \mathbf{y}$$

can be obtained easily. The preconditioners $\mathsf{M}(\mathcal{F})$ can be generated, e.g., by any of the methods described in Sect. 3.6.

Though the performance of the algorithm can be considerably improved by the preconditioning, *preconditioning in face does not result in the improved bound on rate of convergence*. The reason is that such preconditioning affects only the feasible conjugate gradient step, leaving the expansion and the proportioning steps without any preconditioning.

In probably the first application of preconditioning to the solution of bound constrained problems [157], O'Leary considered two simple methods which can be used to obtain the preconditioner for $\mathsf{A}_{\mathcal{FF}}$ from the preconditioner M which approximates A, namely,

$$\mathsf{M}(\mathcal{F}) = \mathsf{M}_{\mathcal{FF}} \quad \text{and} \quad \mathsf{M}(\mathcal{F}) = \mathsf{L}_{\mathcal{FF}}\mathsf{L}_{\mathcal{FF}}^T,$$

where L denotes the factor of the Cholesky factorization $\mathsf{M} = \mathsf{L}\mathsf{L}^T$. It can be proved that whichever method of the preconditioning is used, the convergence bound for the conjugate gradient algorithm applied to the subproblems is at least as good as that of the conjugate gradient method applied to the original matrix [157].

To describe the MPRGP algorithm with the preconditioning in face, let us assume that we are given the preconditioner $\mathsf{M}(\mathcal{F})$ for each set of indices $\mathcal{F}$, and let us denote $\mathcal{F}_k = \mathcal{F}(\mathbf{x}^k)$ and $\mathcal{A}_k = \mathcal{A}(\mathbf{x}^k)$ for each vector $\mathbf{x}^k \in \Omega_B$. To simplify the description of the algorithm, let M_k denote the preconditioner corresponding to the face defined by $\mathcal{F}_k$ padded with zeros so that

$$[\mathsf{M}_k]_{\mathcal{F}\mathcal{F}} = \mathsf{M}(\mathcal{F}_k), \quad [\mathsf{M}_k]_{\mathcal{A}\mathcal{A}} = \mathsf{O}, \quad [\mathsf{M}_k]_{\mathcal{A}\mathcal{F}} = [\mathsf{M}_k]^T_{\mathcal{F}\mathcal{A}} = \mathsf{O},$$

and recall that $\mathsf{M}_k^\dagger$ denotes the Moore–Penrose generalized inverse of M_k which is defined by

$$[\mathsf{M}_k^\dagger]_{\mathcal{F}\mathcal{F}} = \mathsf{M}(\mathcal{F}_k)^{-1}, \quad [\mathsf{M}_k^\dagger]_{\mathcal{A}\mathcal{A}} = \mathsf{O}, \quad [\mathsf{M}_k^\dagger]_{\mathcal{A}\mathcal{F}} = [\mathsf{M}_k^\dagger]^T_{\mathcal{F}\mathcal{A}} = \mathsf{O}.$$

In particular, it follows that

$$\mathsf{M}_k^\dagger \mathbf{g}(\mathbf{x}^k) = \mathsf{M}_k^\dagger \boldsymbol{\varphi}(\mathbf{x}^k).$$

The MPRGP algorithm with preconditioning in face reads as follows.

Algorithm 5.12. MPRGP with preconditioning in face.

Given a symmetric positive definite matrix A *of the order* n*, n-vectors* $\mathbf{b}$*,* $\boldsymbol{\ell}$*,* $\Omega_B = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \boldsymbol{\ell}\}$*; choose* $\mathbf{x}^0 \in \Omega_B$*,* $\Gamma > 0$*,* $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$*, and the rule which assigns to each* $\mathbf{x}^k \in \Omega_B$ *the preconditioner* M_k *which is SPD in the face defined by* $\mathcal{F}(\mathbf{x}^k)$*.*

Step 0. {*Initialization.*}
  Set $k = 0$, $\mathbf{g} = \mathsf{A}\mathbf{x}^0 - \mathbf{b}$, $\mathbf{z} = \mathsf{M}_0^\dagger \mathbf{g}$, $\mathbf{p} = \mathbf{z}$
  while $\|\mathbf{g}^P(\mathbf{x}^k)\|$ *is not small*
    if $\|\boldsymbol{\beta}(\mathbf{x}^k)\|^2 \leq \Gamma^2 \widetilde{\boldsymbol{\varphi}}(\mathbf{x}^k)^T \boldsymbol{\varphi}(\mathbf{x}^k)$
Step 1. {*Proportional* $\mathbf{x}^k$*. Trial conjugate gradient step.*}
      $\alpha_{cg} = \mathbf{z}^T\mathbf{g}/\mathbf{p}^T\mathsf{A}\mathbf{p}$, $\mathbf{y} = \mathbf{x}^k - \alpha_{cg}\mathbf{p}$
      $\alpha_f = \max\{\alpha : \mathbf{x}^k - \alpha\mathbf{p} \in \Omega_B\} = \min\{(x_i^k - \ell_i)/p_i : p_i > 0\}$
      if $\alpha_{cg} \leq \alpha_f$
Step 2. {*Conjugate gradient step.*}
        $\mathbf{x}^{k+1} = \mathbf{y}$, $\mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathbf{p}$, $\mathbf{z} = \mathsf{M}_k^\dagger \mathbf{g}$
        $\beta = \mathbf{z}^T\mathsf{A}\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathbf{p}$, $\mathbf{p} = \mathbf{z} - \beta\mathbf{p}$
      else
Step 3. {*Expansion step.*}
        $\mathbf{x}^{k+\frac{1}{2}} = \mathbf{x}^k - \alpha_f\mathbf{p}$, $\mathbf{g} = \mathbf{g} - \alpha_f\mathsf{A}\mathbf{p}$
        $\mathbf{x}^{k+1} = P_{\Omega_B}\left(\mathbf{x}^{k+\frac{1}{2}} - \overline{\alpha}\boldsymbol{\varphi}(\mathbf{x}^{k+\frac{1}{2}})\right)$
        $\mathbf{g} = \mathsf{A}\mathbf{x}^{k+1} - \mathbf{b}$, $\mathbf{z} = \mathsf{M}_{k+1}^\dagger \mathbf{g}$, $\mathbf{p} = \mathbf{z}$
      end if
    else
Step 4. {*Proportioning step.*}
      $\mathbf{d} = \boldsymbol{\beta}(\mathbf{x}^k)$, $\alpha_{cg} = \mathbf{g}^T\mathbf{d}/\mathbf{d}^T\mathsf{A}\mathbf{d}$
      $\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_{cg}\mathbf{d}$, $\mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathbf{d}$, $\mathbf{z} = \mathsf{M}_{k+1}^\dagger\mathbf{g}$, $\mathbf{p} = \mathbf{z}$
    end if
    $k = k + 1$
  end while
Step 5. {*Return (possibly inexact) solution.*}
  $\widetilde{\mathbf{x}} = \mathbf{x}^k$

5.10.2 Preconditioning by Conjugate Projector

Let $1 \le m < n$ and let the vector of bounds satisfy

$$\ell_{m+1} = -\infty, \ldots, \ell_n = -\infty,$$

so that problem (5.1) is only partially constrained and the feasible set can be described by

$$\Omega_B = \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{x}_{\mathcal{I}} \ge \boldsymbol{\ell}_{\mathcal{I}}\}, \quad \mathcal{I} = \{1, \ldots, m\}. \tag{5.139}$$

Here we show that such partially constrained problems can be preconditioned by the conjugate projector of Sect. 3.7 and that *it is possible to give an improved bound on the rate of convergence of the preconditioned problem*.

Let us assume that $\mathcal{U}$ is the subspace spanned by the full column rank matrix $\mathsf{U} \in \mathbb{R}^{n \times p}$ of the form

$$\mathsf{U} = \begin{bmatrix} \mathsf{O} \\ \mathsf{V} \end{bmatrix}, \quad \mathsf{V} \in \mathbb{R}^{(n-m)\times p}.$$

As in Sect. 3.7.2, we decompose our partially constrained problem by means of the conjugate projectors

$$\mathsf{P} = \mathsf{U}(\mathsf{U}^T\mathsf{A}\mathsf{U})^{-1}\mathsf{U}^T\mathsf{A} \tag{5.140}$$

and $\mathsf{Q} = \mathsf{I} - \mathsf{P}$ onto $\mathcal{U}$ and $\mathcal{V} = \mathrm{Im}\mathsf{Q}$, respectively. Due to our special choice of U, we get that for any $\mathbf{x} \in \mathbb{R}^n$

$$[\mathsf{Q}\mathbf{x}]_{\mathcal{I}} = \mathbf{x}_{\mathcal{I}},$$

and that for any $\mathbf{y} \in \mathcal{U}$ and $\mathbf{z} \in \mathcal{V}$, $\mathbf{y} + \mathbf{z} \in \Omega_B$ if and only if $\mathbf{z} \in \Omega_B$. Using (3.32), (3.33), and the observations of Sect. 3.7.3, we thus get

$$\begin{aligned}
\min_{\mathbf{x}\in\Omega_B} f(\mathbf{x}) &= \min_{\substack{\mathbf{y}\in\mathcal{U},\ \mathbf{z}\in\mathcal{V} \\ \mathbf{y}+\mathbf{z}\in\Omega_B}} f(\mathbf{y}+\mathbf{z}) = \min_{\mathbf{y}\in\mathcal{U}} f(\mathbf{y}) + \min_{\mathbf{z}\in\mathcal{V}\cap\Omega_B} f(\mathbf{z}) \\
&= f(\mathbf{x}^0) + \min_{\mathbf{z}\in\mathcal{V}\cap\Omega_B} f(\mathbf{z}) = f(\mathbf{x}^0) + \min_{\substack{\mathbf{z}\in\mathsf{A}\mathcal{V} \\ \mathbf{z}_{\mathcal{I}}\ge\boldsymbol{\ell}_{\mathcal{I}}}} \frac{1}{2}\mathbf{z}^T\mathsf{Q}^T\mathsf{A}\mathsf{Q}\mathbf{z} - \mathbf{b}^T\mathsf{Q}\mathbf{z} \\
&= f(\mathbf{x}^0) + \min_{\substack{\mathbf{z}\in\mathsf{A}\mathcal{V} \\ \mathbf{z}_{\mathcal{I}}\ge\boldsymbol{\ell}_{\mathcal{I}}}} \frac{1}{2}\mathbf{z}^T\mathsf{Q}^T\mathsf{A}\mathsf{Q}\mathbf{z} + \left(\mathbf{g}^0\right)^T\mathbf{z},
\end{aligned}$$

where $\mathbf{x}^0 = \mathsf{P}\mathsf{A}^{-1}\mathbf{b}$ and $\mathbf{g}^0 = -\mathsf{Q}^T\mathbf{b}$. We have thus reduced our bound constrained problem (5.1) with the feasible set (5.139) to the problem

$$\min_{\substack{\mathbf{z}\in\mathsf{A}\mathcal{V} \\ \mathbf{z}_{\mathcal{I}}\ge\boldsymbol{\ell}_{\mathcal{I}}}} \frac{1}{2}\mathbf{z}^T\mathsf{Q}^T\mathsf{A}\mathsf{Q}\mathbf{z} + \left(\mathbf{g}^0\right)^T\mathbf{z}. \tag{5.141}$$

The following lemma shows that the above problem can be solved by the MPRGP algorithm.

Lemma 5.23. *Let* $\mathbf{z}^1, \mathbf{z}^2, \ldots$ *be generated by the MPRGP algorithm for the problem*

$$\min_{\mathbf{z}_{\mathcal{I}} \geq \boldsymbol{\ell}_{\mathcal{I}}} \frac{1}{2}\mathbf{z}^T \mathsf{Q}^T \mathsf{A}\mathsf{Q}\mathbf{z} + \left(\mathbf{g}^0\right)^T \mathbf{z} \tag{5.142}$$

starting from $\mathbf{z}^0 = P_{\Omega_B}\left(\mathbf{g}^0\right)$. *Then* $\mathbf{z}^k \in \mathsf{A}\mathcal{V},\ k = 0, 1, 2, \ldots$.

Proof. First observe that since $\mathsf{A}\mathcal{V}$ is orthogonal to $\mathcal{U}$ and $\dim \mathsf{A}\mathcal{V} = \dim \mathcal{V}$, it follows that $\mathsf{A}\mathcal{V}$ is the orthogonal complement of $\mathcal{U}$. Thus $\mathsf{A}\mathcal{V}$ is not only an invariant subspace of Q, but it is also an invariant subspace of P_{Ω_B}. Moreover, it also follows that $\mathsf{A}\mathcal{V}$ contains the set $\mathcal{V}_0 \subseteq \mathbb{R}^n$ of all the vectors of $\mathbb{R}^m$ padded with zeros,

$$\mathcal{V}_0 = \{\mathbf{x} \in \mathbb{R}^n :\ \mathbf{x}_{\mathcal{J}} = \mathbf{o},\ \mathcal{J} = \{m+1, \ldots, n\}\}.$$

More formally,

$$P_{\Omega_B}(\mathsf{A}\mathcal{V}) \subseteq \mathsf{A}\mathcal{V} \quad \text{and} \quad \mathcal{V}_0 \subseteq \mathsf{A}\mathcal{V}. \tag{5.143}$$

Let us now recall that by (3.33) $\mathbf{g}^0 \in \mathrm{Im}\mathsf{Q}^T$ and by (3.35) $\mathrm{Im}\mathsf{Q}^T = \mathsf{A}\mathcal{V}$, so that $\mathbf{g}^0 \in \mathsf{A}\mathcal{V}$. Using the definition of $\mathbf{z}^0$ and (5.143), we have $\mathbf{z}^0 \in \mathsf{A}\mathcal{V}$.

To finish the proof by induction, let us assume that $\mathbf{z}^k \in \mathsf{A}\mathcal{V}$. Since

$$\mathbf{g}^k = \mathsf{Q}^T \mathsf{A}\mathsf{Q}\mathbf{z}^k - \mathsf{Q}^T\mathbf{b} = \mathsf{A}\mathsf{Q}\mathbf{z}^k + \mathbf{g}^0,$$

we have $\mathbf{g}^k \in \mathsf{A}\mathcal{V}$. We shall use this simple observation to examine separately the three possible steps of the MPRGP algorithm of Sect. 5.8.1 that can be used to generate $\mathbf{z}^{k+1}$.

Let us first assume that $\mathbf{z}^{k+1}$ is generated by the proportioning step. Then

$$\mathbf{z}^{k+1} = \mathbf{z}^k - \alpha_{cg}\boldsymbol{\beta}(\mathbf{z}^k).$$

Using the definition of the chopped gradient, it is rather easy to check that $\boldsymbol{\beta}(\mathbf{z}^k) \in \mathcal{V}_0$. Since $\mathcal{V}_0 \subseteq \mathsf{A}\mathcal{V}$, $\mathsf{A}\mathcal{V}$ is a subspace of $\mathbb{R}^n$, and $\mathbf{z}^k \in \mathsf{A}\mathcal{V}$ by the assumptions, this proves that $\mathbf{z}^{k+1} \in \mathsf{A}\mathcal{V}$ when it is generated by the proportioning step.

Before examining the other two steps, observe that $\boldsymbol{\varphi}(\mathbf{z}^k) - \mathbf{g}^k \in \mathcal{V}_0$, so that

$$\boldsymbol{\varphi}(\mathbf{z}^k) = \left(\boldsymbol{\varphi}(\mathbf{z}^k) - \mathbf{g}^k\right) + \mathbf{g}^k \in \mathsf{A}\mathcal{V}.$$

Thus

$$\mathbf{z}^k - \alpha\boldsymbol{\varphi}(\mathbf{z}^k) \in \mathsf{A}\mathcal{V}$$

for any $\alpha \in \mathbb{R}$. Using the first inclusion of (5.143), we get that

$$P_{\Omega_B}\left(\mathbf{z}^k - \overline{\alpha}\boldsymbol{\varphi}(\mathbf{z}^k)\right) \in \mathsf{A}\mathcal{V}$$

for any $\overline{\alpha}$ of Algorithm 5.8. This proves that $\mathbf{z}^{k+1} \in \mathsf{A}\mathcal{V}$ for $\mathbf{z}^{k+1}$ generated by the expansion step. To finish the proof, observe that the conjugate direction $\mathbf{p}^k$ is either equal to $\boldsymbol{\varphi}(\mathbf{z}^k)$, or it is defined by the recurrence (see (5.15))

$\mathbf{p}^{k+1} = \boldsymbol{\varphi}(\mathbf{z}^k) - \beta\mathbf{p}^k$ starting from the restart $\mathbf{p}^{s+1} = \boldsymbol{\varphi}(\mathbf{z}^s)$. In any case, $\mathbf{p}^k \in \mathsf{A}\mathcal{V}$. Since we assume that $\mathbf{z}^k \in \mathsf{A}\mathcal{V}$ and the iterate $\mathbf{z}^{k+1}$ generated by the conjugate gradient step is a linear combination of $\mathbf{z}^k$ and $\mathbf{p}^k$, this completes the proof. □

It follows that we can obtain the correction $\widehat{\mathbf{z}}$ which solves the auxiliary problem by the standard MPRGP algorithm. Since the iterations are reduced to the subspace, the *projector preconditions all three types of steps* and we can give an improved bound on the rate of convergence. The solution $\widehat{\mathbf{x}}$ of the bound constrained problem (5.1) with the feasible set (5.139) can be expressed by $\widehat{\mathbf{x}} = \mathbf{x}^0 + \widehat{\mathbf{z}}$. For convenience of the reader, we give here the complete algorithm for the solution of the preconditioned problem (5.142).

Algorithm 5.13. MPRGP projection preconditioning correction.

Given a symmetric positive definite matrix A *of the order* n *and* $\mathbf{b}, \boldsymbol{\ell} \in \mathbb{R}^n$*; choose a full column rank matrix* $\mathsf{U} \in \mathbb{R}^{m\times n}$, $\mathbf{g}^0 = -\mathsf{Q}^T\mathbf{b}$, $\mathbf{x}^0 = \mathsf{P}\mathsf{A}^{-1}\mathbf{b}$, $\mathbf{z}^0 = P_{\Omega_B}(\mathbf{g}^0)$, $\Gamma > 0$, *and* $\overline{\alpha} \in (0, 2\|\mathsf{A}\mathsf{Q}\|^{-1}]$, *where* P *is defined by (5.140) and* $\mathsf{Q} = \mathsf{I} - \mathsf{P}$.

Step 0. {*Initialization.*}
  Set $k = 0$, $\mathbf{g} = \mathsf{A}\mathsf{Q}\mathbf{z}^0 + \mathbf{g}^0$, $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{z}^0)$
  while $\|\mathbf{g}^P(\mathbf{z}^k)\|$ *is not small*
    if $\|\boldsymbol{\beta}(\mathbf{z}^k)\|^2 \le \Gamma^2\widetilde{\boldsymbol{\varphi}}(\mathbf{z}^k)^T\boldsymbol{\varphi}(\mathbf{z}^k)$
Step 1. {*Proportional* $\mathbf{z}^k$*. Trial conjugate gradient step.*}
      $\alpha_{cg} = \mathbf{g}^T\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathsf{Q}\mathbf{p}$, $\mathbf{y} = \mathbf{z}^k - \alpha_{cg}\mathbf{p}$
      $\alpha_f = \max\{\alpha : \mathbf{z}^k - \alpha\mathbf{p} \in \Omega_B\} = \min\{(z_i^k - \ell_i)/p_i : p_i > 0\}$
      if $\alpha_{cg} \le \alpha_f$
Step 2. {*Conjugate gradient step.*}
        $\mathbf{z}^{k+1} = \mathbf{y}$, $\mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathsf{Q}\mathbf{p}$
        $\beta = \boldsymbol{\varphi}(\mathbf{y})^T\mathsf{A}\mathsf{Q}\mathbf{p}/\mathbf{p}^T\mathsf{A}\mathsf{Q}\mathbf{p}$, $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{y}) - \beta\mathbf{p}$
      else
Step 3. {*Expansion step.*}
        $\mathbf{z}^{k+\frac{1}{2}} = \mathbf{z}^k - \alpha_f\mathbf{p}$, $\mathbf{g} = \mathbf{g} - \alpha_f\mathsf{A}\mathsf{Q}\mathbf{p}$
        $\mathbf{z}^{k+1} = P_{\Omega_B}(\mathbf{z}^{k+\frac{1}{2}} - \overline{\alpha}\boldsymbol{\varphi}(\mathbf{z}^{k+\frac{1}{2}}))$
        $\mathbf{g} = \mathsf{A}\mathsf{Q}\mathbf{z}^{k+1} + \mathbf{g}^0$, $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{z}^{k+1})$
      end if
    else
Step 4. {*Proportioning step.*}
      $\mathbf{d} = \boldsymbol{\beta}(\mathbf{z}^k)$, $\alpha_{cg} = \mathbf{g}^T\mathbf{d}/\mathbf{d}^T\mathsf{A}\mathsf{Q}\mathbf{d}$
      $\mathbf{z}^{k+1} = \mathbf{z}^k - \alpha_{cg}\mathbf{d}$, $\mathbf{g} = \mathbf{g} - \alpha_{cg}\mathsf{A}\mathsf{Q}\mathbf{d}$, $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{z}^{k+1})$
    end if
    $k = k + 1$
  end while
Step 5. {*Return (possibly inexact) solution.*}
  $\widetilde{\mathbf{x}} = \mathbf{z}^k + \mathbf{x}^0$

To describe the improved bound on the rate of convergence, let us denote, as in Sect. 3.7.4, the gap

$$\gamma = \|\mathsf{R}_{\mathsf{A}\mathcal{U}} - \mathsf{R}_{\mathcal{E}}\|$$

between $\mathsf{A}\mathcal{U}$ and the m-dimensional subspace $\mathcal{E}$ spanned by the eigenvectors corresponding to the m smallest eigenvalues

$$\lambda_{n-m+1} \geq \cdots \geq \lambda_{\min}$$

of A, so that the smallest nonzero eigenvalue $\overline{\lambda}_{\min}$ of $\mathsf{Q}^T\mathsf{A}\mathsf{Q}$ satisfies by Theorem 3.6

$$\overline{\lambda}_{\min} \geq \sqrt{(1-\gamma^2)\lambda_{n-m}^2 + \gamma^2\lambda_{\min}^2} \geq \lambda_{\min}. \tag{5.144}$$

Recall that by (3.36) and $\mathsf{A}\mathsf{Q} = \mathsf{Q}^T\mathsf{A}\mathsf{Q}$

$$\|\mathsf{A}\mathsf{Q}\| \leq \|\mathsf{A}\|.$$

Theorem 5.24. *Let $\{\mathbf{z}^k\}$ denote the sequence generated by Algorithm 5.7 for problem (5.142) with $\overline{\alpha} \in (0, 2\|\mathsf{A}\mathsf{Q}\|^{-1}]$ and $\Gamma > 0$ starting from $\mathbf{z}^0 = P_{\Omega_B}(\mathbf{g}^0)$. Let us denote*

$$f_{0,\mathsf{Q}}(\mathbf{z}) = \frac{1}{2}\mathbf{z}^T\mathsf{Q}^T\mathsf{A}\mathsf{Q}\mathbf{z} + (\mathbf{g}^0)^T\mathbf{z}.$$

Then

$$f_{0,\mathsf{Q}}(\mathbf{z}^{k+1}) - f_{0,\mathsf{Q}}(\widehat{\mathbf{z}}) \leq \eta_\Gamma \left(f_{0,\mathsf{Q}}(\mathbf{z}^k) - f_{0,\mathsf{Q}}(\widehat{\mathbf{z}})\right), \tag{5.145}$$

where

$$\eta_\Gamma = 1 - \frac{\widehat{\alpha}\overline{\lambda}_{\min}}{\vartheta + \vartheta\widehat{\Gamma}^2}, \qquad \widehat{\Gamma} = \max\{\Gamma, \Gamma^{-1}\}, \tag{5.146}$$

$$\vartheta = 2\max\{\overline{\alpha}\|\mathsf{A}\|, 1\}, \quad \widehat{\alpha} = \min\{\overline{\alpha}, 2\|\mathsf{A}\|^{-1} - \overline{\alpha}\}, \tag{5.147}$$

and $\overline{\lambda}_{\min}$ denote the least nonzero eigenvalue of $\mathsf{Q}^T\mathsf{A}\mathsf{Q}$ which satisfies (5.144).

Proof. It is enough to combine Theorem 5.14 with the bounds given by Theorem 3.6. □

The efficiency of preconditioning by conjugate projector depends on the choice of the matrix U whose columns span the subspace which should approximate an invariant subspace spanned by the eigenvectors which correspond to small eigenvalues of A. For the minimization problems arising from the discretization of variational inequalities, U is typically obtained by aggregation of variables using geometrical information or from the coarse discretization, as in the multigrid methods. A numerical example is given in the next section. For references on related topics see Sect. 5.12.

5.11 Numerical Experiments

Here we illustrate the performance of some CG-based algorithms for the bound constrained problem (5.1) on minimization of the cost functions $f_{\mathrm{L},h}$ and $f_{\mathrm{LW},h}$ introduced in Sect. 3.10 subject to bound constraints. All the computations are carried out with $\Gamma = 1$ and $\mathbf{x}^0 = \mathbf{o}$.

5.11.1 Polyak, MPRGP, and Preconditioned MPRGP

Let us first compare the performance of the CG-based algorithms presented in this chapter on minimization of the quadratic function $f_{\mathrm{L},h}$ defined by the discretization parameter h (see page 98) subject to the boundary obstacle ℓ defined by the upper part of the circle with the radius $R = 1$ and the center $S = (1, 0.5, -1.3)$. The boundary obstacle is placed under $\Gamma_c = 1 \times [0, 1]$. Our benchmark is described in more detail in Sect. 7.1; its solution is in Fig. 7.4. Recall that the Hessian $\mathsf{A}_{\mathrm{L},h}$ of $f_{\mathrm{L},h}$ is ill conditioned with $\kappa(\mathsf{A}_{\mathrm{L},h}) \approx h^{-2}$.

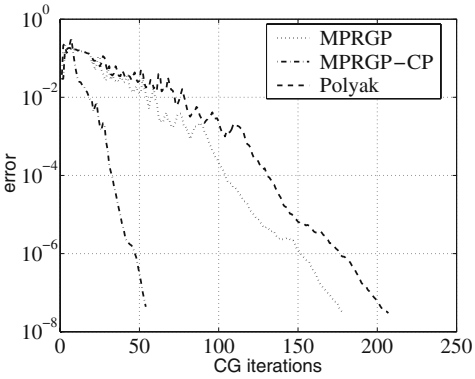

Fig. 5.18. Convergence of Polyak, MPRGP, and MPRGP–CP algorithms

The graph of the norm of the projected gradient (vertical axis) against the numbers of matrix–vector multiplications (horizontal axis) for Algorithm 5.2 (Polyak), Algorithm 5.8 (MPRGP), and MPRGP with preconditioning by the conjugate projector (MPRGP–CP) is in Fig. 5.18. The results were obtained with $h = 1/32$, which corresponds to $n = 1056$ unknowns. The conjugate projector was defined by the aggregation of variables in the squares with 8×8 variables as in Sect. 3.10.1, so that the matrix U has 16 columns. We can see not only that the MPRGP algorithm outperforms Polyak's algorithm, but also that the performance of MPRGP can be considerably improved by preconditioning. The difference between the Polyak and basic MPRGP algorithms is small due to the choice of $\boldsymbol{\ell}$ which makes identification of the active set easy; most iterations of both algorithms were CG steps. The picture can completely change for different $\boldsymbol{\ell}$ as documented in Dostál and Schöberl [74].

5.11.2 Numerical Demonstration of Optimality

To illustrate the concept of optimality, let us consider the class of problems to minimize the quadratic function $f_{\mathrm{LW},h}$ (see page 99) subject to the bound constraints defined by the obstacle as above. The class of problems can be given a mechanical interpretation associated to the expanding spring systems on Winkler's foundation. The spectrum of the Hessian $\mathsf{A}_{\mathrm{LW},h}$ of $f_{\mathrm{LW},h}$ is located in the interval $[2, 10]$. Moreover, $\boldsymbol{\ell} \leq \mathbf{o}$, so that the assumptions of Theorem 5.16 are satisfied.

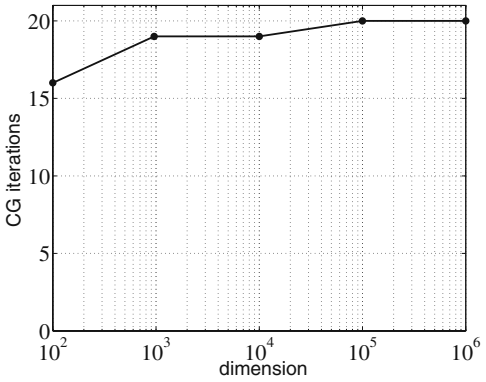

Fig. 5.19. Scalability of MPRGP algorithm

In Fig. 5.19, we can see the numbers of the CG iterations k_n (vertical axis) that were necessary to reduce the norm of the projected gradient by 10^{-6} for the problems with the dimension n ranging from 100 to 1000000. The dimension n on the horizontal axis is in the logarithmic scale. We can see that k_n varies mildly with varying n, in agreement with Theorem 5.16. Moreover, since the cost of the matrix–vector multiplications is in our case proportional to the dimension n of the matrix $\mathsf{A}_{\mathrm{LW},h}$, it follows that the cost of the solution is also proportional to n.

The purpose of the above numerical experiment was just to illustrate the concept of optimality. Realistic classes of problems arise from application of the discretization schemes, such as the finite element method, boundary element method, finite differences, etc., to the elliptic boundary variational inequalities, such as those arising in contact problems of elasticity, in combination with a suitable preconditioning scheme, such as FETI–DP or BETI–DP. An optimal algorithm for the solution of the class of problems arising from the finite element discretization of a model variational inequality with the FETI–DP preconditioning can be found in Chap. 7. More comprehensive related discussion and references can be found in the next section.

5.12 Comments and References

Since the conjugate gradient method was introduced in the celebrated paper by Hestenes and Stiefel [117] as a method for the solution of systems of linear equations, it seems that Polyak [159] was the first researcher who proposed to use the conjugate gradient method to minimize the quadratic cost function subject to bound constraints. Though Polyak assumed the auxiliary problems to be solved exactly, O'Leary [157] observed that this assumption can be replaced by refining the accuracy in the process of solution. In this way she managed to reduce the number of iterations to about a half as compared with the algorithm using the exact solution. The effective theoretically supported strategies for adaptive precision control were presented independently by Friedlander and Martínez with their collaborators [94, 95, 96, 14], and Dostál [41, 42]. Our exposition of inexact Polyak algorithms is based on Dostál [41, 43]. Comprehensive experiments and tests of heuristics can be found in Diniz-Ehrhardt, Gomes-Ruggiero, and Santos [34]. The research was not limited to the convex problems, see also Diniz-Ehrhardt et al. [33].

Many authors fought with an unpleasant consequence of the Polyak strategy which yields a lower bound on the number of iterations in terms of the difference between the numbers of the active constraints in the initial approximation and the solution. Dembo and Tulowitzski [30] proposed the conjugate gradient projection algorithm which could add and drop many constraints in an iteration. Later Yang and Tolle [183] further developed this algorithm with backtracking so that they were able to prove its finite termination property.

An important step forward was development of algorithms with a rigorous convergence theory. On the basis of the results of Calamai and Moré [20], Moré and Toraldo [153] proposed an algorithm that also exploits the conjugate gradients and projections, but its convergence is driven by the gradient projections with the steplength satisfying the sufficient decrease condition. The steplength is found, as in earlier algorithms, by possibly expensive backtracking. In spite of iterative basis of their algorithm, the authors proved that their algorithm preserved the finite termination property of the original algorithm provided the solution satisfies the strict complementarity condition. Friedlander, Martínez, Dostál, and their collaborators combined this result with inexact solution of auxiliary problems [94, 95, 96, 14, 33, 41, 42]. The concept of proportioning algorithm as presented here was introduced by Dostál in [42]. The convergence of the proportioning algorithm was driven by the proportioning step, leaving more room for the heuristic implementation of projections as compared with Moré and Toraldo [153]. The heuristics for implementation of the proportioning algorithm of Dostál [42] can be applied also to the MPRGP algorithm of Sect. 5.8.

The common drawbacks of all the above-mentioned strategies were possible backtracking in search of the gradient projection step and the lack of results on the rate of convergence. A key to further progress were the results by Schöberl [165, 166], who found the bound on the rate of convergence of the

cost function in the energy norm for the gradient projection method with the fixed steplength $\overline{\alpha} \in (0, \|\mathsf{A}\|^{-1}]$ in terms of the spectral condition number of the Hessian matrix. It was observed later by Dostál [45] that this nice result can be plugged into the proportioning algorithm to get a similar result, but with the algorithm which can carry out more efficiently the unconstrained steps. The estimates were extended to $\overline{\alpha} \in (0, 2\|\mathsf{A}\|^{-1}]$ by Dostál [51] (gradient projection) and Dostál, Domorádová, and Sadowská [52] (MPRGP). In our exposition of the MPRGP algorithm, we follow Dostál and Schöberl [74], Dostál [51], and Dostál, Domorádová, and Sadowská [52]. Let us recall that the linear rate of convergence of the cost function for the gradient projection method was proved earlier even for more general problems by Luo and Tseng [146], but they did not make any attempt to specify the constants. Notice that the bound on the coefficient of contraction of the gradient projections in the Euclidean norm is a standard result [12]. The gradient projections were exploited also in the algorithms for more general bound constrained problems, see, e.g., Hager and Zhang [115]. Kučera [138] later modified the algorithm to the minimization of quadratic function subject to separated quadratic constraints.

The attempts to enhance unfeasible iterations into the active set-based methods are usually motivated by an effort to expand effectively the active set, especially in the early stage of computation. Of course, the problem is not how to expand the active set, but how to expand it properly. Our monotonic MPRGP algorithm introduced in Sect. 5.9.3 implements a natural heuristics that any decrease direction is acceptable when we want to expand the active set provided the decrease of the cost function in the unfeasible direction is not surpassed by the increase due to the projection to the feasible set. The algorithm can be considered as a special class of the semismooth Newton method with a globalization strategy. For the semismooth Newton algorithms, see, e.g., Hintermüller, Ito, and Kunisch [118] and Hintermüller, Kovtumenko, and Kunisch [119]. Recent application of Newton-type methods to the contact problem may be found in Hüeber, Stadler, and Wohlmuth [122].

The preconditioning in face was studied by O'Leary [157]. Kornhuber [131, 132] presented nice experimental results and convergence theory for the solution of quadratic programming problems arising from the discretization of boundary variational inequalities with multigrid preconditioning. See also Kornhuber and Krause [133] and Iontcheva and Vassilevski [124]. It turned out that the coarse grid should avoid the constrained variables as in our description of the preconditioning by a conjugate projector, see Domorádová and Dostál [36]. The first implementation of the latter idea can be found in Domorádová [35]. Dostál, Horák, and Stefanica combined the MPRGP algorithm with the FETI–DP domain decomposition method to develop a scalable algorithm for the solution of a boundary variational inequality [70]. For application to contact problems with friction see Dostál and Vondrák [75] and Dostál, Haslinger, and Kučera [63]. A discussion related to application of MPRGP in the cascade algorithm can be found in Braess [16].

Let us finish with a few comments on the bounds on the rates of convergence presented in Sect. 5.6 on the gradient projection method, in Sect. 5.7 on MPGP, and in Sect. 5.8 on MPRGP. Since the coefficient of the Euclidean contraction η_E and the coefficient η_f of the reduction of the cost function for the gradient projection step with the fixed steplength are smaller than the coefficient of reduction of the cost function η_Γ for MPGP and MPRGP, one can doubt superiority of the latter algorithms. However, such doubts are not substantiated. The point is that our estimates are based on the analysis of the worst case for isolated iterations and do not take into account the "global" performance of the conjugate gradient method, which dominates whenever a few consecutive conjugate gradient iterations are carried out; this feature of the CG method is captured by Theorem 3.2. Such global performance is partly captured by our finite termination results and, in the case of MPRGP, also by the result on the rate of convergence of the projected gradient.