A Variational Approach to Sparsity Optimization Based on Lagrange Multiplier Theory

Kazufumi Ito* Karl Kunisch†

October 7, 2013

Abstract

Sparsity optimization for linear least squares problems formulated as non-smooth regularization problems are considered in infinite dimensional sequence spaces ℓ^p with $p \in [0,1]$. Necessary optimality conditions in the format of a complementarity system are obtained. A monotonically convergent scheme is developed for the case $p \in (0,1]$. For the case $p = 0$ a primal dual active set strategy based on the Lagrange multiplier rule is proposed and analyzed for special cases.

Keywords ℓ^p- optimization, sparsity optimization, complementarity condition, non-smooth optimization, Lagrange multipliers, primal-dual active set method.

MSC Classification

1 Introduction

In this paper we discuss optimization problems of the form:

$$\min_{x\in\ell^p} J(x) = \frac{1}{2}|Ax-b|_2^2 + \beta\,|x|_p^p. \tag{1.1}$$

Here $\ell^p = \{x \in \ell^2 : \sum_{k=1}^{\infty} |x_k|^p < \infty\},\ 0 < p \le 1$, is endowed with

$$|x|_p = (\sum_{k=1}^{\infty} |x_k|^p\,)^{1/p},$$

which is a norm if $p = 1$ and a quasi-norm for $0 < p < 1$. We also consider $p = 0$. In this case in (1.1) is replaced by

$$\min_{x\in\ell^p} J(x) = \frac{1}{2}|Ax-b|_2^2 + \beta\,|x|_0$$

where

$$|x|_0 = \sum_{k=1}^{\infty} |x_k|^0 = \text{ number of nonzero elements of } x$$

*Department of Mathematics, North Carolina State University, Raleigh, North Carolina, 27695-8205, USA; research partially supported by the Army Research Office under DAAD19-02-1-0394

†Institut für Mathematik, Karl-Franzens-Universität Graz, A-8010 Graz, Austria, supported in part by the Fonds zur Förderung der wissenschaftlichen Forschung under SFB 32, Mathematical Optimization and Applications in the Biomedical Sciences".

and for a scalar $a \in \mathbb{R}$

$$|a|^0 = \begin{cases} 1 \text{ if } a \neq 0 \\ 0 \text{ if } a = 0 \end{cases}$$

Further we set $\ell^0 = \{x : |x|_0 < \infty\}$. For $x \in \ell^0$ we have that $|x|_p^p \to |x|_0$, as $p \to 0^+$. Throughout it is assumed that $A \in \mathcal{L}(\ell^2)$. We also denote $|x|_2 = (\sum_{k=1}^{\infty} |x_i|^2)^{1/2}$.

Optimization of ℓ^p-functionals as in (1.1) provides an efficient way to extract the essential features of generalized solutions, e.g. in the context of data compression and order reduction methods, with applications arising in signal and image analysis, inverse scattering, de-convolution and tomography problems, and wavelet and generalized Fourier analysis. The literature on sparsity optimization is rapidly increasing. Here we mention e.g., [BL, CT, D, E1, E2, WNF, LZ, RZ, Z, ZDL] and the references therein. More recently sparsity techniques are also receiving increased attention in the optimal control community, we quote e.g. [CCK, HSW, St] in this respect.

One of the objectives of this paper is to derive necessary optimality conditions for (1.1) for $0 \leq p \leq 1$ which are of complementarity type. As a consequence, the nature of the conditions is such that they do not require the a-priori knowledge whether a specific coordinate of an optimal solution is different from zero or not. Rather this distinction is built into the optimality condition itself. To obtain this system we us the quadratic nature of the smooth term of the cost. Another important issue is the development of numerical schemes. This is motivated by fact that due to lack of differentiability of $s \in \mathbb{R} \to |s|^p \in \mathbb{R}^+$ at $s = 0$, the problems (1.1) are non-smooth, non-convex optimization problem, and hence standard algorithms are not readily available.

Let us briefly point to some of the literature that is available for sparsity optimization in the context of ℓ^p, $p \in [0, 1)$ regularization. Especially numerical techniques have been analyzed by many authors and we can therefore only refer to a small selection. In [Z] existence for (1.1) is proven and the asymptotic behavior of solution as the regularization parameter tends to zero is analyzed. The complexity level of the solution to (1.1) as a function of p and β is analyzed in e.g. [CGWY, CXY]. In a very recent paper, solution concepts for the ℓ_0- problem are investigated in [N].- Turning to numerical contribitions, combinatorial techniques are among the natural choices to use for solving the ℓ^0 problem. Greedy algorithms are discussed in [E1, E2], for example. In [LW] mixed integer programming techniques are used aiming at global solutions. An iterative algorithm where each step is obtained by solving an optimization subproblem involving a quadratic term with diagonal Hessian plus the original sparsity-inducing regularizer is proposed and analyzed mostly for the $p = 1$ case in [WNF]. A surrogate functional approach combined with a gradient technique is proposed in [RZ] for the cases $p \in (0, 1]$. In [BL] a general framework for minimization of non-smooth non-convex functionals based on a generalized gradient projection method is analyzed and applied to (1.1), with $p \in (0, 1)$. Iterative thresholding techniques were developed in [BD] and [FW]. In recent papers [LZ, ZDL] penalty decomposition methods are analyzed for wavelet based image restoration and a general class of nonlinear optimization problems with ℓ^0 regularization terms.

Our focus is on the infinite dimensional sequence spaces. The method that we analyze for solving the ℓ^p problem is an iterative algorithm which solves a modified problem where the singularity at the origin is regularized. It is proved that for this algorithm the iterates decrease the cost monotonically. The method that we propose for the ℓ^0 problem is of Newton type and hence distinctly different from previously considered algorithms.

The outline of the paper is as follows. In Section 2 the necessary optimality condition for solutions to (1.1) is derived for $0 \leq p \leq 1$. The complementarity conditions for the cases $p = 0$ and $p = 1$ are given as well. The asymptotic behavior of the minimizers as $\beta \to 0^+$ is analyzed in Section 3. In Section 4 a monotone fixed point algorithm for solving a regularized version of

(1.1), $p \in (0,1)$ is analyzed. Section 5 is devoted to an augmented Lagrangian formulation and a primal-dual active set method for the case $p = 0$. It is based on the necessary optimality and the Lagrange multiplier approach in Section 2.2. In Section 6 we briefly describe some numerical results obtained for the primal-dual active set method.

2 Existence and Necessary Optimality

In this section we establish existence and derive necessary optimality conditions for a minimizer of (1.1). For convenience we recall that ℓ^r with $1 < r < \infty$ are reflexive Banach spaces, that $(\ell^1)' = \ell^\infty$ and $c_0' = \ell^1$, where c_0 is the space of all convergent sequences with limit 0 endowed with the sup-norm. We also have $c' = \ell^1$, where c is the space of all convergent sequences endowed with the sup-norm, [Y], pg. 115. Moreover for $1 \le r < s \le \infty$ we have $\ell^r \subsetneqq \ell^s$ and $|x|_{\ell^s} \le |x|_{\ell^r}$ for all $x \in \ell^r$.

2.1 Case $0 < p \le 1$

To establish existence we use a re-parametrization according to $x = \gamma(y)$, $y \in \ell^2$, where

$$x_i = \gamma(y)_i = |y_i|^{\frac{2}{p}} \operatorname{sgn}(y_i), \text{ for } i = 1, \ldots, \infty.$$

Note that $y_i = |x_i|^{\frac{p}{2}-1} x_i$, and that $\gamma : \ell^2 \to \ell^p$ is an isomorphism satisfying $|\gamma(y)|_p = |y|_2^{\frac{2}{p}}$. In fact, γ is clearly injective and for every $x \in \ell^p$ the sequence $\{y_i\} = \{|x_i|^{\frac{p}{2}} \operatorname{sgn} x_i\} \in \ell^2$ provides a preimage under γ. It follows that (1.1) is equivalent to

$$\min_{y \in \ell^2} J(y) = \frac{1}{2}|A\gamma(y) - b|_2^2 + \beta |y|_2^2. \tag{2.1}$$

Existence for (2.1) and hence to (1.1) for $p \in (0,1]$ was obtained in [Z]. The proof is elegant and hence to be self-contained we repeat it here.

Lemma 2.1. *The mapping $\gamma : \ell^2 \to \ell^2$ is weakly (sequentially) continuous, i.e. $y^n \to \overline{y}$ weakly in ℓ^2 implies that $\gamma(y^n) \to \gamma(\overline{y})$ weakly in ℓ^2.*

Proof. Let $r = \frac{2}{p} + 1 \in [3, \infty)$ and let r^* denote the conjugate exponent given by $r^* = \frac{p}{2} + 1 \in (1, \frac{3}{2}]$. Then γ is the duality mapping from ℓ^r to ℓ^{r^*}, i.e.

$$(\gamma(y), y)_{\ell^{r^*}, \ell^r} = |\gamma(y)|_{r^*} |y|_r, \quad |\gamma(y)|_{r^*} = |y|_r.$$

If $y^n \to \overline{y}$ weakly in ℓ^2, then $y^n \to \overline{y}$ weakly in ℓ^r. Since the duality mapping $\gamma : \ell^r \to \ell^{r^*}$ is weakly sequentially continuous, see [C], pg. 73, we have $\gamma(y^n) \to \gamma(\overline{y})$ weakly in ℓ^{r^*}. Using that $r^* \le 2$, this implies that $\gamma(y^n) \to \gamma(\overline{y})$ weakly in ℓ^2. □

Theorem 2.1. *For any $\beta > 0$ there exists a solution $\overline{y} \in \ell^2$ to (2.1), and hence a solution $\overline{x} = \gamma(\overline{y}) \in \ell^p$ to (1.1).*

Proof. Let y^n be a minimizing sequence of (2.1) and set $x^n = \gamma(y^n) \in \ell^p$. Then $|y^n|_2^2 = |x^n|_p^p$ and thus $x^n \in \ell^2$. It follows that $\{(x^n, y^n)\}_{n=1}^\infty$ is a bounded sequence in $\ell^2 \times \ell^2$. Hence there exists a subsequence, denoted by the same symbols, such that $\{(x^n, y^n)\}$ converges weakly to some $(\overline{x}, \overline{y}) \in \ell^2 \times \ell^2$. From Lemma 2.1 we have that $\overline{x} = \gamma(\overline{y})$, and by weak lower semi-continuity of norms we find

$$\frac{1}{2}|A\gamma(\overline{y}) - b|_2^2 + \beta |\overline{y}|_2^2 = \inf_{y \in \ell^2} J(y).$$

Hence $\overline{y} \in \ell^2$ is a minimizer of (2.1), and $\overline{x} = \gamma(\overline{y})$ minimizes (1.1). □

Since $s \to \gamma(s)$ is continuous differentiable, we have the necessary optimality condition for $\bar{y}$:

$$\gamma'(y_i)(A_i, A\gamma(y) - b) + \beta\, y_i = 0,$$

where $A_i = Ae_i$ and e_i is the sequence with 1 in the $i-th$ coordinate and 0 otherwise. It does not provide the one for $\bar{x}$. But we have the following necessary optimality condition for $\bar{y}$ or $\bar{x}$.

Theorem 2.2. *If $\bar{x}$ is a global minimizer of* (1.1), *then*

$$(2.2)\qquad \begin{cases} \bar{x}_i = 0 & \text{if } |(A_i, f_i)| < \mu_i \\ \\ (A_i, A\bar{x} - b) + \dfrac{\beta p\, \bar{x}_i}{|\bar{x}_i|^{2-p}} = 0 & \text{if } |(A_i, f_i)| > \mu_i, \end{cases}$$

where $f_i = b - A\bar{x} + A_i\bar{x}_i$ and $\mu_i = \beta^{\frac{1}{2-p}}(2-p)(2(1-p))^{-\frac{1-p}{2-p}}\, |A_i|_2^{1-\frac{p}{2-p}}$.

If $|(A_i, f_i)| = \mu_i$, then $\bar{x}_i = 0$ or $\bar{x}_i = (\frac{2\beta(1-p)}{|A_i|_2^2})^{\frac{1}{2-p}}\, sgn((A_i, f_i))$.

Proof. Suppose at first that $p \in (0,1)$. The case $p = 1$ will be considered below. If $\bar{x}$ is a global minimizer of (1.1), then $\bar{x}_i \in \mathbb{R}$ minimizes

$$(2.3)\qquad F(x_i) = \frac{1}{2}|A_i x_i - f_i|_2^2 + \beta\, |x_i|^p = \frac{1}{2}|A_i|_2^2 x_i^2 - (A_i, f_i)x_i + \frac{1}{2}|f_i|_2^2 + \beta\, |x_i|^p.$$

It is convenient to note that f_i can be equivalently expressed as

$$f_i = b - A\tilde{x}, \quad \tilde{x}_k = \begin{cases} 0 & k = i \\ \\ \bar{x}_k & k \neq i. \end{cases}$$

Inspection of $F(x_i)$ shows that 0 is a local minimizer of $F(x_i)$. It is the only minimizer if $A_i = 0$. Henceforth we assume that $A_i \neq 0$. If $x_i = z > 0$ is another local minimizer of (2.3), then

$$(2.4)\qquad |A_i|_2^2\, z - (A_i, f_i) + \frac{\beta p}{z^{1-p}} = 0$$

Equation (2.4) has a solution provided that

$$(2.5)\qquad (A_i, f_i) \geq |A_i|_2^{\frac{2(1-p)}{2-p}}\, (p\,\beta)^{\frac{1}{2-p}}(1-p)^{\frac{p-1}{2-p}}(2-p).$$

This follows by requiring that $F'(\xi) \leq 0$ where $F''(\xi) = 0$, i.e. $\xi = |A_i|_2^{\frac{2}{p-2}}(\beta p)^{\frac{1}{2-p}}(1-p)^{\frac{1}{2-p}}$. If $F'(\xi) < 0$, i.e. if the inequality in (2.5) is strict, then (2.4) has two solutions, the smaller one corresponding the a local maximum, the larger to a local minimum of $F(x_i)$. In this case then $F(x_i)$ has two local minima, 0 and $z > 0$. To decide whether 0 or z is the global minimizer, we first analyze under which condition $F(z) = F(0)$, i.e.,

$$(2.6)\qquad \frac{1}{2}|A_i|_2^2|z|^2 - (A_i, f_i)z + \beta\, z^p = 0.$$

Then (2.3) has the two distinct global minima. Note that (2.4)–(2.6) are equivalent to

$$\begin{cases} |A_i|_2^2\, z^{2-p} - (A_i, f_i)z^{1-p} + \beta\, p = 0 \\ \\ \dfrac{1}{2}|A_i|_2^2\, z^{2-p} - (A_i, f_i)z^{1-p} + \beta = 0. \end{cases}$$

Subtracting these equation we have

$$|A_i|_2^2\, z^{2-p} = 2\beta(1-p)$$

and thus

$$\bar{z} = \left(\frac{2\beta(1-p)}{|A_i|_2^2}\right)^{\frac{1}{2-p}}. \tag{2.7}$$

satisfies $F(\bar{z}) = F(0)$. Let us set

$$\mu_i = |A_i|_2^2 \bar{z} + \frac{\beta p}{\bar{z}^{1-p}} = |A_i|_2^{\frac{2(1-p)}{2-p}}\, \beta^{\frac{1}{2-p}} (2(1-p))^{\frac{p-1}{2-p}} (2-p).$$

Then from (2.4) we have $F(0) = F(\bar{z})$ if $(A_i, f_i) = \mu_i$. A short computation shows that μ_i is larger than the expression on the right hand side of (2.5). Moreover, still denoting by z the second local solution of (2.3), we have $F(z) < F(0)$ if and only if $z > \bar{z}$, which is the cases if and only if

$$(A_i, f_i) = |A_i|_2^2\, z + \frac{\beta p}{z^{1-p}} > \mu_i$$

Similarly, $F(z) > F(0)$ if and only if $z < \bar{z}$, which holds if and only of $(A_i, f_i) < \mu_i$. Thus, if the global minimizer of $F(x_i)$ is nonnegative, then necessarily $\bar{x}_i = 0$, if $(A_i, f_i) < \mu_i$, and $\bar{x}_i = z$ satisfying (2.4), if $(A_i, f_i) > \mu_i$. If $|(A_i, f_i)| = \mu_i$, then $\bar{x}_i = 0$ or $\bar{x}_i = (\frac{2\beta(1-p)}{|A_i|_2^2})^{\frac{1}{2-p}}\, sgn((A_i, f_i))$.
- For $p \to 0^+$, we have $\bar{z} \to \dfrac{\sqrt{2\beta}}{|A_i|_2}$ and $\mu_i \to \sqrt{2\beta}\, |A_i|_2$. The case of non-positive minima can be treated analogously, by noting that $F(x_i) = \tilde{F}(-x_i)$ where $\tilde{F}(x_i) = \frac{1}{2}|A_i|^2 x_i^2 + (A_i, f_i)x_i + \frac{1}{2}|f_i|_2^2 + \beta\, |x_i|^p$.

The case $p = 1$ can be treated along the same lines, if the following modifications are taken into account: $\xi = \bar{z} = 0$. In particular (2.3) has a unique global minimum which is $x_i = 0$ if $|(A_i, f_i)| \le \mu_i$, and $x_i \neq 0$, satisfying the second equation in (2.2) if $|(A_i, f_i)| > \mu_i$. Moreover $\mu_i = \beta$ for $p = 1$.

□

From Theorem 2.2 it follows that a minimizer to (1.1) is not necessarily unique. Moreover from its proof we obtain the following corollary.

Corollary 2.1. *If $\bar{x}_i \neq 0$ then $|\bar{x}_i| \ge \left(\frac{2\beta(1-p)}{|A_i|_2^2}\right)^{\frac{1}{2-p}}$.*

Indeed the second local solution to (2.3) satisfies necessarily $z \ge \bar{z}$ with $\bar{z}$ given in (2.7).

In [BL] a necessary optimality condition is obtained for nonlinear problems regularized by $|x|_p^p$. It considers separately the inactive components with $\bar{x}_i \neq 0$ and the active ones. Here we exploit the quadratic nature of the fit-to-data term to obtain a necessary optimality condition which is of complementarity type, separating the active components of $\bar{x}$ from the inactive ones by the sign of $|(A_i, f_i)| - \mu_i$.

The following result addresses sparsity of the solution $\bar{x}$ of (1.1). The first part is analogous to a result already contained in [BL]. The second part is applicable in the case that A is close to an orthogonal operator.

Proposition 2.1. *Let $\bar{x}$ denote a global minimizer of* (1.1)*. Then we have:*

(a) $\#\{i : \bar{x}_i \neq 0\} \leq |b|_2^2 \big(\frac{(2\beta)^{\frac{2}{p}}}{\sup_{i\in\mathbb{N}} |A_i|_2^2}(1-p)\big)^{\frac{p}{p-2}}$.

(b) If $|(A_i, b)| < \mu_i - \sup_{j\neq i} |(A_i, A_j)| \sum_{j=1}^{\infty} |\bar{x}_j|$, *then* $\bar{x}_i = 0$

Proof. To verify (a) note that by Corollary 2.1

$$\frac{1}{2}|b|_2^2 \geq \frac{1}{2}|A\bar{x} - b|_2^2 + \beta\, |\bar{x}|_p^p \geq \#\{i : \bar{x}_i \neq 0\}\, \beta \,\big(\frac{2\beta(1-p)}{\sup_{i\in\mathbb{N}} |A_i|_2^2} \big)^{\frac{p}{2-p}}$$

and hence

$$\#\{i : \bar{x}_i \neq 0\} \;\leq\; \frac{|b|_2^2}{2\beta} \big(\frac{2\beta(1-p)}{\sup_i |A_i|_2^2} \big)^{\frac{p}{p-2}}.$$

Turning to (b), our assumption implies that

$$|(A_i, f_i)| = |(A_i, b - A\bar{x} + A_i\bar{x}_i)| = |(A_i, b) - \sum_{j\neq i}(A_i, A_j)\bar{x}_j|$$

$$\leq |(A_i, b)| + |\sum_{j\neq i}(A_i, A_j)\bar{x}_j| \leq |(A_i, b)| + \sup_{j\neq i} |(A_i, A_j)| \sum_{j=1}^{\infty} |\bar{x}_j| < \mu_i,$$

and hence by Theorem 2.2 we have $\bar{x}_i = 0$. $\square$

2.2 Case $p = 0$

In this section we consider the case $p = 0$, i.e,

$$\text{(2.8)} \qquad \min \quad \tfrac{1}{2}|Ax - b|_2^2 + \beta\, |x|_0,$$

where $|x|_0 =$ number of nonzero elements of $x \in \ell^2$. We shall assume throughout that $A \in \mathcal{L}(\ell^2)$.

The $|\cdot|_0$ functional satisfies the triangle inequality:

$$|x_1 + x_2|_0 \leq |x_1|_0 + |x_2|_0,$$

and it is also weakly lower semi-continuous in the sense of the following lemma.

Lemma 2.2. *If $x_n \rightharpoonup x$ in ℓ^2 and if $|x_n|_0 \leq L$ for all n sufficiently large, then $|x|_0 \leq L$ and $|x|_0 \leq \liminf_{n\to\infty} |x_n|_0$.*

Proof. First note that weak convergence in ℓ^2 implies convergence of each coordinate, $(x_n)_i \to x_i$. Assume that $|x|_0 \geq L+1$, i.e. x contains more than L nontrivial entries. Let $\{i_\ell\}$ be these entries in increasing order. Consider $\{x_{i_\ell}\}_{\ell=1}^{L+1}$ and let $\epsilon = \{\min |x_{i_\ell}| : \ell = 1, \ldots, L+1\}$, which satisfies $\epsilon > 0$. Since $x_{i_\ell}^n \to x_{i_\ell}$ for $\ell = 1, \ldots, L+1$, there exists $n = n(\epsilon)$ such that $x_{i_\ell}^n \neq 0$ for all $n \geq n(\epsilon)$. This is a contradiction. Hence i_L is the largest index for which $x_i \neq 0$, and $|x|_0 \leq L$. Using coordinate convergence of x_n to x and $|x|_0 < \infty$ it easily follows that $|x|_0 \leq \liminf_{n\to\infty} |x_n|_0$. $\square$

The second part assertion of the previous lemma was also observed in the recent paper [WLMC].

We have the following result on the existence of a solution to (2.8). Here we assume that A has closed range. Otherwise, if A is ill-posed we need to add a quadratic regularization term as done in Section 5. Existence for (2.8) was also considered in [G], where lack coercivity of the ℓ^0-functional was pointed out and in [Lo], where an example for nonexistence is given with a matrix A which does not satisfy the assumptions which are imposed in the following result.

Theorem 2.3. *Assume that A has closed range, and that its nullspace is finite dimensional. Then problem* (2.8) *admits a solution.*

Proof. The proof follows those of abstract existence results given in [AT], Chapter 3, and [BBGT], but none of these results fit directly to the situation given here. We let $\{\epsilon_n\} \subset (0,1)$ be a sequence converging to 0 from above and consider the family of auxiliary problems

$$\min_{x\in H} J(x) + \frac{\epsilon_n}{2}|x|_2^2. \tag{2.9}$$

where

$$J(x) = \frac{1}{2}|Ax-b|_2^2 + \beta\,|x|_0.$$

Since $\epsilon_n > 0$ every minimizing sequence for (2.9) is bounded. Extracting a weakly convergent subsequence and using weak lower semi-continuity of J and the norm functional, existence of a solution $x_n \in \ell^2$ for (2.9) can be argued in a standard manner.

If the sequence of solutions $\{x_n\}$ to (2.9) is bounded in ℓ^2, then there exists a weakly convergent subsequence, denoted by the same symbols, and $\bar{x}$, such that $x_n \rightharpoonup \bar{x}$. Since $|x_n|_0 \le \frac{1}{2\beta}|b|_2^2$ for all n, Lemma 2.2 is applicable. Passing to the limit $\epsilon_n \to 0$ in

$$J(x_n) + \frac{\epsilon_n}{2}|x_n|_2^2 \le J(x) + \frac{\epsilon_n}{2}|x|_2^2 \text{ for all } x \in \ell^2,$$

we have

$$J(\bar{x}) \le J(x) \text{ for all } x \in \ell^2,$$

and thus $\bar{x}$ is a minimizer for J.

Henceforth we show that the assumption that $\{x_n\}$ is unbounded in ℓ^2 leads to a contradiction. In fact, if $\{x_n\}$ is not bounded, then there exists a subsequence, denoted by the same symbol, such that

$$\lim_{n\to\infty} |x_n| = \infty, \text{ and } \frac{x_n}{|x_n|} \rightharpoonup \bar{x}, \text{ for some } \bar{x} \in \ell^2.$$

Using the assumption that A has closed range, every element $x \in \ell^2$ can be uniquely decomposed as $x = x^1 + x^2 \in R(A^*) + N(A)$.

Since $\{x_n\}$ is a minimizing sequence, there exists a constant $K > 0$ such that

$$J(x_n) = \frac{1}{2}|Ax_n - b|_2^2 + \beta\,|x_n|_0 \le K, \text{ for all } n. \tag{2.10}$$

Consequently with $x_n = x_n^1 + x_n^2 \in R(A^*) + N(A)$ we find $0 \le |Ax_n^1|_2 - 2(b, Ax_n^1)_2 + |b|_2^2 \le K$ and

$$0 \le \big|\, A(\frac{x_n^1}{|x_n|_2})\,\big|_2^2 - 2\frac{1}{|x_n|_2}(\,A^*b, \frac{x_n}{|x_n|_2}\,)_2 + \frac{|b|_2^2}{|x_n|_2} \to 0. \tag{2.11}$$

Using that $\{\frac{x_n}{|x_n|_2}\}$ is bounded, we deduce from (2.11) that $|A(\frac{x_n^1}{|x_n|_2})\,| \to 0$. By the closed range theorem this implies that $\frac{x_n^1}{|x_n|_2} \to \bar{x}^1 = 0$ in ℓ^2. Since $\frac{x_n^2}{|x_n|_2} \rightharpoonup \bar{x}^2$ and since by assumption dim $N(A) < \infty$ it follows that

$$\frac{x_n}{|x_n|_2} \to \bar{x} = \bar{x}^2 \text{ strongly in } \ell^2. \tag{2.12}$$

Next we argue that there exists some $\rho > 0$ such that

$$J(x_n - \rho\bar{x}) = \frac{1}{2}|A(x_n - \rho\bar{x}) - b|_2^2 + \beta\,|x_n - \rho\bar{x}|_0 \le J(x_n) = \frac{1}{2}|Ax_n - b|_2^2 + \beta\,|x_n|_0 \tag{2.13}$$

for all n sufficiently large. Since $\bar{x} = \bar{x}^2 \in N(A)$ this will be implied by showing that

$$|(x_n^1 + x_n^2 - \rho\bar{x})_i|^0 \le |(x_n^1 + x_n^2)_i|^0 \quad \text{for all } i,$$

and for all n sufficiently large. Only the coordinates for which $(x_n^1 + x_n^2)_i = 0$ with $(\bar{x}^2)_i \neq 0$ require our attention. Since $\{\left|\frac{x_n}{|x_n|_2}\right|_0\} \le K$ for all n sufficiently large and $\frac{x_n}{|x_n|_2} \to \bar{x}^2$ we have $|\bar{x}^2|_0 \le K$, and hence $\bar{x}^2$ has bounded support, i.e. there exists $\tilde{i}$ such that $(\bar{x}^2)_i = 0$ for all $i \ge \tilde{i}+1$. For $i \in \{1, \dots, \tilde{i}\}$ we define $\mathcal{I}_i = \{n : (x_n^1 + x_n^2)_i = 0, (\bar{x}^2)_i \neq 0\}$. These sets are finite. In fact, if $\mathcal{I}_i$ is infinite for some $i \in \{1, \dots, \tilde{i}\}$, then $\lim_{n\to\infty, n\in\mathcal{I}_i} \frac{1}{|x_n|_2}(x_n^1 + x_n^2)_i = 0$. Since $\lim_{n\to\infty} \frac{1}{|x_n|_2}(x_n^1)_i = 0$ this implies that $(\bar{x}^2)_i = 0$, which is a contradiction. Taking $\tilde{n}$ as the maximal index in $\{\mathcal{I}_i : i \in \{1, \dots, \tilde{i}\}\ \}$ we have that $(x_n^1 + x_n^2)_i \neq 0$ for all $i \in \{1, \dots, \tilde{i}\}$ and $n \ge \tilde{n}$. Summarizing we showed that (2.13) holds for any $\rho > 0$ and $n \ge \tilde{n}$.

From (2.13) we have for $n \ge \tilde{n}$:

$$J(x_n) + \frac{\epsilon}{2}|x_n|_2^2 \le J(x_n - \rho\bar{x}) + \frac{\epsilon}{2}|x_n - \rho\bar{x}|_2^2 \le J(x_n) + \frac{\epsilon}{2}|x_n - \rho\bar{x}|_2^2.$$

It follows that

$$|x_n|_2 \le |x_n - \rho\bar{x}|_2 = |x_n - \rho\frac{x_n}{|x_n|_2} + \rho(\frac{x_n}{|x_n|_2} - \bar{x})|_2 \le |x_n|_2(1 - \frac{\rho}{|x_n|_2}) + \rho|\frac{x_n}{|x_n|_2} - \bar{x}|_2$$

This implies that

$$1 \le |\frac{x_n}{|x_n|_2} - \bar{x}|_2,$$

which give a contradiction to (2.12), and concludes the proof.

□

An alternative to the above assumptions which guarantee existence of a solution to (2.8) is to assume radial unboundedness, i.e.

$$\text{(H1)} \qquad |Ax|_2 \to \infty \text{ for } |x|_2 \to \infty.$$

Theorem 2.4. *Assume that* (H1) *holds. Then problem* (2.8) *admits a solution $\bar{x}$. Moreover any weak-cluster point in ℓ^2 (of which there exists at least one) of solutions $\{x^p\}$ to* (1.1) *as $p \to 0^+$ is a solution to* (2.8).

Proof. (i) Let $\{x^n\}$ be a minimizing sequence of (2.8). Then $|x^n|_0 \le \frac{1}{2}|b|_2$. By (H1) the sequence $\{x^n\}$ is bounded in ℓ^2. Consequently there exists a weakly convergent subsequence, denoted by the same symbol with weak limit $\bar{x} \in \ell^2$. Using weak lower semi-continuity of the $|\cdot|_2$ norm and the $|\cdot|_0$ functional we may pass to the limit in

$$\lim_{n\to\infty} \frac{1}{2}\,|Ax^n - b|_2^2 + \beta\,|x^n|_0 \le \frac{1}{2}\,|Ax - b|_2^2 + \beta\,|x|_0, \text{ for all } x \in l^2,$$

to obtain that

$$\min \frac{1}{2}|A\bar{x} - b|_2^2 + \beta\,|\bar{x}|_0 \;\le\; \min \frac{1}{2}|Ax - b|_2^2 + \beta\,|x|_0, \text{ for all } x \in \ell^2, \tag{2.14}$$

and hence $\bar{x}$ is a solution to (2.8).

Next, let $\{x^p\}$ be a solution sequence for (1.1) with $p > 0$, as $p \to 0^+$. Thus we have

$$\frac{1}{2}|Ax^p - b|_2^2 + \beta\,|x^p|_p^p \le \frac{1}{2}|Ax - b|_2^2 + \beta\,|x|_p^p \text{ for all } x \in \ell_0.$$

By (H1) the sequence $\{x^p\}$ is bounded in ℓ^2. Hence there exists a subsequence, denoted by the same symbols that converges weakly in ℓ^2 to some $\bar{x}$ in ℓ^2. It satisfies that $x_i^p \to \bar{x}_i$ for each $i = 1, \ldots,$. For $\epsilon > 0$ and $x \in \ell^2$ let

$$N_\epsilon(x)_i = \begin{cases} 1 & \text{if } |x_i| > \epsilon \\ 0 & \text{if } |x_i| \le \epsilon. \end{cases}$$

Then

$$\frac{1}{2}|Ax^p - b|_2^2 + \beta\,|N_\epsilon(x^p) \odot x^p|_p^p \le \frac{1}{2}|Ax - b|_2^2 + \beta\,|x|_p^p \ \text{ for all } x \in \ell_0,$$

where $(a \odot b)_i = a_i b_i$. Using the fact that $|N_\epsilon(\bar{x})|_0 < \infty$ we find that

$$\frac{1}{2}|A\bar{x} - b|_2^2 + \beta\,|N_\epsilon(\bar{x}) \odot \bar{x}|_0 \le \frac{1}{2}|Ax - b|_2^2 + \beta\,|x|_0 \ \text{ for all } x \in \ell_0.$$

Since $\epsilon > 0$ was arbitrary this implies that

$$\frac{1}{2}|A\bar{x} - b|_2^2 + \beta\,|\bar{x}|_0 \le \frac{1}{2}|Ax - b|_2^2 + \beta\,|x|_0 \ \text{ for all } x \in \ell_0,$$

and hence $\bar{x}$ is a solution to (2.8).

□

Theorem 2.5. *If $\bar{x}$ is a solution to* (2.8)*, then the following necessary optimality holds:*

$$\begin{cases} \bar{x}_i = 0 & \text{if } |(A_i, f_i)| < \sqrt{2\beta}|A_i|_2 \\ \\ (A_i, A\bar{x} - b) = 0 & \text{if } |(A_i, f_i)| > \sqrt{2\beta}|A_i|_2, \end{cases} \tag{2.15}$$

where $f_i = b - A\bar{x} + A_i\bar{x}_i$. For the second case in (2.15)*, $|(A_i, f_i)| > \sqrt{2\beta}|A_i|_2$ is equivalent to $|\bar{x}_i| > \dfrac{\sqrt{2\beta}}{|A_i|_2}$. If $|(A_i, f_i)| = \sqrt{2\beta}|A_i|_2$, then $\bar{x}_i = 0$ or $\bar{x}_i = \frac{\sqrt{2\beta}}{|A_i|_2}\, sgn((A_i, f_i))$.*

Proof. Formally we obtain these conditions by passing to the limit $p \to 0^+$ in the optimality conditions (2.2) for the ℓ^p problems. Here we proceed as in the proof of Theorem 2.2. If $\bar{x}$ is a global minimizer of (2.8), then for each i the coordinate $\bar{x}_i \in \mathbb{R}$ minimizes

$$F(x_i) = \frac{1}{2}|A_i x_i - f_i|_2^2 + \beta\,|x_i|^0 = \frac{1}{2}|A_i|_2^2 x_i^2 - (A_i, f_i)x_i + \frac{1}{2}|f_i|_2^2 + \beta\,|x_i|^0. \tag{2.16}$$

Therefore $\bar{x}_i$ must be either $\frac{(A_i, f_i)}{|A_i|^2}$ or 0. For these values of $\bar{x}_i$ we have $F(\frac{(A_i, f_i)}{|A_i|^2}) = \frac{-(A_i, f_i)^2}{2|A_i|^2} + \frac{1}{2}|f_i|^2 + \beta$ and $F(0) = \frac{1}{2}|f_i|^2$. A case study implies the claimed optimality condition (2.15). □

We turn to the discussion of sparsity of the solutions to (2.8). To guarantee the necessary a-priori bounds we again utilize (H1). From (2.14) with $x = 0$ we have

$$\frac{1}{2}|A\bar{x} - b|_2^2 + \beta|\bar{x}|_0 \le \frac{1}{2}|b|_2^2$$

and hence by (H1) the family of solutions $\bar{x} = \bar{x}(\beta)$ to (2.8) is bounded in ℓ^2, i.e. there exists a constant M such that

$$|\bar{x}(\beta)|_2 \le M \text{ for all } \beta \in (0, \infty). \tag{2.17}$$

For any $i \in \mathbb{N}$ we introduce

$$\alpha_i = \Big(\sum_{j \ne i} |(A_i, A_j)|^2\Big)^{\frac{1}{2}} \in [0, \infty],$$

and note that $\alpha_i = 0$ for all i, if A is orthogonal.

Proposition 2.2. *Suppose that* (H1) *holds and let* $a^2 = \sup_{i\in\mathbb{N}} |A_i|_2^2$. *We have*

(a) $\#\{i : \bar{x}_i \neq 0\} \le \frac{a^2 M^2}{2\beta}$, *and*

(b) if $|(A_i, b)| \le \sqrt{2\beta}|A_i| - \alpha_i M$, *then* $\bar{x}_i = 0$

Proof. By Theorem 2.4 and (2.17) we have

$$M^2 \ge \sum |\bar{x}_i|_2^2 \ge \sum_{i\in\mathcal{I}} \frac{2\beta}{|A_i|_2^2},$$

where $\mathcal{I} = \{i : \bar{x}_i \neq 0\}$. Consequently

$$M^2 \ge \frac{2\beta}{a^2}\,\#(\mathcal{I})$$

and (a) follows. Next we compute

$$|(A_i, f_i)| \le |(A_i, b)| + |\sum_{j\neq i}(A_i, A_j)\bar{x}_j| \le |(A_i, b)| + (\sum_{j\neq i}(A_i, A_j)^2)^{1/2} M = |(A_i, b)| + \alpha_i M,$$

and hence the first case in (2.15) occurs, if $|(A_i, b)| \le \sqrt{2\beta}|A_i|_2 - \alpha_i M$.

Remark 2.1. For those solutions $\bar{x}$ of (2.8) which can be approximated as weak cluster points of solutions $\{x^p\}$, to (1.1) as $p \to 0^+$, we can utilize Proposition 2.1 to establish that $\#\{i : \bar{x}_i \neq 0\} \le \frac{|b|_2^2}{2\beta}$, which has the same asymptotic as Proposition 2.2 (a).

□

2.3 Complementarity Problem for $p = 0$ and $p = 1$

For $p = 1$ the necessary condition of Theorem 2.2 is equivalent to

$$A^*(Ax - b) + \beta\,\lambda = 0, \tag{2.18}$$

$$\begin{cases} \lambda_i = 1 & x_i > 0 \\ \lambda_i \in [-1,\ 1] & x_i = 0 \\ \lambda_i = -1 & x_i < 0. \end{cases} \tag{2.19}$$

In fact, if (2.18)-(2.19) hold, then

$$-(A_i, f_i) + |A_i|^2 x_i + \beta\lambda_i = 0.$$

If $\lambda_i = 1$, $x_i > 0$, then $|A_i|_2^2\, x_i = (A_i, f_i) - \beta = (A_i, f_i) - \mu_i > 0$, and if $\lambda_i = -1$, $x_i < 0$, then $|A_i|_2^2\, x_i = (A_i, f_i) + \beta = (A_i, f_i) + \mu_i < 0$. Consequently $(A_i, A\bar{x} - b) + \beta\frac{x_i}{|x_i|} = 0$ if $|(A_i, f_i)| > \mu_i$, which is the second line of (2.2) for $p = 1$. If $x_i = 0$, $\lambda_i \in [-1, 1]$, then $(A_i, f_i) - \beta\lambda_i = 0$ and consequently $|(A_i, f_i)| \le \mu_i$. This is consistent with the first line in (2.2) and the statement concerning the case $|(A_i, f_i)| = \mu_i$. Conversely the necessary condition of Theorem 2.2 implies (2.18)-(2.19).

Furthermore, the complementarity condition (2.19) can be expressed as

$$\lambda_i = \frac{\lambda_i + c\,x_i}{\max(1, |\lambda_i + c\,x_i|)}, \quad \text{for all } i = 1, \dots, \tag{2.20}$$

for each $c > 0$ [IK]. In fact, if (2.20) holds then $|\lambda_i| \le 1$. If $|\lambda_i + c\,x_i| \le 1$ then $\lambda_i = \lambda_i + c\,x_i$ and thus $x_i = 0$. If $a = |\lambda_i + c\,x_i| > 1$, then $\lambda_i(a-1) = cx_i$, $|\lambda_i| = 1$ and thus $\lambda_i = sign(x_i)$. The converse can be argued analogously.

For $p = 0$ we introduce the multiplier

$$\lambda_i = (A_i, b - A\bar{x}).$$

Then $|(A_i, f_i)|$ can be expressed as $|(A_i, f_i)| = |\lambda_i + |A_i|^2\bar{x}_i|$ and (2.15) becomes

$$\begin{cases} \bar{x}_i = 0 & \text{if } |\lambda_i + |A_i|^2\bar{x}_i| < \sqrt{2\beta}|A_i|_2 \\ \\ \lambda_i = 0 & \text{if } |\lambda_i + |A_i|^2\bar{x}_i| > \sqrt{2\beta}|A_i|_2, \end{cases} \tag{2.21}$$

3 Asymptotic as $\beta \to 0^+$

In this section we discuss the asymptotic of x_β as $\beta \to 0^+$. Let P be the orthogonal projection of ℓ_2 onto $N(A^*)$ and set $\tilde{b} = (I-P)b$. Then,

$$|Ax - b|_2^2 = |Ax - (I-P)b|_2^2 + |Pb|_2^2 = |Ax - \tilde{b}|_2^2 + |Pb|_2^2.$$

Assume $\tilde{b} = (I-P)b \in R(A)$. Consider the minimum norm problem

$$\min \quad |x|_0 \quad \text{subject to } Ax - \tilde{b} = 0. \tag{3.1}$$

Let x_β be a minimizer of (2.8) over $x \in \ell_0$, given $\beta > 0$.

Theorem 3.1. *Assume that there exists $\tilde{x} \in \ell^0$ such that $A\tilde{x} = \tilde{b}$, and let* (H1) *hold. Then every weak cluster point in ℓ^2 of solutions x_β to* (2.8) *as $\beta \to 0^+$ is a minimizer of* (3.1) *and $|Ax_\beta - \tilde{b}|_0 = O(\sqrt{\beta})$.*

Proof. Let $\tilde{x}$ satisfy $A\tilde{x} - \tilde{b} = 0$. Then,

$$\frac{1}{2}|Ax_\beta - \tilde{b}|_2^2 + \beta\,|x_\beta|_0 \le \frac{1}{2}|A\tilde{x} - \tilde{b}|_2^2 + \beta\,|\tilde{x}|_0 = \beta\,|\tilde{x}|_0.$$

and thus $|x_\beta|_0 \le |\tilde{x}|_0$ and $\lim |Ax_\beta - \tilde{b}|_2^2 = 0$ as $\beta \to 0^+$.

The proof is now similar to that of the first part of Theorem 2.4. By (H1) the sequence $\{x_\beta\}_{\beta>0}$ is bounded in ℓ^2 and hence there exists a weak subsequential limit $\bar{x}$ in ℓ^2. It clearly satisfies $A\bar{x} = \tilde{b}$. Moreover $|\bar{x}|_0 < \infty$. Let i_L denote the largest index such that $\bar{x}_i \neq 0$, and define $(\hat{x}_\beta)_i = (x_\beta)_i$ if $i \le i_L$ and equal to zero otherwise. Then

$$\frac{1}{2}|Ax_\beta - \tilde{b}|_2^2 + \beta\,|\hat{x}_\beta|_0 \le \frac{1}{2}|Ax - \tilde{b}|_2^2 + \beta\,|x|_0 \;\; \text{for all } x \in \ell^2.$$

Taking the limit $\beta \to 0^+$ and choosing x as a solution to (3.1) concludes the proof. ∎

4 Monotone Convergent Algorithm

Here we consider the case $p \in (0,1]$. In order to overcome the singularity of $(|s|^p)' = \dfrac{ps}{|s|^{2-p}}$ near $s = 0$, we consider for $\epsilon > 0$ the regularized problem:

$$J_\epsilon(x) = \frac{1}{2}\,|Ax - b|_2^2 + \beta\Psi_\epsilon(|x|^2), \tag{4.1}$$

where for $t \geq 0$

$$\Psi_\epsilon(t) = \begin{cases} \dfrac{p}{2}\dfrac{t}{\epsilon^{2-p}} + (1 - \dfrac{p}{2})\,\epsilon^p & \text{for } 0 \leq t \leq \epsilon^2 \\[2ex] t^{\frac{p}{2}} & \text{for } t \geq \epsilon^2, \end{cases}$$

and $\Psi_\epsilon(|x|^2)$ is short for $\sum_{i=1}^\infty \Psi_\epsilon(|x_i|^2)$. Note that

$$\Psi_\epsilon'(t) = \frac{p}{2\,\max(\epsilon^{2-p}, t^{\frac{2-p}{2}})} \quad \text{for } t \geq 0,$$

and hence $\Psi \in C^1([0,\infty), \mathbb{R})$. The necessary optimality condition for (4.1) is given by

$$A^*Ax + \frac{\beta p}{\max(\epsilon^{2-p}, |x|^{2-p})}\, x = A^*b, \tag{4.2}$$

where the max-operation is interpreted coordinate-wise. To solve (4.2) we consider the iteration procedure:

$$A^*Ax^{k+1} + \frac{\beta p}{\max(\epsilon^{2-p}, |x^k|^{2-p})}\, x^{k+1} = A^*b, \tag{4.3}$$

where the second addend is short for the vector with components $\frac{\beta p}{\max(\epsilon^{2-p}, |x_i^k|^{2-p})}\, x_i^{k+1}$. Multiplying this by $x^{k+1} - x^k$, we obtain

$$\frac{1}{2}(A^*Ax^{k+1}, x^{k+1}) - \frac{1}{2}(A^*Ax^k, x^k) + \frac{1}{2}(A^*A(x^{k+1} - x^k), x^{k+1} - x^k))$$

$$+ \sum_{i=1}^\infty \frac{\beta p}{\max(\epsilon^{2-p}, |x_i^k|^{2-p})}\,\frac{1}{2}\,(|x_i^{k+1}|^2 - |x_i^k|^2 + |x_i^{k+1} - x_i^k|^2) = (A^*b, x^{k+1} - x^k).$$

Below we use that

$$\frac{1}{\max(\epsilon^{2-p}, |x_i^k|^{2-p})}\,\frac{p}{2}\,(|x_i^{k+1}|^2 - |x_i^k|^2) = \Psi_\epsilon'(|x_i^k|^2)(|x_i^{k+1}|^2 - |x_i^k|^2).$$

Since $t \to \Psi_\epsilon(t)$ is concave, we have

$$\Psi_\epsilon(|x_i^{k+1}|^2) - \Psi_\epsilon(|x_i^k|^2) - \frac{1}{\max(\epsilon^{2-p}, |x_i^k|^{2-p})}\,\frac{p}{2}\,(|x_i^{k+1}|^2 - |x_i^k|^2) \leq 0$$

and thus

$$J_\epsilon(x^{k+1}) + \frac{1}{2}(A^*A(x^{k+1} - x^k), x^{k+1} - x^k)) + \sum_{i=1}^\infty \frac{\beta p}{\max(\epsilon^{2-p}, |x_i^k|^{2-p})}\,\frac{1}{2}\,|x_i^{k+1} - x_i^k|^2 \leq J_\epsilon(x^k). \tag{4.4}$$

We have the following convergence result:

Theorem 4.1. *For $\epsilon > 0$ let $\{x_k\}$ be generated by (4.3). Then, $J_\epsilon(x^k)$ is strictly monotonically decreasing, unless there exists some k such that $x^k = x^{k+1}$ and x^k satisfies the necessary optimality condition (4.2). Moreover every weakly convergent subsequence of x_k, of which there exists at least one, converges weakly in ℓ^2 to a solution of (4.2).*

Proof. From (4.4) it follows that $\{x^k\}_{k=1}^\infty$ is bounded in ℓ^2 and hence in ℓ^∞. Consequently from (4.4) there exists $\kappa > 0$ such that

$$J_\epsilon(x^{k+1}) + \frac{1}{2}(A^*A(x^{k+1}-x^k), x^{k+1}-x^k)) + \kappa|x^{k+1}-x^k|_2^2 \le J_\epsilon(x^k). \tag{4.5}$$

This implies the first part of the theorem. From (4.5) we conclude that

$$\sum_{k=0}^\infty |x^{k+1}-x^k|_2^2 < \infty. \tag{4.6}$$

Since $\{x^k\}_{k=1}^\infty$ is bounded in ℓ^2 there exists $\bar{x} \in \ell^2$ and a subsequence such that $x^{k_\ell} \to \bar{x}$ weakly in ℓ^2. By (4.6) moreover $\lim_{\ell\to\infty} x_i^{k_\ell+1} = \lim_{\ell\to\infty} x_i^{k_\ell} = \bar{x}_i$ for all i. Testing (4.3) with e_i, $i = 1, \dots$ and passing to the limit with respect to k, we find that $\bar{x}$ satisfies (4.2). $\square$

5 Augmented Lagrangian Formulation and Primal-Dual Active Set Method

In this section we develop the augmented Lagrangian formulation and the primal-dual active strategy for the sparsity optimization problem (2.8). Let P be a nonnegative self-adjoint operator P, satisfying $\big((A+\alpha P)x, x\big) \ge \gamma|x|_{\ell^2}^2$ for some $\alpha,\ \gamma > 0$ independent of $x \in \ell^2$. We set

$$\Lambda_k = |A_k|_2^2 + \alpha\, P_{kk},$$

and let Λ denote the invertible diagonal operator with entries Λ_k. Here $P_{kk} = \mathcal{P}_k P \mathcal{P}_k$ with $\mathcal{P}_k$ the projection of ℓ^2 onto the $k-th$ component of ℓ^2. Thus, if A is nearly singular, we use $\alpha > 0$ and the regularization functional $\frac{\alpha}{2}(x, Px)$ to regularize (2.8). Consider the associated augmented Lagrangian functional

$$L(x,v,\lambda) = \frac{1}{2}|Ax-b|_2^2 + \frac{\alpha}{2}(Px,x) + \beta\sum_k |v_k|^0 + \sum_k \big(\frac{\Lambda_k}{2}\,|x_k - v_k|^2 + (\lambda_k, x_k - v_k)\big).$$

Given (x,λ), the Lagrangian L can be minimized coordinate-wise with respect to v by considering the expressions $\beta|v_k|^0 + \frac{\Lambda_k}{2}|x_k - v_k|^2 - \lambda_k v_k$ to obtain

$$v_k = \Phi(x,\lambda)_k = \begin{cases} \dfrac{\lambda_k + \Lambda_k\, x_k}{\Lambda_k} & \text{if } |\lambda_k + \Lambda_k\, x_k|^2 > 2\Lambda_k\beta \\ 0 & \text{otherwise.} \end{cases}$$

Given (v,λ), L is minimized at x that satisfies

$$A^*(Ax-b) + \alpha P x + \Lambda(x-v) + \lambda = 0.$$

where Λ is diagonal operator with entries Λ_k. Thus, the augmented Lagrangian method [IK] uses the update:

$$
\begin{aligned}
&A^*(Ax^{n+1}-b)+\alpha Px^{n+1}+\Lambda(x^{n+1}-v^n)+\lambda^n=0\\
&v^{n+1}=\Phi(x^{n+1},\lambda^n)\\
&\lambda^{n+1}=\lambda^n+\Lambda(x^{n+1}-v^{n+1}).
\end{aligned}
\tag{5.1}
$$

If it converges, i.e. $x^n,\ v^n\to x$ and $\lambda^n\to\lambda$, then

$$
\begin{cases}
A^*(Ax-b)+\alpha\,Px+\lambda=0\\
\lambda_k=0,\quad \text{if } |\lambda_k+\Lambda_k x_k|^2>2\beta\,\Lambda_k,\\
x_k=0,\quad \text{if } |\lambda_k+\Lambda_k x_k|^2\le 2\beta\,\Lambda_k.
\end{cases}
\tag{5.2}
$$

That is, (x,λ) satisfies the necessary optimality condition (2.21) with A replaced by $A+\alpha P^{\frac12}$.

Let us further observe that in the inactive case $\lambda_k=0$ and $|x_k|>\sqrt{\frac{2\beta}{\Lambda_k}}$ and in the active case $x_k=0$ and $|\lambda_k|\le\sqrt{2\beta\Lambda_k}$. Thus on the inactive set only the λ component is 0, the x component is different from 0, on the active the x component is 0 and the λ component may or may not be 0.

Motivated by the augmented Lagrangian formulation we obtain a primal-dual active-set method as follows.

Primal-Dual Active Set Method

1. Initialize: $\lambda^0=0$ and determine x^0 by $A^*(Ax^0-b)+\alpha\,Px^0=0$. Set $n=0$
2. Solve for (x^{n+1},λ^{n+1});

$$
A^*(Ax^{n+1}-b)+\alpha\,Px^{n+1}+\lambda^{n+1}=0,
\tag{5.3}
$$

where

$$
\begin{aligned}
&\lambda_k^{n+1}=0,\qquad \text{if } k\in\{k:|\lambda_k^n+\Lambda_k x_k^n|^2>2\beta\,\Lambda_k\}\\
&x_k^{n+1}=0,\qquad \text{if } k\in\{k:|\lambda_k^n+\Lambda_k x_k^n|^2\le 2\beta\,\Lambda_k\}.
\end{aligned}
\tag{5.4}
$$

3. Converged, or set $k=k+1$ and return to Step 2.

Note that if the active set method converges, then the converged pair (x,λ) satisfies the necessary optimality (5.2). Due to good numerical experience we shall analyze its convergence. First sufficient conditions for uniqueness of solutions to (5.2) will be given and the following remarks are made.

Remark 5.1. Let us point out that (5.3) can be solved efficiently by first determining the solution on the inactive set by solving

$$
\mathcal{R}(A^*A+\alpha P)\mathcal{R}^*(\mathcal{R}x^{n+1})=\mathcal{R}(A^*b),
$$

where $\mathcal{R}$ denotes the restriction to the currently inactive set $\{k:|\lambda_k^n+\Lambda_k x_k^n|^2>2\beta\,\Lambda_k\}$ and then assigning the value for λ^{n+1} according to (5.3). In computations the matrix representation of $\mathcal{R}(A^*A+\alpha P)\mathcal{R}^*$ is simply obtained from $A^*A+\alpha P$ by forming the block sub-matrix corresponding to inactive rows and columns.

Remark 5.2. Since $(x^{n+1}, \lambda^{n+1}) = 0$ for all n, we have

$$((A^*A + \alpha\, P)x^{n+1}, x^{n+1}) \le (Ax^{n+1}, b)$$

and thus $\{|Ax^n|^2 + \alpha\,(Px^n, x^n)\}_{n=1}^{\infty}$ is bounded. Moreover from (5.3) we obtain

$$\begin{aligned} 0 &= (A^*(Ax^{n+1} - b) + Px^{n+1} + \lambda^{n+1}, x^{n+1} - x^n) \\ &= \frac{1}{2}(|Ax^{n+1} - b|^2 + \alpha\,(x^{n+1}, Px^{n+1})) - \frac{1}{2}(|Ax^n - b|^2 + \alpha\,(x_n, Px^n)) \\ &+ \frac{1}{2}\,(|A(x^{n+1} - x^n)|^2 + \alpha\,(x^{n+1} - x^n, P(x^{n+1} - x^n)) - (\lambda^{n+1}, x^n). \end{aligned}$$

The term (λ^{n+1}, x^n) relates to the switching between active and inactive set. Its value must be controlled to obtain convergence results.

5.1 Uniqueness

For any pair (x, λ) we define

$$\mathcal{I}(x,\lambda) = \{k : |\lambda_k + \Lambda_k x_k|^2 > 2\beta\Lambda_k\} \text{ and } \mathcal{A}(x,\lambda) = \{k : |\lambda_k + \Lambda_k x_k|^2 \le 2\beta\Lambda_k\},$$

and we set

$$Q = A^*A + P.$$

The following diagonal dominance condition will be used:

$$\|\Lambda^{-\frac{1}{2}}(Q - \Lambda)\Lambda^{-\frac{1}{2}}\|_\infty \le \rho \ \text{ for some } \rho \in (0,1). \tag{5.5}$$

Theorem 5.1. *(Uniqueness) Assume that* (5.5) *holds and that* $\delta > \frac{2\rho}{1-\rho}$. *Then there exists at most one solution to* (5.2) *satisfying*

$$\inf_{\mathcal{I}(x,\lambda)} |\Lambda^{-\frac{1}{2}}(\lambda + \Lambda x)| \ge (1+\delta)\sqrt{2\beta}. \tag{5.6}$$

An analogous statement holds with (5.6) *replaced by* $\sup_{\mathcal{A}(x,\lambda)} |\Lambda^{-\frac{1}{2}}(\lambda + \Lambda x)| \le (1-\delta)\sqrt{2\beta}$.

Above $\min_{\mathcal{I}(x,\lambda)} |\Lambda^{-\frac{1}{2}}(\lambda + \Lambda x)|$ stands for $\min_{k\in\mathcal{I}(x,\lambda)} |\Lambda_k^{-\frac{1}{2}}(\lambda_k + \Lambda x_k)|$.

Proof. Assume that there two pairs (x, λ) and $(\hat{x}, \hat{\lambda})$ satisfying (5.2) and (5.6). Then we have

$$Q(x - \hat{x}) + \lambda - \hat{\lambda} = 0.$$

and therefore

$$\Lambda^{\frac{1}{2}}x + \Lambda^{-\frac{1}{2}}\lambda - (\Lambda^{\frac{1}{2}}\hat{x} + \Lambda^{-\frac{1}{2}}\hat{\lambda}) = \Lambda^{-\frac{1}{2}}(\Lambda - Q)\Lambda^{-\frac{1}{2}}\,\Lambda^{\frac{1}{2}}(x - \hat{x}). \tag{5.7}$$

First consider the case that $x_k \ne 0$ if and only if $\hat{x}_k \ne 0$. Then, due to the fact that $x_k \ne 0$ implies that $\lambda_k = 0$, diagonal dominance implies that $x = \hat{x}$ and consequently $\lambda = \hat{\lambda}$. If there exists j such that $sign|x_j| \ne sign|\hat{x}_j|$, then without loss of generality we may assume that $x_j \ne 0$ and $\hat{x}_j = 0$.

As mentioned below (5.2) we have

$$|(\Lambda^{-\frac{1}{2}}\lambda)_k| \le \sqrt{2\beta} \ \text{ if } \ x_k = 0, \qquad \text{and } |(\Lambda^{-\frac{1}{2}}\hat{\lambda})_k| \le \sqrt{2\beta} \ \text{ if } \ \hat{x}_k = 0. \tag{5.8}$$

By (5.7) we find

$$\Lambda^{\frac{1}{2}}(x-\hat{x}) = \Lambda^{-\frac{1}{2}}(\Lambda - Q)\Lambda^{-\frac{1}{2}}\,\Lambda^{\frac{1}{2}}(x-\hat{x}) - \Lambda^{-\frac{1}{2}}(\lambda-\hat{\lambda}).$$

Due to (5.6) and (5.8) we obtain

$$|\Lambda^{\frac{1}{2}}(x-\hat{x})|_\infty \le \rho\,|\Lambda^{\frac{1}{2}}(x-\hat{x})|_\infty + 2\sqrt{2\beta},$$

and hence

$$|\Lambda^{\frac{1}{2}}(x-\hat{x})|_\infty \le \frac{2\sqrt{2\beta}}{1-\rho}. \tag{5.9}$$

Again by (5.7) and by (5.9) we have for each k:

$$|\Lambda_k^{-\frac{1}{2}}(\lambda_k - \Lambda_k x_k)| - |\Lambda_k^{-\frac{1}{2}}(\hat{\lambda}_k - \Lambda_k\hat{x}_k)| \le |\Lambda_k^{-\frac{1}{2}}(\lambda_k - \hat{\lambda}_k + \Lambda_k(x_k-\hat{x}_k)|$$

$$|\Lambda^{-\frac{1}{2}}(\lambda - \hat{\lambda} + \Lambda(x-\hat{x}))|_\infty \le \rho|\Lambda^{\frac{1}{2}}(x-\hat{x})|_\infty \le \frac{2\rho\sqrt{2\beta}}{1-\rho},$$

and consequently for the j chosen above

$$|\Lambda_j^{-\frac{1}{2}}(\lambda_j - \Lambda_j x_j)| - |\Lambda_j^{-\frac{1}{2}}(\hat{\lambda}_j - \Lambda_j\hat{x}_j)| \le \frac{2\rho\sqrt{2\beta}}{1-\rho}.$$

The strict complementarity assumption (5.6) implies that

$$(1+\delta)\sqrt{2\beta} - \sqrt{2\beta} \le \frac{2\rho\sqrt{2\beta}}{1-\rho},$$

and hence $\delta \le \frac{2\rho}{1-\rho}$ which is a contraction to the assumption $\delta > \frac{2\rho}{1-\rho}$. – The case $\max_{\mathcal{A}(x,\lambda)} |\Lambda^{-\frac{1}{2}}(\lambda + \Lambda x)| \le (1-\delta)\sqrt{2\beta}$ can be treated analogously.

$\square$

5.2 Convergence: Diagonal dominant case

Here we give a sufficient condition for the convergence of the primal-dual active set method. We shall utilize a diagonal dominance condition and consider a solution to (5.2) which satisfies a strict complementary condition. As such it is unique according to Theorem 5.1. Recall that by Remark (5.2) there exists M such that the iterates are bounded, i.e. $|x^n|_{\ell^2} \le M$ for all n. We set $\tilde{M} = \|\Lambda^{\frac{1}{2}}\|_{\mathcal{L}(\ell^2,\ell^\infty)}\,M$.

Proposition 5.1. *Let $(\bar{x},\bar{\lambda})$ denote a solution to (5.2) which satisfies the strict complementarity condition*

$$\sup_{\mathcal{A}(\bar{x},\bar{\lambda})} |\Lambda^{-\frac{1}{2}}\bar{\lambda}| \le (1-\delta)\sqrt{2\beta} \quad \text{and} \quad \inf_{\mathcal{I}(\bar{x},\bar{\lambda})} |\Lambda^{\frac{1}{2}}\bar{x}| \ge (1+\delta)\sqrt{2\beta} \tag{5.10}$$

and suppose that (5.6) holds. If $0 < \delta < \frac{\rho}{1-\rho}(\frac{2\rho\tilde{M}}{\sqrt{2\beta}} + 1)$, then the sets

$$S^n = \{k \in \mathcal{I}(\bar{x},\bar{\lambda}) : \lambda_k^n = 0\} \quad \text{and} \quad T^n = \{k \in \mathcal{A}(\bar{x},\bar{\lambda}) : x_k^n = 0\}$$

are monotonically nondecreasing. If $S^n = S^{n+1}$ and $T^n = T^{n+1}$ for some n, then $(x^n,\lambda^n) = (\bar{x},\bar{\lambda})$.

The proof will not make use of the particular initialization of the algorithm. In particular, if (5.2) admits a solution satisfying the strict complementarity assumption (5.10) we have global convergence to this solution.

Proof of Proposition 5.2. For two consecutive iterates we have

$$Q(x^n - x^{n-1}) + \lambda^n - \lambda^{n-1} = 0,$$

and thus

$$\Lambda^{-\frac{1}{2}}(Q-\Lambda)\Lambda^{-\frac{1}{2}}\,\Lambda^{\frac{1}{2}}(x^n - x^{n-1}) + \Lambda^{-\frac{1}{2}}(\lambda^n + \Lambda x^n) - \Lambda^{-\frac{1}{2}}(\lambda^{n-1} + \Lambda x^{n-1}) = 0. \tag{5.11}$$

If $x_k^n = 0$, then either $\lambda_k^n = 0$, or $\lambda_k^n = 0$, in which case $|\lambda^{n-1} + \Lambda x^{n-1}|^2 \le 2\beta\Lambda_k$ and by (5.11)

$$\begin{aligned} |\Lambda^{-\frac{1}{2}}\lambda_k^n| &\le |\,[\Lambda^{-\frac{1}{2}}(Q-\Lambda)\Lambda^{-\frac{1}{2}}\,\Lambda^{\frac{1}{2}}(x^n - x^{n-1})]_k\,| + |\Lambda_k^{-\frac{1}{2}}(\lambda_k^{n-1} + \Lambda_k x_k^{n-1})| \\ &\le \rho|\Lambda^{\frac{1}{2}}(x^n - x^{n-1})|_\infty + \sqrt{2\beta} = 2\rho\tilde{M} + \sqrt{2\beta}, \end{aligned} \tag{5.12}$$

with $\tilde{M}$ as defined before the statement of the theorem. We also have

$$\lambda(x^n - \bar{x}) + \lambda^n - \bar{\lambda} = (\Lambda - Q)(x^n - \bar{x}),$$

and hence

$$\Lambda^{\frac{1}{2}}(x^n - \bar{x}) + \Lambda^{-\frac{1}{2}}(\lambda^n - \bar{\lambda}) = \Lambda^{-\frac{1}{2}}(\Lambda - Q)\Lambda^{-\frac{1}{2}}\,\Lambda^{\frac{1}{2}}(x^n - \bar{x}). \tag{5.13}$$

Considering separately the cases $x_k^n \neq 0, \bar{x}_k \neq 0$, and $x_k^n \neq 0, \bar{x}_k = 0$, and $x_k^n = 0, \bar{x}_k \neq 0$, by (5.6), we find for all k and any n, using(5.12) and $|\Lambda_k^{-\frac{1}{2}}\bar{\lambda}_k| \le \sqrt{2\beta}$, that

$$|\Lambda^{\frac{1}{2}}(x_k^n - \bar{x}_k)| \le \rho|\Lambda^{\frac{1}{2}}(x^n - \bar{x})|_\infty + 2\rho\tilde{M} + \sqrt{2\beta}.$$

As a consequence we have

$$|\Lambda^{\frac{1}{2}}(x^n - \bar{x})|_\infty \le \frac{2\rho\tilde{M} + \sqrt{2\beta}}{1-\rho}. \tag{5.14}$$

Considering (5.13) on the set $S^n = \{\lambda_k^n = \bar{\lambda}_k = 0\}$ we find

$$\sup_{S^n} |\Lambda^{\frac{1}{2}}(x_k^n - \bar{x}_k)| \le \rho|\Lambda^{\frac{1}{2}}(x^n - \bar{x})|_\infty \le \frac{\rho}{1-\rho}(2\rho\tilde{M} + \sqrt{2\beta}).$$

For $k \in S^n$ we have $|\Lambda^{\frac{1}{2}}\bar{x}_k| \ge (1+\delta)\sqrt{2\beta}$ and hence

$$\begin{aligned} |\Lambda_k^{-\frac{1}{2}}\lambda_k^n + \Lambda_k^{\frac{1}{2}}x_k^n| &\ge |\Lambda_k^{\frac{1}{2}}\bar{x}_k| - |\Lambda_k^{\frac{1}{2}}(x_k^n - \bar{x}_k)| \ge (1+\delta)\sqrt{2\beta} - \tfrac{\rho}{1-\rho}(2\rho\tilde{M} + \sqrt{2\beta}) \\ &= \sqrt{2\beta}\big(1 + \delta - \tfrac{\rho}{1-\rho}(2\rho\tilde{M} + \sqrt{2\beta})\big) > \sqrt{2\beta}, \end{aligned}$$

and hence $\lambda_k^n = 0$ and $k \in S^{n+1}$. For $k \in T^n$ we have by (5.13) and (5.14)

$$|\Lambda_k^{-\frac{1}{2}}(\lambda_k^n - \bar{\lambda}_k)| \le \rho|\Lambda^{\frac{1}{2}}(x^n - \bar{x})|_\infty \le \frac{\rho}{1-\rho}(2\rho\tilde{M} + \sqrt{2\beta})$$

and hence

$$
\begin{aligned}
|\Lambda_k^{-\frac{1}{2}}\lambda_k^n + \Lambda_k^{\frac{1}{2}}x_k^n| &= |\Lambda_k^{-\frac{1}{2}}\lambda_k^n| \le |\Lambda_k^{-\frac{1}{2}}(\lambda_k^n - \bar\lambda_k)| + |\Lambda_k^{-\frac{1}{2}}\bar\lambda_k| \\
&\le (1-\delta)\sqrt{2\beta} + \tfrac{\rho}{1-\rho}(2\rho\tilde M + \sqrt{2\beta}) = \sqrt{2\beta}\big(1 + \tfrac{\rho}{1-\rho}(2\rho\tilde M + \sqrt{2\beta}) - \delta\big) < \sqrt{2\beta},
\end{aligned}
$$

and hence $x_k^{n+1} = 0$ and $k \in T^{n+1}$.

Assume now that $S^n = S^{n+1} \subset \mathcal{I}(\bar x, \bar\lambda)$ and $T^n = T^{n+1} \subset \mathcal{A}(\bar x, \bar\lambda)$ and that

$$
S^n \cup T^n \subsetneq \mathcal{I}(\bar x, \bar\lambda) \cup \mathcal{A}(\bar x, \bar\lambda). \tag{5.15}
$$

Assume that there exists $k \in \mathcal{A}(\bar x, \bar\lambda) \setminus T^n$. Then

$$
\begin{aligned}
&x_k^{n+1} \neq 0, \quad x_k^n \neq 0, \qquad \bar x_k = 0, \\
&\lambda_k^{n+1} = 0, \quad \lambda_k^n = 0.
\end{aligned}
$$

The update rule of the algorithm and strict complementarity imply that

$$
|\Lambda_k^{\frac{1}{2}}x_k^n| > \sqrt{2\beta} \;\text{ and }\; |\Lambda_k^{-\frac{1}{2}}\bar\lambda_k| \le (1-\delta)\sqrt{2\beta}.
$$

From (5.13) and (5.14)

$$
|\Lambda_k^{\frac{1}{2}}(x_k^n - \bar x_k)| \le \frac{\rho}{1-\rho}(2\rho\tilde M + \sqrt{2\beta}) + (1-\delta)\sqrt{2\beta}
$$

and hence

$$
\sqrt{2\beta} < |\Lambda_k^{\frac{1}{2}}x_k^n| \le \frac{\rho}{1-\rho}(2\rho\tilde M + \sqrt{2\beta}) + (1-\delta)\sqrt{2\beta}.
$$

This implies that $\delta < \frac{\rho}{1-\rho}(\frac{2\rho\tilde M}{\sqrt{2\beta}} + 1)$ which is impossible by the choice of δ and thus $T^n = \mathcal{A}(\bar x, \bar\lambda)$.

Similarly, if there exists $k \in \mathcal{I}(\bar x, \bar\lambda) \setminus S^n$ then

$$
\begin{aligned}
&\lambda_k^{n+1} \neq 0, \quad \lambda_k^n \neq 0, \quad \bar\lambda_k = 0, \\
&x_k^{n+1} = 0, \quad x_k^n = 0, \quad \bar x_k \neq 0.
\end{aligned}
$$

As a consequence

$$
|\Lambda_k^{-\frac{1}{2}}\lambda_k^n| \le \sqrt{2\beta} \;\text{ and }\; |\Lambda_k^{\frac{1}{2}}\bar x_k| > (1+\delta)\sqrt{2\beta}.
$$

Again by (5.13) and (5.14)

$$
(1+\delta)\sqrt{2\beta} < |\Lambda_k^{\frac{1}{2}}(x_k^n - \bar x_k)| \le \frac{\rho}{1-\rho}(2\rho\tilde M + \sqrt{2\beta}) + \sqrt{2\beta},
$$

which implies that $\delta < \frac{\rho}{1-\rho}(\frac{2\rho\tilde M}{\sqrt{2\beta}} + 1)$. This is impossible by the choice of δ and thus $S^n = \mathcal{I}(\bar x, \bar\lambda)$. Once the active set structure is determined the unique solution is determined by (5.3).

□

In the finite dimensional case, we can use the fact that there are only finitely many combinations of active indices and we obtain the following corollary.

Corollary 5.1. *In the finite dimensional case under the assumptions of Proposition 5.1 the algorithm converges in finitely many steps.*

5.3 Convergence: M-operator case

Again we set $Q = A^*A + \alpha$. For $\mathcal{C}$ an arbitrary subset of the index set $\{1, \ldots, \infty\}$, let $P_\mathcal{C}$ be the projection of ℓ^2 onto $\ell^2_\mathcal{C} = \{x \in \ell^2 : x_i = 0 \text{ if } i \neq \mathcal{C}\}$. Further let $\mathcal{C}^c$ denote the complement of $\mathcal{C}$ and set

$$Q_\mathcal{C} = P_\mathcal{C} Q P_\mathcal{C}, \quad Q_{\mathcal{C}^c\mathcal{C}} = P_{\mathcal{C}^c} Q P_\mathcal{C}.$$

We assume that Q is an M-operator , i.e.

$$(e_i, Qe_j)_{\ell^2} \leq 0, \text{ for all } i \neq j, \quad (e_i, Qe_j)_{\ell^2} > 0, \text{ for all } i,$$

and

$$Q_\mathcal{C} : l^2_\mathcal{C} \to l^2_\mathcal{C} \text{ is continuously invertible, with } Q_\mathcal{C}^{-1} x \geq 0 \text{ for } 0 \leq x \in \ell^2_\mathcal{C}.$$

Note that by definition of Λ we have $(Q - \Lambda)_{ii} = 0$ for all i. We further require the property

$$(Q - \Lambda)_{ij} \leq 0 \text{ for all } i \neq j, \tag{5.16}$$

which is clearly satisfied if Q is an M-operator and P is a diagonal operator.

Proposition 5.2. *If Q is an M-operator , (5.16) holds, $A^*b \geq 0$ and initialization is carried out with $x^0 > 0$, $\lambda^0 = 0$, then $x^n \geq 0, \lambda^n \geq 0$ for all n and*

$$\mathcal{A}^n = \{k : |\lambda^n_k + \Lambda_k x^n_k|^2 \leq 2\beta\Lambda_k\}$$

is monotonically decreasing. If $\mathcal{A}^n = \mathcal{A}^{n+1}$ for some $n \geq 0$, then (x^{n+1}, λ^{n+1}) satisfy the necessary optimality condition.

Corollary 5.2. *In the finite dimensional case under the assumptions of Proposition 5.2 the algorithm converges in finitely many steps.*

Proof of Proposition 5.2. Let $\mathcal{I}^n$ denote the complement of $\mathcal{A}^n$.

(i) Note that $\lambda^1 = \lambda^0 = 0$ on $\mathcal{I}^0$ and $x^0 \geq x^1 = 0$ on $\mathcal{A}^0$. Hence

$$Q(x^1 - x^0) = -(\lambda^1 - \lambda^0) = 0 \text{ on } \mathcal{I}^0.$$

Since Q is an M-operator and $x^0 \geq x^1$ on $\mathcal{A}^0$ it follows that $x^1 \leq x^0$ on $\mathcal{I}^0$.

Since

$$Qx^1 = A^*b \geq 0 \text{ on } \mathcal{I}^0,$$

splitting $P_{\mathcal{I}^0}Q$ according to $Q_{\mathcal{I}^0} + Q_{\mathcal{I}^0,\mathcal{A}^0}$, the M-property of Q implies that $x^1 \geq 0$. Analogously

$$Qx^1 + \lambda^1 = A^*b \geq 0 \text{ on } \mathcal{A}^0$$

and the fact that $x^1 = 0$ on $\mathcal{A}^0, x^1 \geq 0$ on $\mathcal{I}^0$ and the M-property of Q imply that $\lambda^1 \geq 0$ on $\mathcal{A}^0$ and consequently $\lambda^1 \geq 0$.

Next observe that by (5.16)

$$\lambda^1 + \Lambda x^1 - (\lambda^0 + \Lambda x^0) = -(Q - \Lambda)(x^1 - x^0) \leq 0$$

and hence $A^0 \supseteq \mathcal{A}^1$ and, taking complements, $\mathcal{I}^0 \subseteq \mathcal{I}^1$.

(ii) Assume now that $x^n \geq 0, \lambda^n \geq 0$ and let $\mathcal{I}^n, \mathcal{A}^n$ be determined according to (5.4). Let

$$\hat{\mathcal{I}}^n = \{k \in \mathcal{I}^n : x^n_k \neq 0\},\ \hat{\mathcal{A}}^n = (\hat{\mathcal{I}}^n)^c.$$

Then

$$Q_{\hat{\mathcal{I}}^n} P_{\hat{\mathcal{I}}^n}(x^{n+1} - x^n) = -Q_{\hat{\mathcal{I}}^n \hat{\mathcal{A}}^n} P_{\hat{\mathcal{A}}^n}(x^{n+1} - x^n) - P_{\hat{\mathcal{I}}^n}(\lambda^{n+1} - \lambda^n) \leq 0 \tag{5.17}$$

where we use that due to complementarity of (x^n, λ^n) we have $P_{\hat{\mathcal{I}}^n}\lambda^n = 0$, $P_{\hat{\mathcal{I}}^n}\lambda^{n+1} = 0$, and $P_{\hat{\mathcal{A}}^n} x^{n+1} = 0$. By (5.17) therefore $x^{n+1} \leq x^n$, on $\hat{\mathcal{I}}^n$ and consequently $x^{n+1} \leq x^n$. As in (i) above we now argue that $x^{n+1} \geq 0$ and $\lambda^{n+1} \geq 0$. Finally

$$\lambda^{n+1} + \Lambda x^{n+1} - (\lambda^n + \Lambda x^n) = -(Q - \Lambda)(x^{n+1} - x^n) \leq 0$$

and hence $\mathcal{A}^{n+1} \subseteq \mathcal{A}^n$.

(iii) If $\mathcal{A}^{n+1} = \mathcal{A}^n$ then for all $i \in \mathcal{A}^{n+1} = \{k : |\lambda^{n+1}_k + \Lambda_k x^n_k|^2 \leq 2\beta\Lambda_k\}$ we have $x^{n+1}_k = 0$ and for $k \in \mathcal{I}^{n+1} = \mathcal{I}^n = \{k : |\lambda^{n+1}_k + \Lambda_k x^n_k|^2 > 2\beta\Lambda_k\}$ we find $\lambda^{n+1}_k = 0$. Moreover $Qx^{n+1} + \lambda^{n+1} = A^*b$ and thus (x^{n+1}, λ^{n+1}) satisfies the first order conditions.

□

6 Examples

In this section we discuss examples that demonstrate the efficiency of the primal-dual active set algorithm for the case $p = 0$. We stress that in each case the algorithm is terminated when two successive iterates coincide. Thus we obtain an exact solution for a finite-dimensional approximation to the necessary condition (5.2), rather than an approximate solution, which is obtained by most other algorithms which are used to solve optimization problems involving ℓ^0 terms. Note that each iteration requires to solve one linear system. Our numerical experiments show that total number of iterations is small.- A detailed comparison among different methods is planed further work.

6.1 Sparsity in a Control Problem

We consider the linear control system:

$$\frac{d}{dt} y(t) = \mathcal{A}y(t) + Bu(t), \quad y(0) = 0,$$

i.e.

$$y(T) = \int_0^T e^{\mathcal{A}(T-s)} Bu(s)\, ds.$$

where the linear closed operator $\mathcal{A}$ generates a C_0-semigroup $e^{\mathcal{A}t}$, $t \geq 0$ on the state space X. Specifically, we discuss the (normalized) one dimensional controlled heat equation for $y = y(t,x)$:

$$y_t = y_{xx} + b_1(x)\, v_1(t) + b_2(x) v_2(t), \quad x \in (0,1),$$

with homogeneous boundary condition $y(t,0) = y(t,1) = 0$, where the differential operator $\mathcal{A}y = y_{xx}$ is discretized in space by the fourth order finite difference approximation [LI] with $n = 49$

Table 1:

β	.001	.003	.005	.007	.01	.03	.05	.07	.1	.3	.5
no of iterates	1	3	16	7	4	29	7	4	2	2	2
$\|u\|_0$	98	95	79	72	68	22	13	9	7	2	1

interior spatial nodes (mesh-size $\Delta x = 1/50$). We utilize two time dependent controls $u = (v_1, v_2)$ with corresponding spatial control distributions b_i chosen as step functions:

$$b_1(x) = \chi_{(.3,.4)}, \quad b_2(x) = \chi_{(.6,.7)}.$$

The control problem consists in finding the control function u that steers the state $y(0) = 0$ to a neighborhood of the desired state y_d at the terminal time $T = 1$. We discretize the problem in time by the mid-point rule, i.e,

$$Au = \sum_{k=1}^{m} e^{\mathcal{A}(T-t_{k+1/2})} B u_k \, \Delta t, \tag{6.1}$$

where $u \in \mathbb{R}^{2m}$ is a discretized control vector with coordinates $u_k \in \mathbb{R}^2$ which represent the control values at the mid-point of the intervals (t_k, t_{k+1}). A uniform step-size $\Delta t = 1/50$, (m=50), is utilized. The solution of the control problem is based on the sparsity formulation (1.1) where b is the discretized target function chosen as the Gaussian distribution $y_d(x) = exp(-100(x-.7)^2)$ centered at $x = .7$. That is, we apply our proposed algorithm for the discretized optimal control problem in time and space where x from (2.8) is the discretized control vector $u \in \mathbb{R}^{2m}$ which is mapped by A to the discretized output y at time 1 by means of (6.1). Further b from (2.8) is the descretized desired state y_d with respect to the spatial grid Δx.

The primal-dual active set formulation (5.4) with $p = 0$ was tested, where the weight matrix P was chosen as the derivative norm, i.e.,

$$(u, Pu) = \sum_{k=1}^{m-1} \left|\frac{u_{k+1} - u_k}{\Delta t}\right|^2 \Delta t.$$

Since the second control distribution is well within the support of the desired state y_d we expect the authority of this control to be much stronger than that of the first one, which is a distance away from the target. Our tests were conducted by incrementally increasing β from $\beta = .001$ to .5. For the results of the table below we initialized by $u_0 = (A^*A + \alpha\, P)^{-1}A^*b$ for the smallest β value and with the solution of the smaller β value for the next larger one. Moreover $\lambda_0 = 0$ for all cases. Consistent with our expectation, the ℓ^0 norm increases with β.

The method globally converges to a solution for each β. If we modify the initialization and also choose $u = (A^*A + \alpha\, P)^{-1}A^*b$ and $\lambda = 0$ for $\beta = .03$ and $\beta = .1$ the algorithm requires 31 and 27 iterates respectively, and converged to the same solution. As we expected the sparsity increases much faster on the first control v_1. Also, we added 10 % noise to f and tested the method. It converges globally and the number of iterates is actually smaller for this example.

6.2 M-matrix Example

Here we report on computations corresponding to Section 5.3. We consider

$$\min_{x\in\mathbb{R}^{n^2}} \frac{1}{2}|Ax-b|_2^2+\beta\,|x|_0. \tag{6.2}$$

where A is the forward finite difference gradient

$$A=\begin{pmatrix} G_1 \\ G_2 \end{pmatrix},$$

with $G_1\in\mathbb{R}^{n(n+1)\times n^2}$, $G_2\in\mathbb{R}^{n(n+1)\times n^2}$ given by

$$G_1=I\otimes D, \qquad G_2=D\otimes I.$$

Here I the $n\times n$ identity matrix, $\otimes$ denotes the tensor product, and $D\in\mathbb{R}^{(n+1)\times n}$ is given by

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ & & & & \\ 0 & \dots & 0 & -1 & 1 \\ 0 & \dots & 0 & 0 & -1 \end{pmatrix}.$$

Then A^TA is an M matrix coinciding with the 5 - point star discretization on a uniform mesh on a square of the Laplacian with Dirichlet boundary conditions. Moreover (6.2) can be equivalently expressed as

$$\min_{x\in\mathbb{R}^{n^2}} \frac{1}{2}|Ax|_2^2-(x,f)+\beta\,|x|_0, \tag{6.3}$$

where $f=A^Tb$. If $\beta=0$ this is the discretized variational form of the elliptic equation

$$-\Delta y=f \text{ in } \Omega, \qquad y=0 \text{ on } \partial\Omega. \tag{6.4}$$

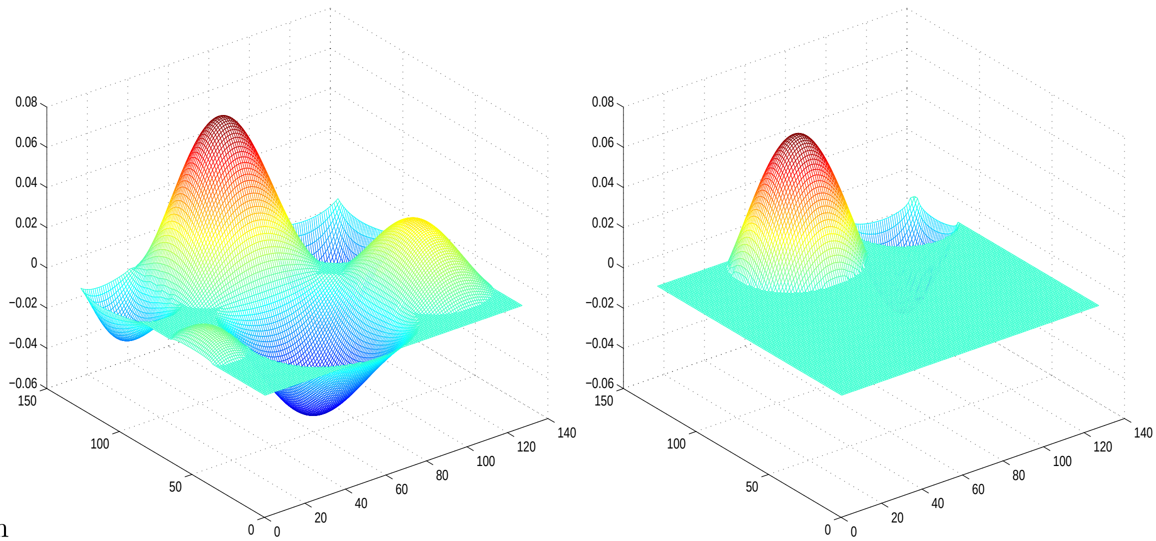
For $\beta>0$ the variational problem (6.3) gives a sparsity enhancing solution for this elliptic equation, i.e. the displacement y will be 0 where the forcing f is small.

In Table 2 we present the results of the primal-dual active set method for an $n=128$ mesh and f chosen as discretization of $f=10x_1 sin(5x_2)cos(7x_1)$. The matrix P is constructed with $\alpha=0$. The active sets convergence monotonically in spite of the fact that f does not have uniform sign. It stops with two consecutive iterates coinciding. In the third row of Table 2 we see that the cardinality of the active set increases with β. For $\beta=1$ the solution to (6.3) is 0. The last row of Table 2 exhibits the ℓ^2-norm of the difference between the 'free' solution x^* to $A^TAx=f$ and the sparsity enhancing solutions x_β to (6.3). For $\beta\le 10^{-9}$ the sparse solution x_β coincides with the free solution x^*.

If n is increased to $n=256$ then for $\beta=.1$ the algorithm requires 90 iterations to reach the solution for which $|x|_0=52543$.

Table 2:

β	.00001	.0001	.001	.01	.1	1
no of iterates	4	9	16	38	40	18
$\|x\|_0$	1634	16226	15874	13842	3437	0
$\|x^* - x_\beta\|_2$	$3.2 * 10^{-4}$	$2.02 * 10^{-3}$	$1.24 * 10^{-2}$	$1.86 * 10^{-1}$	2.05	3.04



m

Figure 1: Left: $\beta = .01$, Right $\beta = .1$

6.3 Sparsity and SVM

An important use of the sparsity optimization is for the Support Vector Machine (SVM) for the classification. We are given training data $\mathcal{D}$, a set of n points of the form

$$\mathcal{D} = \{(x_i, d_i) \mid x_i \in \mathbb{R}^m,\, d_i \in \{-1, 1\}\}_{i=1}^n$$

where the d_i is either 1 or -1. We want to find the maximum-margin hyperplane that divides the points having $d_i = 1$ from those having $d_i = -1$. Any hyperplane can be written as the set of points x satisfying

$$w \cdot x - \gamma = 0.$$

To this end a linear SVM determines the hyperplane (w, γ) by unconstrained minimization of the form [HIB]:

$$\min \quad \frac{1}{2}|\max(0, y)|^2 + \beta\,(|w|_p + \frac{1}{2}\gamma^2) \tag{6.5}$$

over $u = (w, \gamma)$, where $y = e - D(Aw - \gamma e)$ and $\max(0, y)$ measures the degree of misclassification. D is the $n \times n$ diagonal matrix with diagonal $D_{ii} = d_i$. That is, the SVM algorithm classifies data into two categories, $\mathcal{R}^-$ and $\mathcal{R}^+$, geometrically separated by the plane $\{x : x \cdot w - \gamma = 0\}$, and clustered around the two planes

$$\begin{aligned} \mathcal{R}^- &= \{x \in \mathbb{R}^m :\; x \cdot w - \gamma \le -1\} \\ \mathcal{R}^+ &= \{x \in \mathbb{R}^m :\; x \cdot w - \gamma \ge +1\}. \end{aligned} \tag{6.6}$$

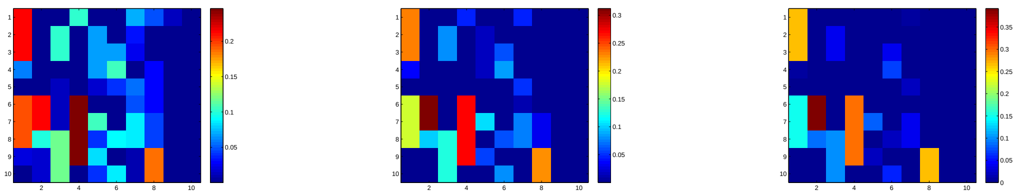


Figure 2: Heat map using $\ell^{1/2}$, $\ell^{1/5}$ and $\ell_{.0002}$ norms

The weight w can largely consist of insignificant coefficients, which one may desire to "weed" out, or effectively remove. We introduce sparsity by using ℓ^p which weeds out unnecessary weights and selects the responsible data. That is, if $w_j = 0$, the jth descriptor is not used for the classifier (6.6). We tested the sparsity formulation (6.5) for detecting neural activities by the Braingate technology [B]. The goal of the technology is to classify neural activity that correlates to specific imagined movements given data recorded from an electrode array implanted in the primary motor cortex of the human brain. The vector $x_i \in R^m$ describes the firing rate at time i, i.e. the j-th component of x_i corresponds to the firing rate measurement at the j-th electrode. Test example we used has 96 (m=96) neural channels data $\mathcal{D}$ that records neural firing rate data in time. Figure 2 shows a heat map for weights w for neural nodes with different p with a fixed $\beta = 1.e-3$. The sparsity optimization formulation provides a method to identify the active neurons that are most responsible for each movement. The smaller p the more sparsity is enhanced. We used the monotone convergent algorithm of Section 4 to solve the optimization problem (6.5), given $0 < p < 1$. The method is terminated within 5-6 interacts after attaining the desirable accuracy of solutions.

References

[AT] A. Auslender and M. Teboulle: Asymptotic Cones and Functions in Optimization and Variational Inequalities, Springer, Heidelberg, 2003.

[BBGT] C. Baiocchi, G. Buttazzo, F. Gastaldi, F. Tomarelli: General existence theorems for unilateral problems in continuum mechanics. Arch. Rational Mech. Anal. 100(1988), 149-189.

[BD] T. Blumensath and M. Davies: Iterative thresholding for sparse approximations, Journal of Fourier Analysis and Applications 14 (5) (2008) 629–654.

[B] C. Bouton: Decoding neural activity from an intracortical implant in humans with tetraplegia. In Biomedical Science & Engineering Conference, 2009. BSEC 2009. First Annual ORNL, Oak Ridge, TN, March 2009.

[BL] K. Bredies and D.A. Lorenz: Minimization of non-smooth, nonconvex functionals by iterative thresholing, preprint.

[CT] E. J. Candès and T. Tao: Decoding by linear programming. IEEE Trans. Inform. Theory, 51(12)(2005):4203-4215.

[CCK] E. Casas, C. Clason and K. Kunisch: Approximation of elliptic control problems in measure spaces with sparse solutions, SIAM Journal on Control and Optimization, to appear.

[CGWY] X. Chen, D. Ge, Z. Wang, and Y. Ye: Complexity of Unconstrained L2-Lp Minimization, Math. Program., Ser. A DOI 10.1007/s10107-012-0613-0

[CXY] X. Chen, F.Xu, and Y. Ye: Lower bound theory of nonzero entries in solutions of $\ell 2 - \ell_p$ minimization, SIAM J. Sci. Comput. 32 (2010), 2832-2852.

[C] I. Cioranescu: Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems, Mathematics and Its Applications, Volume 62, Kluwer Academic Publishers, 1990.

[D] D. L. Donoho: Compressed Sensing, IEEE Trans.Inform. Theory, 52(2006), 1289-1306.

[E1] M. Elad: Five Lectures on Sparse and Redundant Representations Modelling of Images, Book Chapters in "Mathematics in Image Processing" edited by Hongkai Zhao, AMS Publishing.

[E2] M. Elad: Sparse and Redundant Representations Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing, Springer-Verlag, 2010.

[FW] M. Fornasier and R. Ward: Iterative thresholing meets free-discontinuity problems, Found. Comput. Math., 10(2010), 527-567.

[G] M. Grasmair: Non-convex sparse regularisation. J. Math. Anal. Appl., 365(2010),19-28.

[HSW] R. Herzog, G. Stadler, and G. Wachsmuth: Directional Sparsity in Optimal Control of Partial Differential Equations, SIAM J. Control Optim., to appear.

[HIB] C. Humber, K. Ito, and C. Bouton: Nonsmooth Formulation of the Support Vector Machine for a Neural Decoding Problem, http://arxiv.org/abs/1012.0958v1.

[IK] K. Ito and K. Kunisch: *Lagrange Multiplier Approach to Variational Problems and Applications* SIAM Advances in Design and Control, 2008.

[LJ] R.M. Leahy and B.D. Jeffs: On the design of maximally sparse beamforming array IEEE Transaction on Antennas and Propagation, 39(1991), 1178-1187.

[LI] Z. Li and K. Ito: *The Immersed Interface Method*, SIAM Frontiers in Applied Mathematics, SIAM, Philadelphia, 2006.

[Lo] D.A. Lorenz: Convergence rates and source conditions for Tikhonov regularization with sparsity constraints. J. Inverse Ill-Posed Probl., 16(2008), 463-78.

[LW] Y. Liu and Y. Wu: Variable selection via a combination of the ℓ_0 and ℓ_1 penalties, J. Comput. Graph. Statist., 16(2007), 782-798.

[LZ] Z. Lu and Y. Zhang: Sparse approximation via penalty decomposition methods, preprint.

[N] M. Nikolova: Description of the minimizers of least squares regularized with ℓ_0- norm. Uniqueness of the global minimizer, SIAM J. Imaging Sciences, 6(2013), 904-937.

[RZ] R. Ramlau and C. Zarzer: On the optimization of a Tikhonov functional with non-convex sparsity constraints, ETNA Vol. 39: 476-507, 2012

[St] G. Stadler: Elliptic Optimal Control Problems with L^1-Control Cost and Applications for the Placement of Control Devices,Comp. Optim. Appls. 44(2), (2009),159–181.

[TLY] A.N. Tikhonov, A.S. Leonov, A.G. Yagola: *Nonlinear Ill-Posed Problems. Vol. 1, 2*, Applied Mathematics and Mathematical Computation, 14, Translated from the Russian, Chapman & Hall, 1998.

[WLMC] W. Wang, S. Lu, H. Mao and J. Cheng: Multi-parameter Tikhonov regularization with the ‘0 sparsity constraint. Inverse Problems, 29(2013), 065018 (18pp).

[WNF] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo: Sparse Reconstruction by Separable Approximation, IEEE Trans. Signal Process, 57 (2009), 2479-2493.

[Y] K. Yosida: *Functional Analysis*, Springer Verlag, Berlin, 1974.

[Z] C. A. Zarzer: On Tikhonov regularization with non-convex sparsity constraints. Inverse Problems, 25:025006, 2009, 13 pp.

[ZDL] Y. Zhang, B. Dong, and Z. Lu: Minimization for wavelet frame based image restoration, Math. Comp, to appear.