Bregman iterative methods for ℓ_1-minimization

Alessandro La Farciola

August 8, 2022

Contents

1 Introduction

In this seminar we introduce some efficient methods based on Bregman distance that can be used to solve ℓ_1-minimization problems. In particular, we, firstly, discuss about Bregman iteration in general introduced in [1]. Then, we analyse on one side Bregman iterative algorithms proposed by Yin, Osher, Goldfarb and Darbon in [3] for a specific problem, on the other the methods proposed by Goldstein and Osher in [4] with a special focus on the *Split Bregman method*. These techniques are used extensively in compressed sensing (CS) in order to solve efficiently ℓ_1- regularized problems.

1.1 General problem

Many important problems in engineering, computer science, and imaging science can be posed as a ℓ_1- regularized optimization problem and the general form for such problem is

$$\min_{u\in\mathbb{R}^n}\{\|u\|_1 \ : \ Au = f\} \qquad (Basis\ Pursuit) \tag{1.1}$$

where $A \in \mathbb{R}^{m\times n}$ and $f \in \mathbb{R}^m$. Moreover, the system $Au = f$ is typically underdetemined, i.e. taking $m < n$, $Au = f$ has more than one solution. CS is based on the idea that a signal can be reconstructed, i.e. through optimization, the sparsity of a signal can be used for recovering that signal from a small amount of data (for example magnetic resonance imaging). Indeed, let $\bar{u} \in \mathbb{R}^n$ denote a highly sparse signal, i.e.

$$k = \|\bar{u}\|_0 = |\{i : \bar{u}_i \neq 0\}| \ll n,$$

then one can encode $\bar{u}$ by a linear transform $f = A\bar{u} \in \mathbb{R}^m$ for $k < m \ll n$, and recover $\bar{u}$ from f by solving (1.1).

It is well known that instead of solving the constrained problem (1.1), is more preferable to solve the unconstrained problem

$$\min_{u\in\mathbb{R}^n} \mu \|u\|_1 + \frac{1}{2}\|Au - f\|_2^2 \tag{1.2}$$

where $\mu \in \mathbb{R}$. A method widely used to solve (1.2) is the *Proximal gradient algorithm* and in particular the *Iterative Shrinkage Thresholding Algorithm* (ISTA). In this seminar we analyse new methods to solve (1.2) based on *Bregman iterative regularization.*

1.2 Bregman iterative regularization

Bregman iteration is a concept that was introduced by Olsher et al. in [1] in the context of image processing and then it was applied to other optimization problems like compressed sensing. The authors were motivated by the following classical problem in image restoration: given a noisy image $f : \Omega \longrightarrow \mathbb{R}$, with $\Omega \subseteq \mathbb{R}^2$ open bounded, we want to obtain a decomposition

$$f = u + v$$

where u is the true signal and v the noise. One of the most popular method for approximating the solution was due to Rudin, Osher, Fatemi (ROF) and it was solving the problem

$$\min_{u \in BV(\Omega)} \left\{ \int_\Omega |\nabla u| + \lambda \, \|f - u\|_{L^2}^2 \right\} \tag{1.3}$$

for some $\lambda > 0$ and with $BV(\Omega)$ that denotes the space of bounded variation on Ω.

The main aim of [1] is to propose an iterative regularization procedure designed to improve ROF restoration. The idea was to iterate the problem (1.3), calling the first minimizer u_1 and use it to compute u_2, u_3, etc. To do that, the authors decided to use the Bregman distance based on a convex functional $J(\cdot)$ between two points, i.e.

$$D_J^p(u, v) := J(u) - J(v) - \langle p, u - v\rangle \tag{1.4}$$

where $p \in \partial J(v)$ is a subgradient in the subdifferential of J at the point v. The quantity $D_J^p(u, v)$ is not a distance in the usual sense, since $D_J^p(u, v) \neq D_J^p(v, u)$ in general and the triangle inequality does not hold. However, we can say that it measures the closeness between u and v in the sense that $D_J^p(u, v) \geq 0$, and $D_J^p(u, v) \geq D_J^p(w, v)$ for each w on the line segment connecting u and v. Taking $w = \lambda u + (1 - \lambda)v$ for a suitable $\lambda \in (0, 1)$ and due to the convexity of $J(\cdot)$, the last statement follows from

$$\begin{aligned} D_J^p(w, v) = J(w) - J(v) - \langle p, w - v\rangle &\leq \lambda J(u) - \lambda J(v) - \langle p, \lambda u - \lambda v\rangle \\ &\leq \lambda \left(J(u) - J(v) - \langle p, u - v\rangle\right) \\ &\leq D_J^p(u, v). \end{aligned}$$

Now, instead of the only (1.3), taking $J(u) = \int_\Omega |\nabla u|$, the Bregman iterative regularization method of Osher et al. is to solve convex problems

$$\begin{aligned} u^{k+1} &= \underset{u \in BV(\Omega)}{\arg\min} \left\{ D_J^{p^k}(u, u^k) + \lambda \, \|f - u\|_{L^2}^2 \right\} &(1.5) \\ &= \underset{u \in BV(\Omega)}{\arg\min} \left\{ J(u) - \langle p^k, u - u^k\rangle + \lambda \, \|f - u\|_{L^2}^2 \right\} &(1.6) \end{aligned}$$

for $k = 0, 1, \ldots$, starting with $u^0 = 0$ and $p^0 = 0$. Hence, one can note that for $k = 0$ it solves the original problem (1.3). Since $J(u)$ is not differentiable everywhere it could happen that its subdifferential contain more than one element. However, from the optimality of u^{k+1} in (1.6), it follows that $0 \in \partial J(u^{k+1}) - p^k + u^{k+1} - f$ and so they set

$$p^{k+1} := p^k + f - u^{k+1} \ \in \partial J(u^{k+1})$$

and it is possible to iterate the process.

Moreover, in [1] were proved some key results for the sequence $\{u^k\}$. Firstly, it is shown that $\left\|u^k - f\right\|_{L^2}$ converges to 0 monotonically; secondly, for λ sufficiently small, calling $\tilde{u}$ the truth noise-free image, until

$$\left\|u^k - f\right\|_{L^2} < \tau \left\|\tilde{u} - f\right\|_{L^2} \qquad \forall \tau > 1,$$

the sequence $\{u^k\}$ also gets monotonically closer to $\tilde{u}$ in terms of the Bregman distance. In fact, considering the following more general problem

$$\min_u J(u) + H(u)$$

with $J(\cdot)$ a convex function and $H(\cdot)$ a convex and differentiable function. The *Bregman iterative regularization* algorithm becomes

$$u^{k+1} = \arg\min_u \left\{D_J^{p^k}(u, u^k) + H(u)\right\} \tag{1.7}$$

$$p^{k+1} = p^k - \nabla H(u^{k+1}) \in \partial J(u^{k+1}), \tag{1.8}$$

and the authors showed the next significant theorem of convergence.

Theorem 1.1. *Let $J(\cdot)$ be a convex function, let $H(\cdot)$ be a convex and differentiable function and suppose that the solutions u^{k+1} of* (1.7) *exist. Therefore, the iterate sequence $\{u^k\}$ satisfies the following:*

1. *Monotonic decrease in H: $H(u^{k+1}) \leq H(u^{k+1}) + D_J^{p^k}(u^{k+1}, u^k) \leq H(u^k)$.*

2. *Convergence to the original in H with exact data: if $\tilde{u}$ minimizes $H(\cdot)$ and $J(\tilde{u}) < \infty$, then $H(u^k) \leq H(\tilde{u}) + J(\tilde{u})/k$.*

3. *Convergence to the original in D with noisy data: Let $H(\cdot) = H(\cdot, f)$ and suppose that $H(\tilde{u}, f) \leq \delta^2$ and $H(\tilde{u}, g) = 0$, where $f, g, \tilde{u}, \delta$ represent noisy data, noiseless data, perfect recovery, noise level, respectively. Then*

$$D_J^{p^{k+1}}(\tilde{u}, u^{k+1}) < D_J^{p^k}(\tilde{u}, u^k)$$

as long as $H(u^{k+1}, f) > \delta^2$.

2 Bregman iterative regularization by Yin, Osher, Goldfarb and Darbon

The principal aim of [3] is to show that the Bregman iterative method can be a simple but efficient algorithm for solving the basis pursuit problem (1.1) and the authors prove that in a finite number of iterations, u^k becomes a minimizer of $\|u\|_1$ among $\{u : Au = f\}$.

2.1 Formulations

Consider the general *Bregman iterative regularization* algorithm (1.7)-(1.8). We want to apply that algorithm to the problem (1.2) with $J(u) = \mu \|u\|_1$ and $H(u) = \frac{1}{2} \|Au - f\|_2^2$. Hence, we obtain

Algorithm 1:

$$u^0 = 0, \; p^0 = 0 \tag{2.9}$$
$$For \; k = 0, 1, \dots \; do \tag{2.10}$$
$$u^{k+1} = \arg\min_u D_J^{p^k}(u, u^k) + \frac{1}{2} \|Au - f\|_2^2 \tag{2.11}$$
$$p^{k+1} = p^k - A^T(Au^{k+1} - f) \; \in \partial J(u^{k+1}) \tag{2.12}$$

since $\nabla H(u^{k+1}) = A^T(Au^{k+1} - f)$. Now, consider

Algorithm 2:

$$u^0 = 0, \; f^0 = 0 \tag{2.13}$$
$$For \; k = 0, 1, \dots \; do \tag{2.14}$$
$$f^{k+1} = f + (f^k - Au^k) \tag{2.15}$$
$$u^{k+1} = \arg\min_u J(u) + \frac{1}{2} \left\|Au - f^{k+1}\right\|_2^2 \tag{2.16}$$

We are now ready to show that the previous algorithms are equivalent.

Theorem 2.1. *The Bregman iterative Algorithm 1* (2.9)-(2.12) *and the Bregman iterative Algorithm 2* (2.13)-(2.16) *are equivalent in the sense that* (2.11) *and* (2.16) *have the same objective functions (up to a constant) for any k.*

Proof. For any k, let u^k and $\bar{u}^k$ be the solutions of Algorithm 1 and 2, respectively. Firstly, at iteration $k = 0$ we have that $D_J^{p^0}(u, u^0) = D_J^0(u, 0) = J(u)$, while $f^1 = f$. Therefore, they have the same optimization problem

$$\min_u J(u) + \frac{1}{2} \|Au - f\|_2^2 .$$

We can note that at each iteration k the optimazation problem to solve may have more than one solution. Hence, in this case we don't assume that u^1 (Algorithm 1) is equal to $\bar{u}^1$ (Algorithm 2). While, we use the fact from [2] that the quantity $A^T(f - Au)$ is constant for all optimal solutions u [i]. So we have that $A^T(f - Au^1) = A^T(f - A\bar{u}^1)$. Now, due to (2.12), $p^0 = 0$, and $f = f^1$, we have

$$p^1 = p^0 - A^T(Au^1 - f) = A^T(f - Au^1) = A^T(f - A\bar{u}^1) = A^T(f^1 - A\bar{u}^1).$$

By induction it turns out that

$$p^k = A^T(f^k - A\bar{u}^k). \tag{2.17}$$

In fact, given $p^k = A^T(f^k - A\bar{u}^k)$ we have that

$$\begin{aligned} p^{k+1} = p^k - A^T(Au^{k+1} - f) &= p^k - A^T(A\bar{u}^{k+1} - f) \\ &= A^T(f^k - A\bar{u}^k) - A^T(A\bar{u}^{k+1} - f) \\ &= A^T\left(f + (f^k - A\bar{u}^k) - A\bar{u}^{k+1}\right) \\ &= A^T(f^{k+1} - A\bar{u}^{k+1}). \end{aligned}$$

We are now ready to prove that the optimization problems in (2.11) and (2.16) at iteration k are equivalent. It follows from

$$\begin{aligned} D_J^{p^k}(u, u^k) + \frac{1}{2}\|Au - f\|_2^2 &= J(u) - \langle p^k, u\rangle + \frac{1}{2}\|Au - f\|_2^2 + \underbrace{J(u^k) + \langle p^k, u^k\rangle}_{C_1} \\ &= J(u) - \langle f^k - A\bar{u}^k, Au\rangle + \frac{1}{2}\|Au - f\|_2^2 + C_1 \\ &= J(u) + \frac{1}{2}\left\|Au - (f + (f^k - A\bar{u}^k)\right\|_2^2 + C_2 \\ &= J(u) + \frac{1}{2}\left\|Au - f^{k+1}\right\|_2^2 + C_2 \end{aligned}$$

where $C_2 = C_1 - \frac{1}{2}\langle f^k - A\bar{u}^k, f\rangle + \left\|f^k - Au^k\right\|_2^2$ is constant respect to u. $\square$

2.2 Convergence results

In this section, following what the authors did in [3], we show that the Bregman iterative algorithm (2.9)-(2.12) (or, equivalently, (2.13)-(2.16)) generates a sequence $\{u^k\}$ that converges to u_{opt}, a solution of the basis pursuit problem (1.1), in a finite number of steps. Indeed, they prove that there exists a finite K such that every u^k for $k > K$ is a solution of the basis pursuit problem.

[i] In *Theorem 2.1* of [2] it is, firstly, shown that for any u^* optimal solution it turns out that Au^* is constant. Then, $A^T(f - Au^*)$ must be constant too.

The convergence analysis is divided into two theorems. The first proves that if u^k satisfies the condition $Au^k = f$, then it is a minimum for $J(\cdot) = \mu \|\cdot\|_1$; the second shows that such an u^k is obtained for a finite k.

Theorem 2.2. *Suppose that some u^k from* (2.11) *satisfies $Au^k = f$; then u^k is a solution of the basis pursuit problem* (1.1).

Proof. Due to nonnegative of the Bregman distance, for any u we have that

$$0 \le D_J^{p^k}(u, u^k) = J(u) - J(u^k) - \langle p^k, u - u^k\rangle \quad \Rightarrow \quad J(u^k) \le J(u) - \langle p^k, u - u^k\rangle.$$

Hence, from (2.17) and the hypothesis

$$\begin{aligned} J(u^k) &\le J(u) - \langle p^k, u - u^k\rangle \\ &= J(u) - \langle A^T(f^k - Au^k), u - u^k\rangle \\ &= J(u) - \langle f^k - Au^k, Au - Au^k\rangle \\ &= J(u) - \langle f^k - f, Au - f\rangle. \end{aligned}$$

Now, for any u such that $Au = f$ it turns out that $J(u^k) \le J(u)$ and so u^k is a solution of the basis pursuit problem. □

Theorem 2.3. *There exists a finite K such that any u^k with $k \ge K$ is a solution of the basis pursuit problem* (1.1).

Proof. Let (I_+^j, I_-^j, E^j) be a partition of the index set $\{1, 2, \dots, n\}$, and define

$$U^j := U(I_+^j, I_-^j, E^j) = \{u \ : \ u_i \ge 0, \ i \in I_+^j; \ u_i \le 0 \ i \in I_-^j; \ u_i = 0 \ i \in E^j\} \tag{2.18}$$

$$H^j := \min_u \left\{ \frac{1}{2} \|Au - f\|_2^2 \ : \ u \in U^j \right\}. \tag{2.19}$$

At iteration k let (I_+^j, I_-^j, E^j) be defined as follows:

$$I_+^k = \{i \ : \ p_i^k = \mu\}, \ I_-^k = \{i \ : \ p_i^k = -\mu\}, \ E^k = \{i \ : \ p_i^k \in (-\mu, \mu)\}. \tag{2.20}$$

Due to definition (2.18) and the fact that $p^k \in \partial J(u^k) = \partial(\mu\|u^k\|_1)$, it follows that $u^k \in U^k$. In fact, due to definition of subgradient we have that for any $u \in \mathbb{R}^n$

$$\mu \|u\|_1 \ge \mu \left\|u^k\right\|_1 + \langle p^k, u - u^k\rangle. \tag{2.21}$$

If we take $u = (u_1, \dots, u_n)$, $u^k = (u_1^k, \dots, u_n^k)$ and $p^k = (p_1^k, \dots, p_n^k)$, (2.21) becomes

$$\mu\left(|u_1| + \cdots + |u_n|\right) \ge \mu\left(|u_1^k| + \cdots + |u_n^k|\right) + p_1^k(u_1 - u_1^k) + \cdots + p_n^k(u_n - u_n^k).$$

Hence,

- if $p_i^k = \mu$ we can take $u = u^k - u_i^k e_i$, i.e. u_k with 0 in the i-th component, and we obtain that

$$\mu\left(|u_1| + \cdots + |u_{i-1}| + |u_{i+1}| + \cdots + |u_n|\right) \geq \mu\left(|u_1^k| + \cdots + |u_n^k|\right) - p_i^k u_i^k.$$

  Then,

$$0 \geq \mu|u_i^k| - \mu\ u_i^k \quad \Rightarrow \quad u_i^k \geq 0,$$

  in fact if it was < 0, therefore $|u_i^k| - u_i^k > 0$ and it is absurd.

- similarly, if $p_i^k = -\mu$, we obtain that

$$0 \geq \mu|u_i^k| + \mu\ u_i^k \quad \Rightarrow \quad u_i^k \leq 0.$$

- if $p^k \in (-\mu, \mu)$, it turns out that

$$0 \geq \mu|u_i^k| - p_i^k\ u_i^k,$$

  and from here it is easy to show that $u_i^k = 0$.

Now, we want to apply Theorem (1.1). Let $\tilde{u}$ be a minimizer of $H(u)$, i.e. $H(\tilde{u}) = \frac{1}{2}\left\|A\tilde{u} - f\right\|_2^2 = 0$. Due to the 2nd statement of Theorem (1.1), we see that for each j with $H^j > 0$ there exists a K_j such that u^k is not in U^j for each $k \geq K_j$. This follows since if we take j $s.t.$ $H^j = a > 0$, from Theorem (1.1)

$$H(u^k) \leq H(\tilde{u}) + \frac{J(\tilde{u})}{k} = \frac{\mu}{k}\left\|\tilde{u}\right\|_1,$$

then there exists a large K_j such that the last term is less than a for $k \geq K_j$, and so also $H(u^k) < a$, which means that $u^k \notin U^j$ for any $k \geq K_j$.Therefore, taking

$$K := \max_j\{K_j \ :\ H^j > 0\},$$

we have that $H(u^k) = 0$ for any $k \geq K$, which means that $Au^k = f$ for any $k \geq K$ and for the previous Theorem they solves the basis pursuit problem. $\square$

2.3 Equivalence to the augmented Lagrangian method

An interesting remark that the authors, initially, did not notice is the equivalence of the Bregman iterative method to the well-known augmented Lagrangian method (also known as the method of multipliers). Consider the optimization problem

$$\min_u s(u) \quad \textit{subject to } c_i(u) = 0, \quad i = 1, \ldots, m, \tag{2.22}$$

the augmented Lagrangian procedure consist of minimize the augmented Lagrangian function

$$L(u, \lambda^k, \nu) := s(u) + \sum_{i=1}^{m} \lambda_i^k c_i(u) + \frac{1}{2} \sum_{i=1}^{m} \nu_i c_i^2(u) \tag{2.23}$$

with respect to u at each iteration k, and uses the minimizer u^{k+1} to update the multipliers

$$\lambda_i^{k+1} = \lambda_i^k + \nu_i c_i(u^{k+1}). \tag{2.24}$$

Now, we can see that this method is equivalent to Algorithm 1 (2.9)-(2.12) letting

$$\begin{aligned} s(u) &= J(u) \\ c = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} &= Au - f \\ p^k &= -A^T \lambda^k \\ \nu_i &\equiv 1 \quad \forall i. \end{aligned}$$

Hence, it turns out that

$$\begin{aligned} L(u, \lambda^k, \nu) &= s(u) + \sum_{i=1}^{m} \lambda_i^k c_i(u) + \frac{1}{2} \sum_{i=1}^{m} \nu_i c_i^2(u) \\ &= J(u) + \langle \lambda^k, Au \rangle + \frac{1}{2} \|Au - f\|_2^2 + C_1 \\ &= J(u) - \langle p^k, u \rangle + \frac{1}{2} \|Au - f\|_2^2 + C_1 \\ &= D_J^{p^k}(u, u^k) + \frac{1}{2} \|Au - f\|_2^2 + C_2, \end{aligned}$$

which is the objective function at iteration k of Algorithm 1 (up to a constant). Moreover, $C_1,\ C_2$ are suitable constants in u and in the last equality we can add the term $-J(u^k)$ since it does not depend on u. Finally, due to (2.24) it follows also the condition (2.12), by multiplying by A^T. Therefore, taking $u^0 = 0$ and $\lambda^0 = 0$, the augmented Lagrangian method is equivalent to Algorithm 1. However, it is possible to note that Bregman iterative regularization is generally not equivalent to the augmented Langrangian method when the constraints are not linear.

3 Bregman iterative regularization by Goldstein and Osher

After reviewing the most significant results about Bregman iteration we have discussed above, the authors Goldstein and Osher introduce in [4] an innovative formulation for ℓ_1-regularized problems, that can broaden the previous results. In this way, it is possible to use an efficient algorithm in order to solve also some more general problems.

3.1 Constrained optimization via Bregman iteration

Firstly, they show a sort of generalization of Theorem (2.2). Instead of the basis pursuit problem, consider

$$\min_u E(u) \text{ } \textit{such that } Au = f, \tag{3.25}$$

where $E(\cdot)$ is a convex energy functional, A is a linear operator and $f \in \mathbb{R}^m$. As we have seen before, this problem is equivalent to the following unconstrained optimization problem

$$\min_u E(u) + \frac{\lambda}{2} \|Au - f\|_2^2 . \tag{3.26}$$

The conventional solution to this problem is to let $\lambda \to \infty$. Now, we apply the Bregman iteration methods (Algorithm 1) and we obtain the following procedure:

$$u^{k+1} = \arg\min_u D_E^{p^k}(u, u^k) + \frac{\lambda}{2} \|Au - f\|_2^2 \tag{3.27}$$

$$p^{k+1} = p^k - \lambda A^T(Au^{k+1} - f). \tag{3.28}$$

But, as we know, it is equivalent to Algorithm 2:

$$f^{k+1} = f + f^k - Au^k \tag{3.29}$$

$$u^{k+1} = \arg\min_u E(u) + \frac{\lambda}{2} \left\|Au - f^{k+1}\right\|_2^2 \tag{3.30}$$

We are now ready to prove that a solution $u^\star$ of $Au = f$ obtained through to (3.29)-(3.30) is a solution of the original constrained problem (3.25). Instead of the previous article [3] where it was assumed a specific form for the objective function, here the authors broaden that result.

Theorem 3.1. *Let $E : \mathbb{R}^n \longrightarrow \mathbb{R}$ be convex, let $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear operator and consider the algorithm* (3.29)-(3.30)*. Suppose that some iterate, $u^\star$, satisfies $Au^\star = f$. Therefore, $u^\star$ is a solution to the original constrained problem* (3.25)*.*

Proof. Let $u^\star$ and $b^\star$ such that $Au^\star = f$ and

$$u^\star = \arg\min_u E(u) + \frac{\lambda}{2} \|Au - f^\star\|_2^2 \tag{3.31}$$

Now, let $\hat{u}$ be a true solution of (3.25). Then, $Au^\star = b = A\hat{u}$ and it turns out that

$$\|Au^\star - f^\star\|_2^2 = \|A\hat{u} - f^\star\|_2^2 . \tag{3.32}$$

Due to (3.31), we have that

$$E(u^\star) + \frac{\lambda}{2}\|Au^\star - f^\star\|_2^2 \leq E(\hat{u}) + \frac{\lambda}{2}\|A\hat{u} - f^\star\|_2^2 . \tag{3.33}$$

Finally, putting together (3.32) and (3.33), we obtain that

$$E(u^\star) \leq E(\hat{u})$$

and since $\hat{u}$ solves the original optimization problem, the opposite inequality holds too and so $u^\star$ solves (3.25). □

3.2 Split Bregman method

In this section, following the authors Goldstein and Osher, we want to introduce the application of the Bregman method to the following general ℓ_1-optimization problem

$$\min_u \|\Phi(u)\|_{L^1} + H(u) \tag{3.34}$$

where both $\Phi(\cdot)$ and $H(\cdot)$ are convex functionals. We also assume $\Phi(\cdot)$ to be differentiable. Now, instead of considering the previous problem, we consider the equivalent

$$\min_{u,d} \|d\|_1 + H(u) \quad \textit{such that } d = \Phi(u). \tag{3.35}$$

As we did before, we convert it into an unconstrained problem:

$$\min_{u,d} \|d\|_1 + H(u) + \frac{\lambda}{2}\|d - \Phi(u)\|_2^2 . \tag{3.36}$$

Now, taking $E(u,d) = \|d\|_1 + H(u)$ and $A(u,d) := d - \Phi(u)$, it turns out that (3.36) is an application of (3.26), with $f = 0$. Applying the Bregman iteration (3.27)-(3.28), we obtain

$$\begin{aligned}
(u^{k+1}, d^{k+1}) &= \arg\min_{u,d} D_E^{p^k}(u, u^k, d, d^k) + \frac{\lambda}{2}\|d - \Phi(u)\|_2^2 && (3.37)\\
&= \arg\min_{u,d} E(u,d) - \langle p_u^k, u - u^k\rangle - \langle p_d^k, d - d^k\rangle + \frac{\lambda}{2}\|d - \Phi(u)\|_2^2 && (3.38)\\
p_u^{k+1} &= p_u^k - \lambda(\nabla\Phi)^T(\Phi(u^{k+1}) - d^{k+1}) && (3.39)\\
p_d^{k+1} &= p_d^k - \lambda(d^{k+1} - \Phi(u^{k+1})). && (3.40)
\end{aligned}$$

While, using the simplifaction (3.29)-(3.30), with $f = 0$, $E(u, d)$ and $A(u, d)$ as above, we have the following algorithm.

Split Bregman Iteration

$$f^{k+1} = f^k + (\Phi(u^k) - d^k). \tag{3.41}$$

$$(u^{k+1}, d^{k+1}) = \arg\min_{u,d} \|d\|_1 + H(u) + \frac{\lambda}{2} \left\| d - \Phi(u) - f^{k+1} \right\|_2^2. \tag{3.42}$$

Now, in order to implement the *Split Bregman Iteration* algorithm, we want to solve the problem

$$\min_{u,d} \|d\|_1 + H(u) + \frac{\lambda}{2} \left\| d - \Phi(u) - f^{k+1} \right\|_2^2. \tag{3.43}$$

Because of the way it is formulated, we can minimize, firstly, respect to u and, secondly, respect to d, separately. Hence, we have two steps:

$$\begin{aligned} \textit{Step } 1: \; u^{k+1} &= \arg\min_{u} H(u) + \frac{\lambda}{2} \left\| d^k - \Phi(u) - f^{k+1} \right\|_2^2 \\ \textit{Step } 2: \; d^{k+1} &= \arg\min_{d} \|d\|_1 + \frac{\lambda}{2} \left\| d - \Phi(u^{k+1}) - f^{k+1} \right\|_2^2. \end{aligned}$$

We note that the optimization problem that we have to solve for u^k (Step 1) is differentiable, and so there exist already several efficient techniques to solve it. The particular method to use depends on the nature of H and Φ. While, to solve step 2 it is possible to use the *Proximal gradient algorithm* and in particular the *Iterative Shrinkage Thresholding Algorithm* (ISTA) in order to explicitly compute the optimal value of d.

Finally, it is easy to show that any fixed point of the *Split Bregman Iteration* algorithm is a minimizer of the original constrained problem (3.35). Let $(u^\star, f^\star)$ be a fixed point of (3.41)-(3.42), then it satisfies

$$f^\star = f^\star + \Phi(u^\star) - d^\star$$

which implies that

$$d^\star = \Phi(u^\star).$$

Therefore,

$$A(u^\star, d^\star) = d^\star - \Phi(u^\star) = 0 = f$$

and thanks to Theorem (3.1) we have taht $(u^\star, f^\star)$ is a solution of (3.35).

4 Conclusion

In this seminar we presented some efficient methods of compressed sensing for solving ℓ_1-minimization problems. In particular, we, firstly, considered the general *Bregman iterative regularization* introduced in [1]. Here, Osher et al., motivated by classical problem in image restoration, proposed an iterative regularization procedure based on the Bregman distance. Then, we moved on to analyse two different methods found in [3] and [4] based on the previous work about Bregman iterative technique. In [3] the authors wanted to apply the Bregman iterative procedure to a specific ℓ_1-optimization problem, the Basis Pursuit problem (1.1), and we discussed two equivalent algorithms with some significant convergence results. We also noted the equivalence of this method to the augmented Lagrangian method. Finally, we investigated the procedures proposed by Goldstein and Osher in [4]. Here, we found a sort of generalization of the basis pursuit problem faced in [3] with a similar convergence result of the Bregman iterative algorithm applied to this problem. Moreover, they introduced a new riformulation of the general ℓ_1-optimization problem with a solution through the so-called *Split Bregman method*. All these techniques are used extensively in many important problems of engineering, computer science, imaging science in order to solve efficiently ℓ_1-regularized problems.

References

[1] S. Osher, M. Burger, D. Goldfarb, J. Xu, W. Yin, *An iterative regularization method for total variation-based image restoration*, Multiscale Model. Simul., 4 (2005), pp. 460–489.

[2] E. Hale, W. Yin, Y. Zhang, *A Fixed-Point Continuation Method for ℓ1-Regularization with Application to Compressed Sensing*, CAAM Technical report TR07-07, Rice University, Houston, TX, 2007.

[3] W. Yin, S. Osher, D. Goldfarb, J. Darbon, *Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing*, SIAM J. Imaging Sci., 1 (2008), pp. 143–168.

[4] T. Goldstein, S. Osher, *The Split Bregman Method for L1-Regularized Problems*, SIAM J. Imaging Sci., 2 (2009), pp. 323–343.