SIAM J. Imaging Sciences
Vol. 2, No. 2, pp. 323–343

The Split Bregman Method for L1-Regularized Problems*

Tom Goldstein† and Stanley Osher†

Abstract. The class of L1-regularized optimization problems has received much attention recently because of the introduction of "compressed sensing," which allows images and signals to be reconstructed from small amounts of data. Despite this recent attention, many L1-regularized problems still remain difficult to solve, or require techniques that are very problem-specific. In this paper, we show that Bregman iteration can be used to solve a wide variety of constrained optimization problems. Using this technique, we propose a "split Bregman" method, which can solve a very broad class of L1-regularized problems. We apply this technique to the Rudin–Osher–Fatemi functional for image denoising and to a compressed sensing problem that arises in magnetic resonance imaging.

Key words. constrained optimization, L1-regularization, compressed sensing, total variation denoising

AMS subject classification. 65K05

Bregman distance/divergence associated with a convex function E at the point v:

$$D_E^p(u, v) = E(u) - E(v) - \langle p, u - v \rangle,$$

Metric on a functional …

where p is in the sub-gradient of E at v.

Again, consider two convex energy functionals, E and H, defined over R^n with $\boxed{\min_{u \in R^n} H(u)}$ $= 0$. The associated unconstrained minimization problem is

$$\min_u \boxed{E(u)} + \lambda H(u). \tag{2.1}$$

… Used to reformulate a 2-functional variational problem for a single variable:

We can modify this problem by iteratively solving

$$u^{k+1} = \min_u \boxed{D_E^p(u, u^k)} + \lambda H(u) \tag{2.2}$$

"Gradient" on E

$$= \min_u \boxed{E(u) - \langle p^k, u - u^k \rangle} + \lambda H(u), \tag{2.3}$$

$$p^{k+1} = p^k - \nabla H(u^{k+1}).$$

"Gradient" on H

Goldstein-Osher: Special case

$$\min_u E(u) \text{ such that } Au = b$$

$$\min_u \boxed{E(u)} + \boxed{\frac{\lambda}{2}\|Au - b\|_2^2.}$$

$$u^{k+1} = \min_u D_E^p(u, u^k) + \frac{\lambda}{2}\|Au - b\|_2^2 \tag{2.6}$$

$$= \min_u E(u) - \langle p^k, u - u^k \rangle + \frac{\lambda}{2}\|Au - b\|_2^2, \tag{2.7}$$

$$p^{k+1} = p^k - \boxed{\lambda A^T(Au^{k+1} - b)}. \tag{2.8}$$

simplification

Simplification If A is linear:

$$u^{k+1} = \min_u \boxed{E(u)} + \frac{\lambda}{2}\|Au - \boxed{b^k\|_2^2,} \tag{2.9}$$

$$\boxed{b^{k+1}} = b^k + b - Au^k. \tag{2.10}$$

➔ we simply add the **error** in the constraint back to the right-hand side

Split Bregman—A better formulation for L1-regularized problems

Decouple the L1 and L2 parts:

$$\min_{u,d} |d| + H(u) \text{ such that } d = \Phi(u).$$

2 variables, inter-related

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

$$(u^{k+1}, d^{k+1}) = \min_{u,d} D_E^p(u, u^k, d, d^k) + \frac{\lambda}{2}\|d - \Phi(u)\|_2^2 \tag{3.3}$$

$$= \min_{u,d} E(u,d) - \langle p_u^k, u - u^k\rangle - \langle p_d^k, d - d^k\rangle + \frac{\lambda}{2}\|d - \Phi(u)\|_2^2, \tag{3.4}$$

$$p_u^{k+1} = p_u^k - \lambda(\nabla\Phi)^T(\Phi u^{k+1} - d^{k+1}), \tag{3.5}$$

$$p_d^{k+1} = p_d^k - \lambda(d^{k+1} - \Phi u^{k+1}). \tag{3.6}$$

simplification

Split Bregman Iteration

$$(u^{k+1}, d^{k+1}) = \boxed{\min_{u,d} |d| + H(u)} + \frac{\lambda}{2}\|d - \Phi(u) - \boxed{b^k\|_2^2,} \tag{3.7}$$

$$b^{k+1} = b^k + (\Phi(u^{k+1}) - d^{k+1}). \tag{3.8}$$

Split Bregman Iteration

$$(u^{k+1}, d^{k+1}) = \boxed{\min_{u,d} |d| + H(u)} + \frac{\lambda}{2}\|d - \Phi(u) - b^k\|_2^2, \tag{3.7}$$

$$b^{k+1} = b^k + (\Phi(u^{k+1}) - d^{k+1}). \tag{3.8}$$

iteratively minimizing with respect to u and d

Step 1: $u^{k+1} = \boxed{\min_{u} H(u)} + \frac{\lambda}{2}\|d^k - \Phi(u) - b^k\|_2^2,$ solve for u^k is now differentiable.

Step 2: $d^{k+1} = \boxed{\min_{d} |d|} + \frac{\lambda}{2}\|d - \Phi(u^{k+1}) - b^k\|_2^2.$ no coupling between elements of d.

explicitly compute the optimal value of d using shrinkage operators. We simply compute

$$d_j^{k+1} = shrink(\Phi(u)_j + b_j^k, 1/\lambda),$$

where

$$shrink(x, \gamma) = \frac{x}{|x|} * \max(|x| - \gamma, 0). \tag{3.10}$$

Shrinkage operator:

- ➔ **L1 sparse approximation**
- ➔ **Keep strongest “coeff”**