TITLE

AUTHORS

Abstract. Abstract

Key words. key words

AMS subject classifications. AMS ref numbers

1. Introduction.

1.1. Contributions.

1.2. Previous studies / State-of-the-art.
[4]: does not handle box constraint with regularization. In contrast, we can do it using primal-dual optimization.

1.3. Organization of the paper.

1.4. Notation.

2. .

3. .

$$\begin{aligned} \min_{\boldsymbol{\beta},\boldsymbol{\tau}} \quad & \frac{1}{2}\sum_{n=1}^{N}\left(F_n(\boldsymbol{\beta})\boldsymbol{\tau} - \tilde{w}_n\right)^2 + \frac{1}{2}\lambda_\alpha \left\|\boldsymbol{\tau}\right\|_2^2 + \lambda_\beta \mathcal{R}(\boldsymbol{\beta}) \\ \text{s.t.} \quad & \boldsymbol{\tau} \in \mathbb{R}_+^M \,,\, \boldsymbol{\beta} \in]0,1[^M \,. \end{aligned} \tag{3.1}$$

where $\mathcal{R}$ is a regularization that is proper, convex and lower-semicontinuous. For instance, $\mathcal{R}$ can be the total variation (TV) $\mathcal{R}(\mathbf{x}) = \left\|\mathbf{D}\mathbf{x}\right\|_1$ with $\mathbf{D}$ the discrete gradient matrix defined as the idendity matrix with coefficients -1 on its subdiagonal.

Equation (3.1) has the typical form of a constrained variable projection problem [2]. Indeed, it can be formulated as

$$\begin{aligned} \min_{\boldsymbol{\beta},\boldsymbol{\tau}} \quad & \frac{1}{2}\left\|\begin{pmatrix}\mathbf{F}(\boldsymbol{\beta}) \\ \sqrt{\lambda_\alpha}\,\mathrm{Id}\end{pmatrix}\boldsymbol{\tau} - \begin{pmatrix}\tilde{\mathbf{w}} \\ \mathbf{0}\end{pmatrix}\right\|_2^2 + \lambda_\beta \mathcal{R}(\boldsymbol{\beta}) \\ \text{s.t.} \quad & \boldsymbol{\tau} \in \mathbb{R}_+^M \,,\, \boldsymbol{\beta} \in]0,1[^M \,. \end{aligned} \tag{3.2}$$

Solving Problem 3.1 can be performed in a block-descent scheme. Minimization with respect to $\boldsymbol{\tau}$ consists in solving a non-negative least-squares problem. This problem can be solved using for instance trust region reflective algorithm described in [1] and implemented in SciPy library [5]. Minimization with respect to $\boldsymbol{\beta}$ can be performed with primal-dual methods [3]. Indeed, setting f as the indicator function of the set $[\epsilon, 1-\epsilon]$, g as the ℓ_1-norm multiplied by λ_β, and h as the least-squares term, Problem 3.1 can be written as an unconstrained problem of the form

$$\min_{\boldsymbol{\beta}\in\mathbb{R}^M} f(\boldsymbol{\beta}) + g(\mathbf{D}\boldsymbol{\beta}) + h(\boldsymbol{\beta})\,, \tag{3.3}$$

where f, g, and h are proper, convex, and lower-semicontinuous functions, h is gradient 1-Lipschitz, and D is a linear operator from $\mathbb{R}^M$ to $\mathbb{R}^M$. The dual problem of (3.3) is hence given by

$$\min_{\mathbf{v}\in\mathbb{R}^M} (f^* \star_{\inf} h^*)(\mathbf{L}^\top \mathbf{v}) + g^*(\mathbf{v})\,, \tag{3.4}$$

Algorithm 3.1 Implementation of RFBPD to solve (3.1) in $\boldsymbol{\beta}$

Input: Initial value of $\boldsymbol{\beta}_0$
Input: $(\tau, \sigma) \in]0, +\infty[^2$, and $\omega \in \mathbb{R}_+$.
Output: Estimate of $\boldsymbol{\beta}$.

```
1: Initialize i to 0.
2: Initialize v to 0.
3: repeat
4:    Primal update:
          p = β − τ∇g(β) − σ Dᵀ v
          p = Π_{[ε,1−ε]^M}(p)
5:    Dual update:
          q = (Id − Soft_{λ_β/σ})(v + S(2p − β))
6:    Inertial update:
          β = β + ω(p − β)
          v = v + ω(q − v)
7:    Increment i.
8: until stopping criterion is met
9: return β
```

where $\mathbf{v}$ is the dual variable, * denotes the Fenchel transform such that $f^*(\mathbf{v}) = \sup_{\mathbf{x}\in\mathbb{R}^M}(\langle \mathbf{x} \mid \mathbf{v}\rangle - f(\mathbf{x}))$, and $\star_{\inf}$ is the infinimal convolution defined by $(f \star_{\inf} g)(\mathbf{x}) = \inf_{\mathbf{y}\in\mathbb{R}^M} f(\mathbf{y}) + g(\mathbf{x}-\mathbf{y})$. Note that the infinimal convolution acts to the Fenchel transform as the classic convolution to the Fourier transform [3]. For instance, the two following equalities hold: $(f \star_{\inf} g)^* = f^* + g^*$ and $(f+g)^* = f^* \star_{\inf} g^*$. The main idea of primal-dual algorithms consists then to solve both the primal (3.3) and the dual (3.4) problems at the same time. A parallel to solving signal processing problems using time-frequency approach can be made. Exploiting information of both problems allows primal-dual techniques to yield computational advantages as well as to tackle more general problems.

In order to solve Problem (3.3) (and consequently (3.4)), we propose to use the Rescaled Primal Dual Forward-Backward (RFBPD) [3] whose iteration scheme is displayed in Equation (3.5)

$$
\begin{aligned}
\mathbf{p}_n &= \operatorname{prox}_{\rho f}(\boldsymbol{\beta}_n - \rho(\nabla h(\boldsymbol{\beta}_n) + \sigma \mathbf{D}^\top \mathbf{v}_n)) \\
\mathbf{q}_n &= (\mathrm{Id} - \operatorname{prox}_{\lambda g/\sigma})(\mathbf{v}_n + \mathbf{D}(2\mathbf{p}_n - \boldsymbol{\beta}_n)) \\
(\boldsymbol{\beta}_{n+1}, \mathbf{v}_{n+1}) &= (\boldsymbol{\beta}_n, \mathbf{v}_n) + \omega_n((\mathbf{p}_n, \mathbf{q}_n) - (\boldsymbol{\beta}_n, \mathbf{v}_n)),
\end{aligned}
\tag{3.5}
$$

where ω is a positive relaxation or inertial factor, and τ and σ are the two step-sizes. Noting that $\mathrm{Id} - \operatorname{prox}_{\lambda g/\sigma}$ is, up to a rescaling, equal to $\operatorname{prox}_{\sigma g^*}$, the iteration (3.5) closely looks like a Forward-Backward (FB) step performs on the primal, followed by a FB step performs on the dual.

The implementation of the RFBPD algorithm to solve (3.1) is given in Algorithm 3.1 where Soft_γ is the soft thresholding operator with positive parameter γ applied element-wise $\mathrm{Soft}_\gamma(x) = \mathrm{sgn}(x)\max(0, |x| - \gamma)$.

4. .

5. Conclusion.

Conclusion

Appendix A. Proof of blablabla.

REFERENCES

[1] T. F. Coleman and Y. Li, *An interior trust region approach for nonlinear minimization subject to bounds*, SIAM J. Optim., 6 (1996), pp. 418–445, https://doi.org/10.1137/0806023.

[2] G. Golub and V. Pereyra, *Separable nonlinear least squares: the variable projection method and its applications*, Inverse Problems, 19 (2003), pp. R1–R26, https://doi.org/10.1088/0266-5611/19/2/201.

[3] N. Komodakis and J.-C. Pesquet, *Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems*, IEEE Signal Process. Mag., 32 (2015), pp. 31–54, https://doi.org/10.1109/MSP.2014.2377273.

[4] T. van Leeuwen and A. Y. Aravkin, *Variable projection for nonsmooth problems*, SIAM J. Sci. Comput., 43 (2021), pp. S249–S268, https://doi.org/10.1137/20m1348650.

[5] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, Nat. Meth., 17 (2020), pp. 261–272, https://doi.org/10.1038/s41592-019-0686-2.