Solver for ℓ_1 plus ℓ_2 minimization

March 17, 2014

Consider

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \|x\|_1 + \frac{\beta}{2}\|x - w\|^2 \\ \text{subject to} & Ax = b\,. \end{array} \tag{1}$$

We might attempt to solve (1) with a method similar to the solver for the ℓ_1-ℓ_1 problem, i.e., when the second term of the objective is replaced by $\beta\|x - w\|_1$ (see folder `../../basisPursuitPlusL1`). However, such method, which is implemented in the file `../basisPursuitPlusL2.m` turns out to be unstable and does not converge always (see experiments in `../TestL2.m`). Here, we design a method based on the Barzilai-Borwein algorithm, namely on the method described in [1].

Since the objective of (1) is strictly convex, we can solve its dual problem instead:

$$\begin{aligned} &\underset{\lambda}{\text{maximize}}\;\; b^\top\lambda + \inf_x\Big[\|x\|_1 + \frac{\beta}{2}\|x - w\|^2 - \lambda^\top A x\Big] \\ \Longleftrightarrow\qquad &\underset{\lambda}{\text{minimize}}\;\; -b^\top\lambda + \underbrace{\sup_x\Big[\lambda^\top A x - \|x\|_1 - \frac{\beta}{2}\|x - w\|^2\Big]}_{=:\Psi(\lambda)}\,. \end{aligned} \tag{2}$$

Note that (2) is unconstrained, and its objective is convex, continuously differentiable, and its gradient is Lipschitz continuous. Therefore, we can apply any gradient-based method such as Nesterov's algorithm or Barzilai-Borwein. To compute the gradient of Φ and a point λ, we need first to find $x(\lambda)$ that minimizes the problem defining Ψ, that is,

$$\underset{x}{\text{minimize}}\;\; \|x\|_1 + v^\top x + \frac{\beta}{2}\|x\|^2\,,$$

where $v = -(\beta w + A^\top\lambda)$. This problem decomposes across components in x_i, whose optimality condition is

$$0 \in \partial|x_i| + v_i x_i + \frac{\beta}{2}x_i^2\,.$$

Let us then find the ith component:

- If $x_i > 0$, then

  $$0 = 1 + v_i + \beta x_i \qquad \Longleftrightarrow \qquad x_i = -\frac{v_i + 1}{\beta}\,.$$

  And this is positive whenever $v_i < -1$.

- If $x_i < 0$, then

  $$0 = -1 + v_i + \beta x_i \qquad \Longleftrightarrow \qquad x_i = -\frac{v_i - 1}{\beta}\,.$$

  And this is negative whenever $v_i > 1$.

- Finally, if $x_i = 0$, then $v_i \in [-1, 1]$, which is $|v_i| \le 1$.

The gradient of the objective of (2) at a point λ is $Ax(\lambda) - b$. It can be proven that this gradient is Lipschitz-continuous with constant $1/\beta$.

References

[1] E. Birgin, J. Martínez, and M. Raydan, "Nonmonotone Spectral Projected Gradient Methods on Convex Sets" *SIAM J. Optim.*, vol. 10, no. 4, pp. 1196-1211, 2000.