A cyclic block coordinate gradient projection algorithm

Silvia Bonettini, Dipartimento di Matematica, Università di Ferrara

Motivation: Many real life problems can be modeled as smooth, large scale optimization problems whose variable is constrained in a cartesian product of convex sets.

$$\min \; f(x) \quad \text{s.t.} \quad x \in \Omega = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_m \subseteq \mathbb{R}^n$$

$$x = (x_1, x_2, \ldots, x_m) \quad x_i \in \Omega_i$$

Basic idea: Decouple the optimization problem by cycling over the blocks

Nonlinear Gauss-Seidel method

$$x_i^{(k+1)} = \arg\min_{y \in \Omega_i} f(x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}, y, x_{i+1}^{(k)}, \ldots, x_m^{(k)})$$

Drawbacks of the nonlinear GS method:

- no convergence for m>2 without strict convexity assumptions;
- exact solution of a constrained minimization problem required at each partial iteration

Recent developments:

- convergence result for m=2 in the nonconvex, constrained case [5]
- globally convergent line-search based schemes for the unconstrained case [4]

Proposed approach: a descent method based on the projected gradient properties.

The Cyclic Block Gradient Projection (CBGP) method

Choose the starting point $x^{(0)} \in \Omega$ and a positive integer L

For $k = 0, 1, 2, \ldots$

For $i = 1, \ldots, m$

- Choose the inner iterations number $\ell_i^{(k)} \leqslant L$
- Apply $\ell_i^{(k)}$ SGP iterations to the problem

$$\min_{y \in \Omega_i} J_i(y) \equiv f(x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}, y, x_{i+1}^{(k)}, \ldots, x_m^{(k)})$$

to compute $x_i^{(k+1)}$

End

End

The Scaled Gradient Projection (SGP) method

$$\min_{y \in \mathscr{Y}} J(y)$$

Choose the starting point $y^{(0)} \in \mathscr{Y}$ and set the parameters $\beta \in (0,1)$ $0 < \alpha_{min} < \alpha_{max}$

For $\ell = 0, 1, 2, \ldots$

- Choose the stepleng $\alpha_\ell \in [\alpha_{min}, \alpha_{max}]$ and a positive definite scaling matrix D_ℓ
- Compute the descent direction

$$d^{(\ell)} = P_{\mathscr{Y}, D_\ell^{-1}}(y^{(\ell)} - \alpha_\ell D_\ell \nabla J(y^{(\ell)})) - y^{(\ell)}$$

- Armijo Backtracking loop: compute λ_ℓ such that

$$J(y^{(\ell)} + \lambda_\ell d^{(\ell)}) \leqslant J(y^{(\ell)}) + \beta \lambda_\ell \nabla J(y^{(\ell)})^T d^{(\ell)}$$

- Set $y^{(\ell+1)} = y^{(\ell)} + \lambda_\ell d^{(\ell)}$

End

Convergence analysis: [1]

Let $\{x^{(k)}\}$ the sequence generated by the CBGP algorithm and assume that x^* is a limit point of $\{x^{(k)}\}$. Then, x^* is a stationary point.

Remarks: no convexity assumptions, no limitations on the numbers of blocks, convergence for any choice of inner iterations number (just bounded).

Implementation issues: exploit scaling matrix and steplenght choices to improve convergence (e.g. Barzilai-Borwein steplength selection rules).

Applications: [1],[2],[3]

Nonnegative Matrix Factorization (NMF)

Given a data matrix $V \in \mathbb{R}^{n \times m}$ and a positive integer $r < m$, find $W \in \mathbb{R}^{n \times r}, H \in \mathbb{R}^{r \times m}$ s.t.

$$\min_{W \geq 0, H \geq 0} \frac{1}{2} \|WH - V\|_F^2$$

Finds sparse representation of data

Semi blind deconvolution from sparse Fourier data

Given the data $g \in \mathbb{C}^N$, find the image $f \in \mathbb{R}^{n \times n}$ $f_{jh} = f(x_{jh})$ $x_{jh} \in \mathbb{R}^2, j, h = 1, \ldots, n$ and the frequencies $\omega \in \mathbb{R}^{N \times 2}$ s.t.

$$\min_{\substack{f \geqslant 0 \\ l \leqslant \omega \leqslant u}} \frac{1}{2} \|\mathscr{A}(\omega)[f] - g\|^2$$

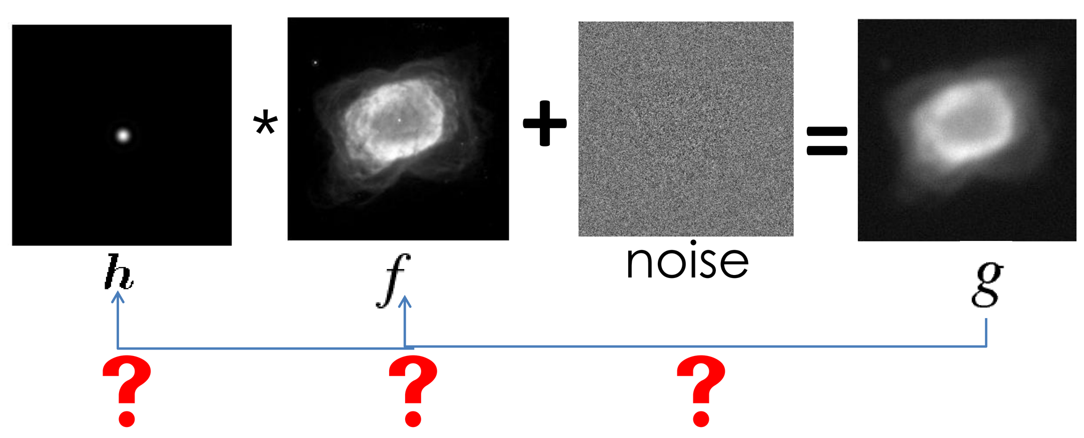
where $\mathscr{A}(\omega)[f]_k = \sum_{j,h} f_{jh} e^{-2\pi i \, x_{jh}^T \omega_k}$

Applications to solar flares image reconstruction [2]

Blind deconvolution

Given the observed image $g \in \mathbb{R}^{n \times n}$, find the true image $f \in \mathbb{R}^{n \times n}$ and the psf $h \in \mathbb{R}^{n \times n}$ s.t.

$$\min_{\substack{f \geqslant 0 \\ 0 \leqslant h \leqslant u \\ \sum_{ij} h_{ij} = 1}} \text{dist}(f * h, g)$$



References:

[1] S. Bonettini (2011). *Inexact block coordinate descent methods with application to nonnegative matrix factorization*, IMA J. Numer. Anal., **31**:1131-1452

[2] S. Bonettini, A. Cornelio, M. Prato, (2012). *Semi blind deconvolution for Fourier based image restoration*, submitted

[3] M. Bertero, S. Bonettini, A. La Camera, M. Prato, (2012). *A new blind deconvolution approach with Strehl constrained psf*, in preparation

[4] L. Grippo, M. Sciandrone (1999). *Globally convergent block coordinate techniques for unconstrained optimization*, Optim. Methods Softw., **10**:587-637

[5] L. Grippo, M. Sciandrone (2000). *On the convergence of block nonlinear Gauss-Seidel method under convex constraints*, Oper. Res. Lett., **26**:127-136