Fast Alternating Direction Augmented Lagrangian Methods for Solving Convex Optimization Problems

DOE Grant #: DE-FG02-08ER-25856

P.I. Donald Goldfarb[†], Co P.I. Garud Iyengar[†]

[†] Department of Industrial Engineering and Operations Research, Columbia University

Minimizing Sum of Convex Functions

- Compressive Sensing: $\{\min_{x\in\mathbb{R}^n} \rho\|x\|_1 + \frac{1}{2}\|Ax-b\|^2\}$
- Matrix Rank Minimization: $\{\min_{X\in\mathbb{R}^{m\times n}} \rho\|X\|_* + \frac{1}{2}\|\mathcal{A}(X)-b\|^2\}$
- Total Variation Image Restoration: $\{\min_{x\in\mathbb{R}^n} \alpha TV(x) + \beta\|Wx\|_1 + \frac{1}{2}\|Ax-b\|^2\}$
- Sparse Covariance Selection: $\{\min_{X\in S^n_+} -\log\det(X) + \langle\Sigma, X\rangle + \rho\|X\|_1\}$
- Robust PCA: $\{\min_{X,Y\in\mathbb{R}^{m\times n}} \|X\|_* + \rho\|Y\|_1, \text{s.t. } X+Y=M\}$

ALM and FALM: Alternating Linearization Methods for Minimizing the Sum of Two Convex Functions

ALM: Alternating Linearization Method

Algorithm: ALM (FALM)

$x^{k+1} := \arg\min_x f(x) + g(z^k) + \langle\nabla g(z^k), x - z^k\rangle + \frac{1}{2\mu}\|x - z^k\|^2$;

$y^{k+1} := \arg\min_Y f(x^{k+1}) + \langle\nabla f(x^{k+1}), y - x^{k+1}\rangle + \frac{1}{2\mu}\|x^{k+1} - y\|^2 + g(y)$;

$z^{k+1} := y^{k+1}$;

if *FALM* **then**

$w^{k+1} := (x^{k+1} + y^{k+1})/2$;

$t_{k+1} := (1 + \sqrt{1 + 4t_k^2})/2$;

$z^{k+1} := w^{k+1} + \frac{1}{t_{k+1}}[t_k(y^{k+1} + w^k) - (w^{k+1} - w^k)]$;

Interpret ALM as symmetric ADAL

$$\begin{cases} x^{k+1} := \arg\min_x \mathcal{L}(x, y^k; \lambda^k) \\ \lambda^{k+\frac{1}{2}} := \lambda^k - (x^{k+1} - y^k)/\mu \\ y^{k+1} := \arg\min_y \mathcal{L}(x^{k+1}, y; \lambda^{k+\frac{1}{2}}) \\ \lambda^{k+1} := \lambda^{k+\frac{1}{2}} - (x^{k+1} - y^{k+1})/\mu \end{cases}$$

Properties of ALM and FALM:

- ALM: $O(1/\epsilon)$ iterations to obtain an ϵ-optimal solution.
- FALM: $O(1/\sqrt{\epsilon})$ iterations to obtain an ϵ-optimal solution. Optimal first-order method in terms of iteration complexity.
- Gauss-Seidel like method.

MSA and FaMSA: Multiple Splitting Algorithms for Minimizing the Sum of K Convex Functions

Solve $\min\sum_{i=1}^K f_i(x)$. Use $\min f(x) + g(x) + h(x)$ to illustrate MSA and FaMSA for simplicity. Define

$Q_{gh}(u,v,w) := f(u) + g(v) + \langle\nabla g(v), u - v\rangle + \frac{1}{2\mu}\|u-v\|^2 + h(w) + \langle\nabla h(w), u - w\rangle + \frac{1}{2\mu}\|u - w\|^2$.

$Q_{fh}(u,v,w)$ and $Q_{fg}(u,v,w)$ are defined similarly.

MSA (FaMSA-s): (Fast) Multiple Splitting Algorithm

Algorithm: MSA (FaMSA-s)

$x^{k+1} := \arg\min_x Q_{gh}(x, \bar{w}^k, \bar{w}^k)$;

$y^{k+1} := \arg\min_y Q_{fh}(\bar{w}^k, y, \bar{w}^k)$;

$z^{k+1} := \arg\min_z Q_{fg}(\bar{w}^k, \bar{w}^k, z)$;

$w^{k+1} := (x^{k+1} + y^{k+1} + z^{k+1})/3$, $\bar{w}^{k+1} := w^{k+1}$;

if *FaMSA-s* **then**

$t_{k+1} := (1 + \sqrt{1 + 4t_k^2})/2$;

$\bar{w}^{k+1} := w^{k+1} + \frac{t_k - 1}{t_{k+1}}(w^{k+1} - w^k)$;

Properties of MSA and FaMSA-s

- MSA: $O(1/\epsilon)$ iterations to obtain an ϵ-optimal solution.
- FaMSA: $O(1/\sqrt{\epsilon})$ iterations to obtain an ϵ-optimal solution. Optimal first-order method in terms of iteration complexity.
- Jacobi like method
- parallelizable

ALB: Accelerated Linearized Bregman method

(Accelerated) Linearized Bregman method (LB/ALB)

Algorithm: LB/ALB

$w^{k+1} := \arg\min_w J(w) + \frac{1}{2\mu}\|w\|^2 - \langle y^k, Aw - b\rangle$;

$y^{k+1} := \tilde{y}^k - \tau(Aw^{k+1} - b)$, $\quad \tilde{y}^{k+1} := y^{k+1}$;

if *ALB* **then**

$\tilde{y}^{k+1} := \alpha_k y^{k+1} + (1 - \alpha_k)y^k$, where $\alpha_k = \frac{3}{k+2}$;

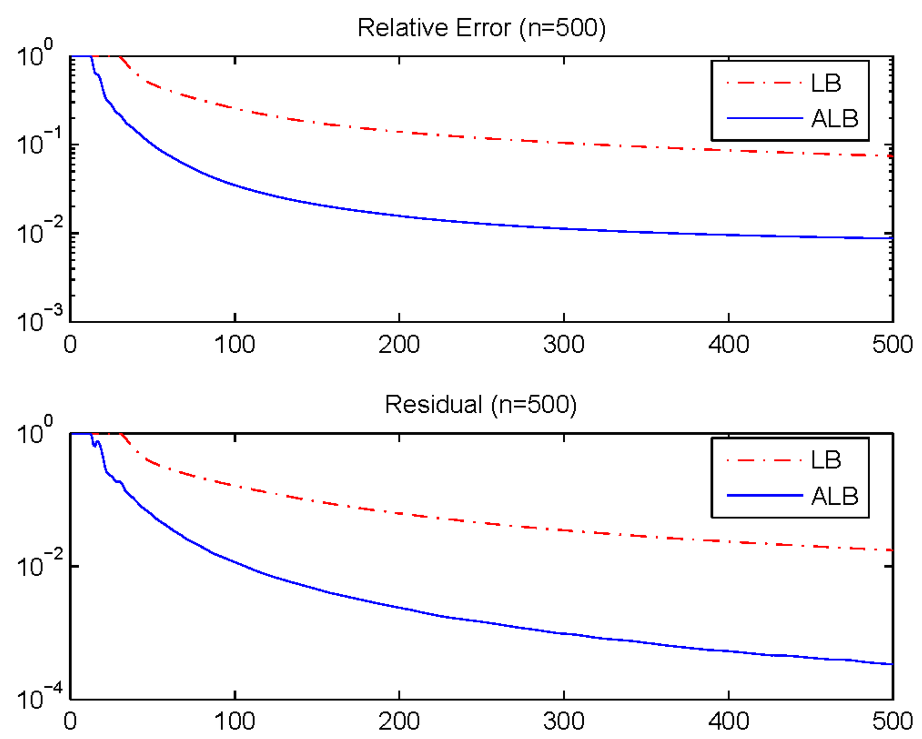


Figure: Comparison of LB and ALB on a matrix completion problem with $n = 500$, rank=10, FR=0.3

Stable Principal Component Pursuit (SPCP)

- Stable PCA: $\min_{X,Y\in\mathbb{R}^{m\times n}}\{\text{rank}(X) + \rho\|Y\|_0, \text{s.t. } \|X + Y - M\|_F \le \delta\}$
- SPCP Convex Relaxation: $\min_{X,Y\in\mathbb{R}^{m\times n}}\{\|X\|_* + \rho\|Y\|_1, \text{s.t. } \|X + Y - M\|_F \le \delta\}$
- Application: Background/foreground decomposition of noisy surveillance video

ADAL: Alternating Direction Augmented Lagrangian method

- Solve $\min_{X,Y,Z\in\mathbb{R}^{m\times n}}\{\|X\|_* + \rho\|Y\|_1 \mid X = Z, (Y,Z)\in\Omega\}$, where $\Omega \equiv \{Y, Z \mid \|Z + Y - M\|_F \le \delta\}$ by solving $\min_{X,Y,Z}\{\mathcal{L}(X,Y,Z;\Gamma) \equiv \|X\|_* + \rho\|Y\|_1 + \langle\Gamma, X - Z\rangle + \frac{\mu}{2}\|X - Z\|_F^2 \mid (Y,Z)\in\Omega\}$ first w.r.t. X and then w.r.t. (Y,Z).
- Subproblem for X corresponds to an SVD.
- Subproblem for (Y,Z) corresponds to a vector shrinkage operation, plus $O(mn\log mn)$ operations to compute the Lagrange multiplier for the constraint $(Y,Z)\in\Omega$.

Table: Comparison of NSA and ASALM (Tao and Yao)

Alg.	NSA	ASALM
CPU	160.8	910.0
no. SVD	19	94
rank(X)	81	89
$\frac{\|X+S-D\|_F}{\|D\|_F}$	0.00068	0.00080

ADAL for Total Variation (TV) Denoising

TV Denoising: $\min_u \quad \lambda\|u\|_{TV} + \frac{1}{2}\|u - b\|^2$

Variable-splitting:

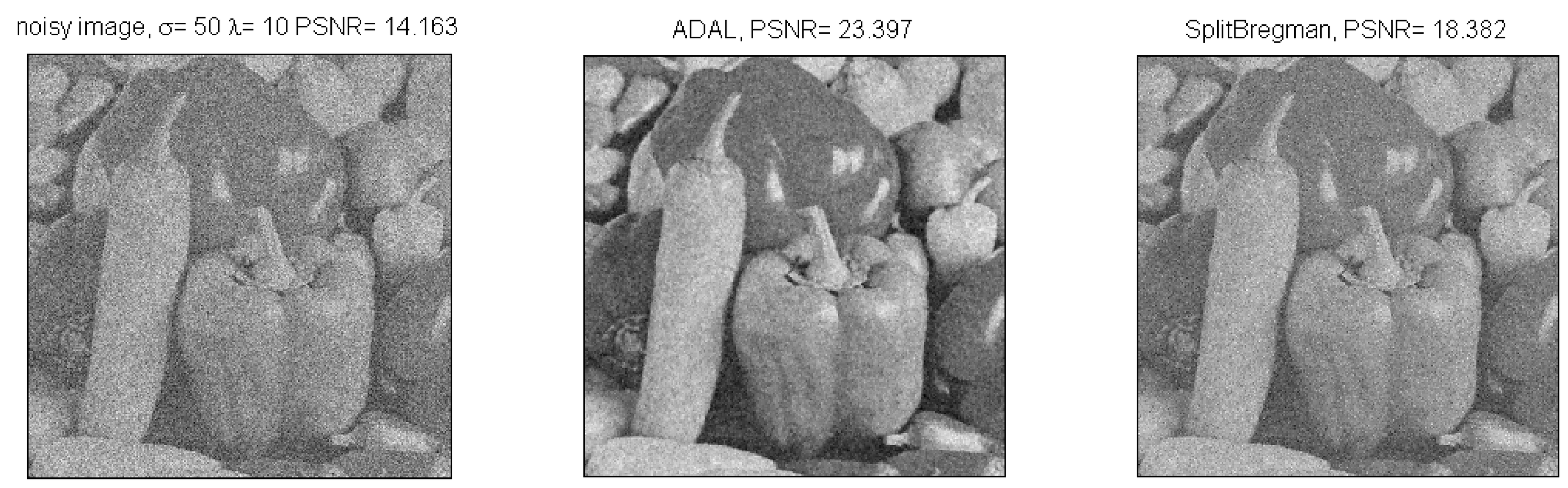
$$\min_{d,u,v} \lambda(\|d_x\|_1 + \|d_y\|_1) + \frac{1}{2}\|u - b\|^2 \quad \text{s.t.} \quad d_x = Du, d_y = Dv, v = Pu$$

ADAL (TV denoising):

$$\begin{cases} d_x^{(k+1)} \leftarrow \mathcal{T}(Du^{(k)} + \mu_1\gamma_x^{(k)}, \lambda\mu_1) \\ d_y^{(k+1)} \leftarrow \mathcal{T}(Dv^{(k)} + \mu_1\gamma_y^{(k)}, \lambda\mu_1) \\ v^{(k+1)} \leftarrow v(d_y^{(k+1)}, u^{(k)}, \gamma_y^{(k)}, \gamma_z^{(k)}) \\ u^{(k+1)} \leftarrow u(d_x^{(k+1)}, v^{(k+1)}, \gamma_x^{(k)}, \gamma_z^{(k)}) \end{cases} \qquad \begin{aligned} \gamma_x^{(k+1)} &\leftarrow \gamma_x^{(k)} + \frac{1}{\mu_1}(Du^{(k+1)} - d_x^{(k+1)}) \\ \gamma_y^{(k+1)} &\leftarrow \gamma_y^{(k)} + \frac{1}{\mu_1}(Dv^{(k+1)} - d_y^{(k+1)}) \\ \gamma_z^{(k+1)} &\leftarrow \gamma_z^{(k)} + \frac{1}{\mu_2}(Pu^{(k+1)} - v^{(k+1)}), \end{aligned}$$

where $\mathcal{T}(x,\lambda)_i := \max\{|x_i| - \lambda, 0\}sign(x_i)$, D is an upper bidiagonal matrix, and P is a permutation matrix.

Denoising Results (after 25 iterations)



Collaborators

Shiqian Ma, Katya Scheinberg, Necdet S. Aybat, Zhiwei Tony Qin, Bo Huang

Department of Industrial Engineering and Operations Research, Columbia University   Email: goldfarb@columbia.edu   WWW: http://www.columbia.edu/~goldfarb