# Block Coordinate Descent on Smooth Manifolds


Liangzu Peng     René Vidal

Center for Innovation in Data Engineering and Science (IDEAS)


## General Optimization Problem

$$\underset{\text{differentiable}}{\min_{x\in\mathcal{M}} F(x)} \qquad \mathcal{M} := \underbrace{\mathcal{M}_1 \times \cdots \times \mathcal{M}_b}_{\text{non-empty}}$$

**Challenge**: non-convexity, constraints

**Hope**: leverage the product structure of $\mathcal{M}$

## Algorithm: Block Coordinate Descent

The idea is to iteratively update one block of variables $x_i^{t+1} \in \mathcal{M}_i$ at a time. Specifically:

- Initialize $x^0 = [x_1^0; \dots; x_b^0] \in \mathcal{M}$
- For $t = 1, \dots, T$ and for $i = 1, \dots, b$:

  Update $x_i^{t+1}$ from $\underbrace{x_{1:i-1}^{t+1}}_{[x_1^{t+1};\dots;x_{i-1}^{t+1}]}$ and $\underbrace{x_{i+1:b}^{t}}_{[x_{i+1}^{t};\dots;x_b^{t}]}$

We consider two update rules:

- Block exact minimization

$$x_i^{t+1} \in \operatorname*{argmin}_{\xi\in\mathcal{M}_i} F\left(x_{1:i-1}^{t+1}, \xi, x_{i+1:b}^{t}\right)$$

- Block Riemannian gradient descent

A map from the tangent space $T_{x_i^t}\mathcal{M}_i$ to $\mathcal{M}_i$

$$x_i^{t+1} = \mathrm{Retr}_{x_i^t}\left(s_i^t\right), \text{where}$$

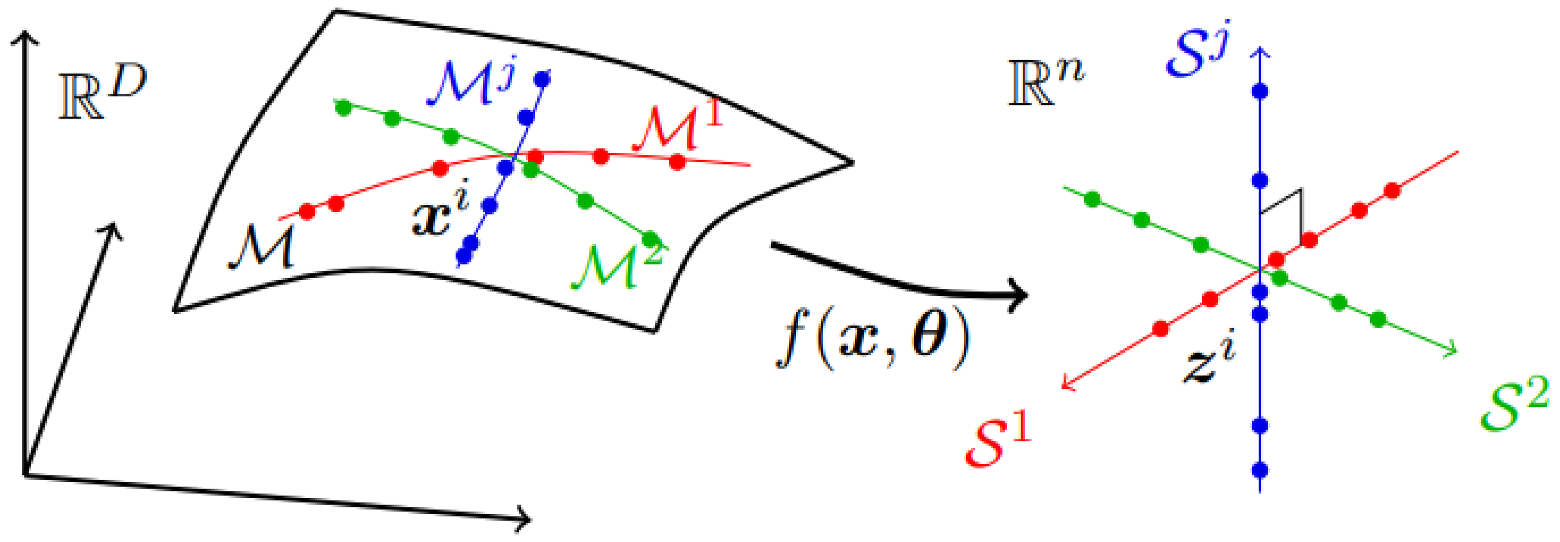
$$s_i^t := -\lambda_i^t \cdot \widetilde{\nabla}_i F\left(x_{1:i-1}^{t+1}, x_i^t, x_{i+1:b}^t\right)$$

stepsize     Riemannian gradient w.r.t. $x_i^t$

## Examples

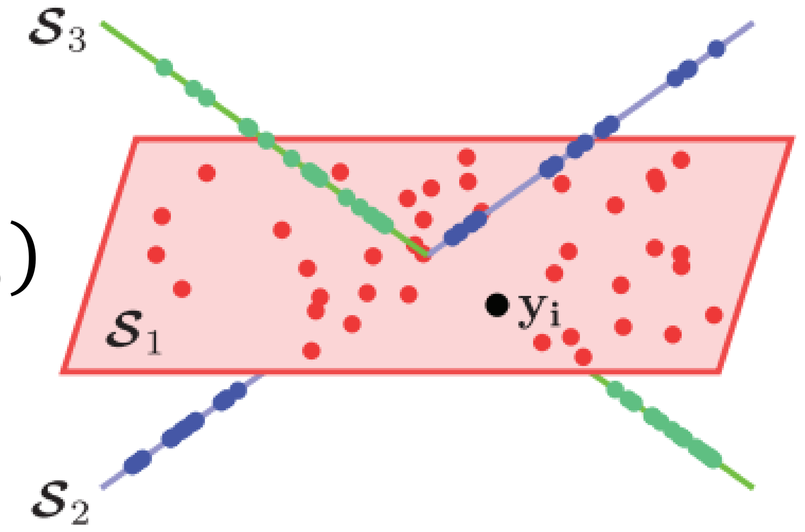
- *Maximum Coding Rate Reduction*

$$\min_{z_i\in\mathcal{M}} F(z_1, \dots, z_n; x_1, \dots, x_n)$$




- *Generalized PCA*

$$\min_{\mathcal{S}_i:subspace} F(\mathcal{S}_1, \dots, \mathcal{S}_b)$$

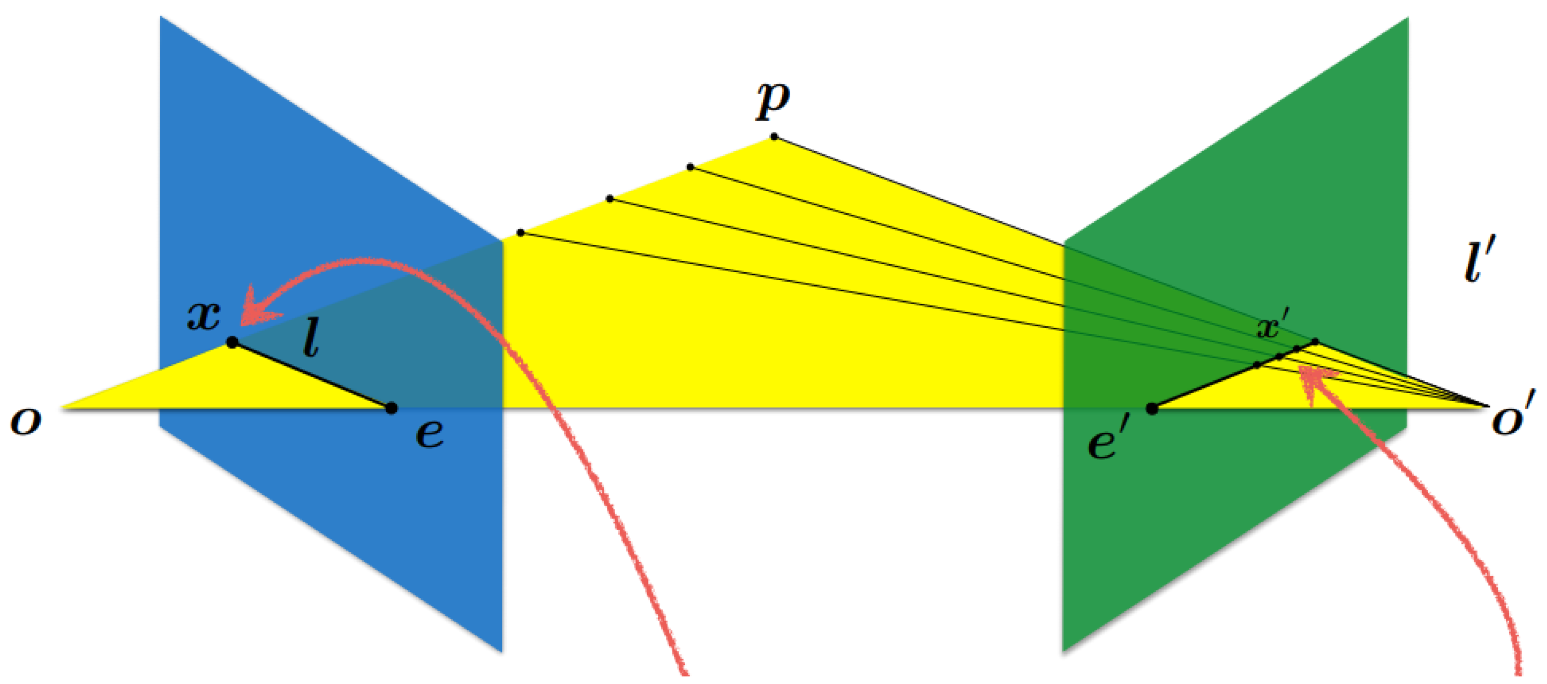



- *Essential Matrix Estimation*

$$\min_E (x_1^\top E x_1')^2 + \cdots + (x_m^\top E x_m')^2$$

Here, the essential matrix $E$ is defined by some 3D rotation $R$ and unit vector $t$, i.e.,

$$Ex' := t \times (Rx')$$




Caution: Figures shown above are not original.

## Convergence Theory

**Theorem 1 (Asymptotic Convergence)**

Assume:
- $\mathcal{M}_i$ is a closed subset of $\mathbb{R}^{n_i}$ ($\forall i$)
- Set $\{x: F(x) \le \gamma\}$ is bounded ($\forall \gamma \in \mathbb{R}$)
- Minimizing $F(x_{1:i-1}, \xi, x_{i+1:b})$ over $\xi \in \mathcal{M}_i$ has a unique minimizer ($\forall x_j \in \mathcal{M}_j$)

Then the sequence $\{x^t\}_t$ generated by block exact minimization has limit points, and each limit point is a stationary point.

**Definition (Block-$i$ Smoothness)**

$F$ is block-$i$ $L_i$-smooth with constant $L_i$ if $\nabla_i F(x_{1:b-1}, \xi, x_{i+1:b}) - \nabla_i F(x_{1:b-1}, \zeta, x_{i+1:b})$ has norm smaller than or equal to $L_i \|\xi - \zeta\|$ for every $\zeta \in \mathcal{M}_i$ and $[x_{1:b-1}; \xi; x_{i+1:b}] \in \mathcal{M}$.

**Theorem 2 (Sublinear Convergence)**

Assume ($\forall i$):
- $F$ is block-$i$ $L_i$-smooth with constant $L_i$
- $\mathcal{M}_i$ is a compact manifold of $\mathbb{R}^{n_i}$

Then block exact minimization and block Riemannian gradient descent converge at a sublinear rate: Their iterates $\{x^t\}_t$ satisfy

$$\min_{t=0,\dots,T} \left\|\widetilde{\nabla} F(x^t)\right\| = O\left(\frac{1}{\sqrt{T}}\right).$$

## Ongoing Work

- Prove stronger convergence guarantees
- Discover more applications & examples