E9 203 Compressed Sensing & Sparse SP

19 Mar. 2021

Last time:

- Extreme pts, BFS, concave opt.
- Properties of local minima
- Reweighting based methods (MM)

Today:

- Majorization-minimization (continued)
- Reweighting based algos for sparse recovery.

Recall

$$\min_{x \in \mathbb{R}^N} \sum_{i=1}^{N} g(x_i) \quad \text{s.t. } Ax = y$$

1. $g(x) = g(-x) = g(|x|)$
2. $g(x)$ monotone $\uparrow$ for $x \in \mathbb{R}^+$
3. $g(x)$ strictly concave for $x \in \mathbb{R}^+$

Noisy data:

$$\min_{x \in \mathbb{R}^+} \lambda \sum_{i=1}^{N} g(x_i) + \|Ax - y\|_2^2$$

or $\min_{x \in \mathbb{R}^N} \sum_{i=1}^{N} g(x_i)$ s.t. $\|Ax - y\|_2^2 \leq \varepsilon$.

Majorization-minimization:

Let $f(\theta)$ be the fn. to be minimized.

Let $g(\theta \mid \theta^{(m)}) \geq f(\theta) \;\; \forall \theta$ | Cleverly choose such a $g(\theta \mid \theta^{(m)})$.

$g(\theta^{(m)} \mid \theta^{(m)}) = f(\theta^{(m)})$

$g(\theta \mid \theta^{(m)})$ easy to optimize

Algo.:

Init.: $\theta^{(0)} =$ something convenient

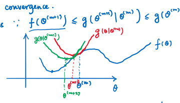
Iterate

$$\theta^{(m+1)} = \arg\min_{\theta} g(\theta \mid \theta^{(m)}) \leftarrow$$

Determine $g(\theta \mid \theta^{(m+1)})$

Until convergence.

Works $\because$ $f(\theta^{(m+1)}) \leq g(\theta^{(m+1)} \mid \theta^{(m)}) \leq g(\theta^{(m)} \mid \theta^{(m)}) = f(\theta^{(m)})$



An important property of differentiable convex fns.:
(Useful in selecting $g(\theta \mid \theta^{(m)})$)

Let $f \in C^1$. Then f is convex over a convex set Ω iff $f(y) \geq f(x) + \nabla f(x)(y-x) \;\; \forall x, y \in \Omega$

Proof: f convex $\Rightarrow \forall\, 0 \leq \alpha \leq 1$

$$f(\alpha y + (1-\alpha)x) \leq \alpha f(y) + (1-\alpha) f(x)$$

$$\frac{f(x + \alpha(y-x)) - f(x)}{\alpha} \leq f(y) - f(x)$$

Letting $\alpha \to 0$

$$\nabla f(x)(y-x) \leq f(y) - f(x). \quad \text{[only if]}$$

Now, assume

$$f(y) \geq f(x) + \nabla f(x)(y-x)$$

Fix $x_1, x_2 \in \Omega$, $\alpha \in [0,1]$

Then $x = \alpha x_1 + (1-\alpha)x_2 \in \Omega$

Set $y = x_1$ & $y = x_2$ alternatively to get

$$f(x_1) \geq f(x) + \nabla f(x)(x_1 - x)$$

$$f(x_2) \geq f(x) + \nabla f(x)(x_2 - x)$$

$$\alpha f(x_1) + (1-\alpha) f(x_2) \geq f(x) + \nabla f(x)\underbrace{(\alpha x_1 + (1-\alpha)x_2 - x)}_{=0}$$

$$\alpha f(x_1) + (1-\alpha) f(x_2) \geq f(\alpha x_1 + (1-\alpha)x_2)$$

$\Rightarrow f$ is convex. $\square$

Iteratively reweighted least squares (IRLS)

$$g(x) = |x|^p, \quad 0 < p < 2$$

$$= (x^2)^{p/2}$$

$$h(\tilde{y}) = \tilde{y}^{p/2}, \quad \tilde{y} \geq 0.$$

$$g(x) = h(x^2)$$

Check: $h(\tilde{y})$ strictly concave for $0 < p < 2$.

$$\Rightarrow h(\tilde{y}) < h(\tilde{y}_0) + h'(\tilde{y}_0)(\tilde{y} - \tilde{y}_0)$$

$$= h(\tilde{y}_0) + \frac{p}{2}(\tilde{y}_0)^{\frac{p}{2}-1}(\tilde{y} - \tilde{y}_0)$$

$$= \frac{p}{2}(\tilde{y}_0)^{\frac{p}{2}-1}\tilde{y} + \underbrace{h(\tilde{y}_0) - \frac{p}{2}(\tilde{y}_0)^{\frac{p}{2}-1}\tilde{y}_0}_{\text{Not imp. for opt.}}$$

Hence,

$$\|Ax - y\|_2^2 + \lambda \sum_{\ell=1}^{N} h(x_\ell^2)$$

$$\leq \|Ax - y\|_2^2 + \lambda \sum_{\ell=1}^{N} \frac{p}{2}\left((x_\ell^{(k)})^2\right)^{\frac{p}{2}-1} x_\ell^2$$

$+$ Const. terms $\perp\!\!\!\perp$ of x

let $\lambda' = \frac{\lambda p}{2}$.

$$= \|Ax - y\|_2^2 + \lambda' \|W_k^{-1} x\|_2^2$$

where $W_k^{-1} \triangleq \operatorname{diag}\left(|x_\ell^{(k)}|^{\frac{p}{2}-1}\right)$: $N \times N$

Let $W_k^{-1} x = z$ or $x = W_k z$

$$\|A W_k z - y\|_2^2 + \lambda' \|z\|_2^2$$

$$\Rightarrow W_k^T A^T (A W_k z - y) + \lambda' z = 0$$

$$(W_k^T A^T A W_k + \lambda' I_N) z = W_k^T A^T y$$

$$\Rightarrow x^{(k+1)} = W_k z = W_k \underbrace{(W_k^T A^T A W_k + \lambda' I_N)^{-1}}_{N \times N \text{ matrix}} W_k^T A^T y$$

Matrix inversion Lemma:

$$(A + UCV)^{-1} = A^{-1} - A^{-1} U (C^{-1} + V A^{-1} U)^{-1} V A^{-1}$$

$$(\lambda' I_N + (AW_k)^T I_m (AW_k))^{-1} = \frac{1}{\lambda'} I_N - \frac{I}{\lambda'} (AW_k)^T \left(I_m + \frac{A W_k W_k^T A^T}{\lambda'}\right)^{-1} \frac{AW_k}{\lambda'}$$

$$x^{(k+1)} = \frac{W_k W_k^T A^T y}{\lambda'} - \frac{W_k^2 A^T}{\lambda'} (\lambda' I_m + A W_k^2 A^T)^{-1} A W_k^2 A^T y$$

$$= \frac{W_k^2 A^T}{\lambda'} \Big(\underbrace{I_m}_{(\lambda' I_m + A W_k^2 A^T)^{-1}(\lambda' I_m + A W_k^2 A^T)} - (\lambda' I_m + A W_k^2 A^T)^{-1} A W_k^2 A^T \Big) y$$

$$= \frac{W_k^2 A^T}{\cancel{\lambda'}} \left(\cancel{\lambda'} (\lambda' I_m + A W_k^2 A^T)^{-1} \right) y$$

$$\boxed{x^{(k+1)} = W_k^2 A^T (A W_k^2 A^T + \lambda' I_m)^{-1} y} \leftarrow 2$$

$$\boxed{W_k^{-1} = \operatorname{diag}\left(|x_\ell^{(k)}|^{\frac{p}{2}-1}\right)}. \leftarrow 1$$

FOCUSS algorithm (Focal Underdetermined System Solver)

Another option for the regularizer [Chatrand & Yin]:

$$g(x) = (x^2 + \beta)^{p/2}, \quad 0 < p < 2, \quad \beta > 0, \quad x \in \mathbb{R}$$

$$h(\tilde{y}) = (\tilde{y} + \beta)^{p/2}, \quad \tilde{y} \geq 0$$

$$h'(\tilde{y}) = \frac{p}{2}(\tilde{y} + \beta)^{\frac{p}{2}-1}$$

$$h''(\tilde{y}) = \frac{p}{2}\underbrace{\left(\frac{p}{2} - 1\right)}_{<0}(\tilde{y} + \beta)^{\frac{p}{2}-2} < 0 \text{ : concave.}$$

As before,

$$h(\tilde{y}) \leq h'(y_0) \cdot (\tilde{y} - \tilde{y}_0) + h(\tilde{y}_0)$$

Thus, at the $(k+1)^{\text{th}}$ update:

$$\min_x \|Ax - y\|_2^2 + \lambda \sum_{\ell=1}^{N} \left[\left(|x_\ell^{(k)}|^2 + \beta\right)^{\frac{p}{2}-1}\right] x_\ell^2$$

$$= \min_x \|Ax - y\|_2^2 + \lambda \|W_k^{-1} x\|_2^2$$

$$W_k^{-2} \triangleq \operatorname{diag}\left(\left(|x_\ell^{(k)}|^2 + \beta\right)^{\frac{p}{2}-1}\right).$$

... as before.

Same approach as before.

Reweighted ℓ_1 :

Bound $g(x)$, $x \geq 0$ as a fn. of x, not x^2.

Let $g(x) \leq f_1(x)$, $x \geq 0$

Then $x \in \mathbb{R}$ covered by $f(x) = f_1(|x|)$ $\forall x \in \mathbb{R}$.

Then $g(x) = g(|x|) \leq f(x) = f_1(|x|)$ $\forall x \in \mathbb{R}$.

Concentrate on $x \geq 0$

Concave : $g(x) \leq g'(x_0)(x - x_0) + g(x_0)$

E.g. $g(x) = |x|^p$, $0 < p < 1$

$g'(x) = p\, x^{p-1}$

For $x \geq 0$, $g(x) \leq p\, x_0^{p-1}(x - x_0) + g(x_0)$

$\Rightarrow$ $f_1(x) = p\,|x_0|^{p-1}(x - |x_0|) + g(|x_0|)$

$f(x) = f_1(|x|) = p\,|x_0|^{p-1}(|x| - |x_0|) + g(|x_0|)$

Hence, $g(|x|) \leq \underline{p\,|x_0|^{p-1}|x|} - \underbrace{p\,|x_0|^p + g(|x_0|)}_{\perp\!\!\!\perp \text{ of } x \text{, not imp. for opt.}}$

At $(k+1)^{th}$ update

$$\min_x \; \|Ax - y\|_2^2 + \underbrace{\lambda \, \|W_k^{-1} x\|_1}_{\lambda \sum_{i=1}^{N} |x_i^{(k)}|^{p-1} |x_i|} \leftarrow \text{Note!}$$

where $W_k^{-1} \triangleq \text{diag}\left(|x_i^{(k)}|^{p-1}\right)$.

Similarly, if $g(|x|) = (|x| + \beta)^p$, $\beta > 0$, $0 < p < 1$

$g'(x) = p\,(x + \beta)^{p-1}$

$$\min_x \; \|Ax - y\|_2^2 + \lambda \, \|W_k^{-1} x\|_1 \quad (\star)$$

with $W_k^{-1} \triangleq \text{diag}\left((|x_i^{(k)}| + \beta)^{p-1}\right)$.

Can use the "weighted ℓ_1 min" approach we discussed previously to efficiently solve $(\star)$.