In Defense of Sparse Tracking: Circulant Sparse Tracker (Supplementary material)

Say $\mathbf{A}$ is block wise circulant, such that $\mathbf{A}_i$ is circulant $\forall i$. $\mathbf{A} = [\mathbf{A}_1|\mathbf{A}_2|\mathbf{A}_3|...|\mathbf{A}_N]$ where $\mathbf{A}_i \in \mathbb{R}^{d\times d}$. Then, $\mathbf{A} \in \mathbb{R}^{d\times Nd}$ and N denotes the number of samples (Training examples). $\mathbf{y} \in \mathbb{R}^{d\times 1}$, and $\mathbf{x} \in \mathbb{R}^{Nd\times 1}$ Then the primal formulation for the Lasso is given by:

$$P1 : \underset{\mathbf{x}}{\text{minimize}} \quad ||\mathbf{A}\mathbf{x} - \mathbf{y}||_2^2 + \lambda||\mathbf{x}||_1 \tag{1}$$

By setting $\mathbf{r} = A\mathbf{x} - \mathbf{y}$, then P1 can be re-written as:

$$\begin{aligned} P2 : \underset{\mathbf{x},\mathbf{r}}{\text{minimize}} &\quad ||\mathbf{r}||_2^2 + \lambda||\mathbf{x}||_1 \\ \text{subject to} &\quad \mathbf{r} = \mathbf{A}\mathbf{x} - \mathbf{y} \end{aligned} \tag{2}$$

1. Dual Formulation

To obtain the dual formulation, we minimize the lagrangian over the primal variables $\mathbf{x}, \mathbf{r}$ as follows:

$$\begin{aligned} \min_{\mathbf{x},\mathbf{r}} \mathcal{L}(\mathbf{x},\mathbf{r},\Psi) &= \min_{\mathbf{x},\mathbf{r}} ||\mathbf{r}||_2^2 + \lambda||\mathbf{x}||_1 + \Psi^T(\mathbf{A}\mathbf{x} - \mathbf{y} - \mathbf{r}) && (3) \\ &= \min_{\mathbf{r}} (||\mathbf{r}||_2^2 - \Psi^T\mathbf{r}) + \min_{\mathbf{x}}(\lambda||\mathbf{x}||_1 + \Psi^T\mathbf{A}\mathbf{x}) - \Psi^T\mathbf{y} && (4) \\ &= -\Psi^T\mathbf{y} - \frac{1}{4}||\Psi||_2^2 - \lambda\max_{\mathbf{x}}(\frac{\mathbf{x}^T(-\mathbf{A}^T\Psi)}{\lambda} - ||\mathbf{x}||_1) && (5) \\ &= -\Psi^T\mathbf{y} - \frac{1}{4}||\Psi||_2^2 - \mathbb{1}_{(\frac{||-\mathbf{A}^T\Psi||_\infty}{\lambda} \leq 1)} && (6) \end{aligned}$$

Note that we used the conjugate function definition of the ℓ_1 norm to obtain Eq (6) from (5). Therefore, the dual problem of 1 and 2 is given by:

$$\begin{aligned} P3 : \underset{\Psi}{\text{minimize}} &\quad \frac{1}{4}||\Psi||_2^2 + \Psi^T\mathbf{y} \\ \text{subject to} &\quad ||\mathbf{A}^T\Psi||_\infty \leq \lambda \end{aligned} \tag{7}$$

2. ADMM Formulation

To solve the problem using ADMM, problem 7 will be set in the standard format by letting:

$$\begin{aligned} f(\Psi) &= \frac{1}{4}||\Psi||_2^2 + \Psi^T\mathbf{y} \\ g(\zeta) &= \mathbf{1}_{(||\zeta||_\infty \leq \lambda)} \end{aligned} \tag{8}$$

where $\zeta = A^T\Psi$. Then the dual problem can be re-witten as:

$$\begin{aligned} \underset{\Psi,\zeta}{\text{minimze}} &\quad f(\Psi) + g(\zeta) \\ \text{subject to} &\quad \zeta = A^T\Psi \end{aligned} \tag{9}$$

The augmented lagrangian is given as follows:

$$\mathcal{L}(\Psi, \zeta, \gamma) = f(\Psi) + g(\zeta) + \gamma^T(A^T\Psi - \zeta) + \frac{\rho}{2}||A^T\Psi - \zeta||_2^2 \tag{10}$$

Note that: $\gamma \in \mathbb{R}^{Nd\times 1}$ and $\zeta \in \mathbb{R}^{Nd\times 1}$

2.1. Updating Ψ

$$\Psi^{k+1} = \underset{\Psi}{\text{armin}} \quad f(\Psi) + (\gamma^k)^T(\mathbf{A}^T\Psi - \zeta^k) + \frac{\rho}{2}||\mathbf{A}^T\Psi - \zeta^k||_2^2 \tag{11}$$

Then the solution is given by:

$$(\rho\mathbf{A}\mathbf{A}^T + \frac{1}{2}\mathbf{I})\Psi^{k+1} = \mathbf{A}(\rho\zeta^k - \gamma^k) - \mathbf{y} \tag{12}$$

$\mathbf{A}\mathbf{A}^H$ can actually be computed very efficiently as follows:

$$\begin{aligned}\mathbf{A}\mathbf{A}^H &= \sum_i^N A_i A_i^H = \sum_i^N(\mathbf{F}diag(\hat{\mathbf{a}}_i)diag(\hat{\mathbf{a}}_i^*)\mathbf{F}^H)\\ &= \mathbf{F}diag\Big(\sum_i^N \hat{\mathbf{a}}_i \odot \hat{\mathbf{a}}_i^*\Big)\mathbf{F}^H\end{aligned} \tag{13}$$

Then the update rule will collapse to be:

$$\begin{aligned}\Psi^{k+1} &= (\rho\mathbf{F}diag\Big(\sum_i^n \hat{\mathbf{a}}_i \odot \hat{\mathbf{a}}_i^*\Big)\mathbf{F}^H + \frac{1}{2}\mathbf{F}\mathbf{F}^H)^{-1}(\mathbf{A}(\rho\zeta^k - \gamma^k) - \mathbf{y})\\ &= \mathbf{F}(\rho diag\Big(\sum_i^N \hat{\mathbf{a}}_i \odot \hat{\mathbf{a}}_i^*\Big) + \frac{1}{2}\mathbf{I})^{-1}\mathbf{F}^H\mathbf{A}(\rho\zeta^k - \gamma^k) - \mathbf{F}(\rho diag\Big(\sum_i^N \hat{\mathbf{a}}_i \odot \hat{\mathbf{a}}_i^*\Big) + \frac{1}{2}\mathbf{I})^{-1}\hat{\mathbf{y}}^*\\ &= \mathbf{F}\frac{\Gamma^k}{(\rho\Big(\sum_i^N \hat{\mathbf{a}}_i \odot \hat{\mathbf{a}}_i^*\Big) + \frac{1}{2})} - \mathbf{F}\frac{\hat{\mathbf{y}}^*}{(\rho\Big(\sum_i^N \hat{\mathbf{a}}_i \odot \hat{\mathbf{a}}_i^*\Big) + \frac{1}{2})}\end{aligned} \tag{14}$$

Where the division is element wise operation. Since:

$$\begin{aligned}\mathbf{A}(\rho\zeta^k - \gamma^k) &= \mathbf{F}[diag(\hat{\mathbf{a}}_1) \mid diag(\hat{\mathbf{a}}_2) \mid .... \mid diag(\hat{\mathbf{a}}_N)][\mathbf{F}^H \otimes \mathbf{I}](\rho\zeta^k - \gamma^k)\\ &= \mathbf{F}\sum_i^N diag(\hat{\mathbf{a}}_i)(\rho\hat{\zeta}^{*k} - \hat{\gamma}^{*k})_i = \mathbf{F}\sum_i^N \hat{\mathbf{a}}_i \odot (\rho\hat{\zeta}^{*k} - \hat{\gamma}^{*k})_i = \mathbf{F}\Gamma^k\end{aligned} \tag{15}$$

Then,

$$\begin{aligned}\Psi^{k+1} &= \mathbf{F}\frac{\Gamma^k}{(\rho\Big(\sum_i^n \hat{\mathbf{a}}_i \odot \hat{\mathbf{a}}_i^*\Big) + \frac{1}{2})} - \mathbf{F}\frac{\hat{\mathbf{y}}^*}{(\rho\Big(\sum_i^n \hat{\mathbf{a}}_i \odot \hat{\mathbf{a}}_i^*\Big) + \frac{1}{2})}\\ &= \mathbf{F}\frac{(\Gamma^k - \hat{\mathbf{y}}^*)}{(\rho\Big(\sum_i^N \hat{\mathbf{a}}_i \odot \hat{\mathbf{a}}_i^*\Big) + \frac{1}{2})}\end{aligned} \tag{16}$$

2.2. Updating ζ

$$\begin{aligned}
\zeta^{k+1} &= \underset{\zeta^k}{\text{argmin}} \quad g(\zeta^k) + \frac{\rho}{2}||A^T\Psi^{k+1} - \zeta^k||_2^2 + (\gamma^k)^T(A^T\Psi^{k+1} - \zeta^k) \\
&= \underset{\zeta^k}{\text{argmin}} \quad \frac{\rho}{2}||A^T\Psi^{k+1} - \zeta^k||_2^2 + (\gamma^k)^T(A^T\Psi^{k+1} - \zeta^k) \\
&\quad \text{subject to} \quad ||\zeta^k||_\infty \leq \lambda \\
&= \underset{\zeta^k}{\text{argmin}} \quad ||\zeta^k - A^T\Psi^{k+1}||_2^2 - \frac{2}{\rho}(\gamma^k)^T\zeta^k \\
&\quad \text{subject to} \quad ||\zeta^k||_\infty \leq \lambda
\end{aligned} \tag{17}$$

Note that:

$$\begin{aligned}
&\underset{\zeta^k}{\text{argmin}} \; ||(\zeta^k - A^T\Psi^{k+1}) - \frac{1}{\rho}\gamma^k||_2^2 \\
&= \underset{\zeta^k}{\text{argmin}} \; ||\zeta^k - A^T\Psi^{k+1}||_2^2 - \frac{2}{\rho}(\gamma^k)^T(\zeta^k - A^T\Psi^{k+1}) + \frac{1}{\rho^2}||\gamma^k||_2^2 \\
&= \underset{\zeta^k}{\text{argmin}} \; ||\zeta^k - A^T\Psi^{k+1}||_2^2 - \frac{2}{\rho}(\gamma^k)^T\zeta^k
\end{aligned} \tag{18}$$

Then the dual problem can be re-written as:

$$\begin{aligned}
&\underset{\zeta^k}{\text{argmin}} \; ||\zeta^k - (A^T\Psi^{k+1} + \frac{1}{\rho}\gamma^k)||_2^2 \\
&\text{subject to} \quad ||\zeta^k||_\infty \leq \lambda
\end{aligned} \tag{19}$$

Then the solution for problem 19 is given by projecting on to the L_∞ ball:

$$\zeta^{k+1} = \begin{cases} A^T\Psi^{k+1} + \frac{1}{\rho}\gamma^k & \text{if } ||A^T\Psi^{k+1} + \frac{1}{\rho}\gamma^k||_\infty \leq \lambda \\ \text{proj}(A^T\Psi^{k+1} + \frac{1}{\rho}\gamma^k)_i & \text{if } \forall\, i \text{ such that} \\ & ||A^T\Psi^{k+1} + \frac{1}{\rho}\gamma^k||_\infty > \lambda \end{cases} \tag{20}$$

where:
$\text{proj}(A^T\Psi^{k+1} + \frac{1}{\rho}\gamma^k)_i = \lambda sign(A^T\Psi^{k+1} + \frac{1}{\rho}\gamma^k)_i$

Therefore, the update for ζ^{k+1} is given by:

$$A^H\Psi^{k+1} + \frac{1}{\rho}\gamma^k = \begin{bmatrix} \mathbf{F}diag(\hat{\mathbf{a}}_1^* \odot \hat{\Psi}^{k+1^*}) + \frac{1}{\rho}\gamma_1^k \\ \mathbf{F}diag(\hat{\mathbf{a}}_2^* \odot \hat{\Psi}^{k+1^*}) + \frac{1}{\rho}\gamma_2^k \\ \vdots \\ \mathbf{F}diag(\hat{\mathbf{a}}_N^* \odot \hat{\Psi}^{k+1^*}) + \frac{1}{\rho}\gamma_N^k \end{bmatrix} \tag{21}$$

And the final update is given by 20 with the vector condition defined for efficacy as in 21.

2.3. Updating γ

$$\gamma^{k+1} = \gamma^k + \rho(A^T\Psi^{k+1} - \zeta^{k+1}) \tag{22}$$