ELEC70037
Topics in Large Dimensional Data Processing
Exercise

Wei Dai

November 16, 2023

EEE Department, Imperial College London

Part IV.

Coursework 3

8.3. Blind Deconvolution: Convex Relaxation

In this section, we observe the output of the cyclic convolution

$$\boldsymbol{y} = \boldsymbol{x} \circledast_n \boldsymbol{h}$$

and try to find the signal $\boldsymbol{x}$ and the channel $\boldsymbol{h}$ using convex optimization. The assumption is that $\boldsymbol{x}$ has small total variation

$$\mathrm{TV}(\boldsymbol{x}) = \sum_m |x_{m+1} - x_m| \,,$$

and $\boldsymbol{h}$ is sparse and hence small l_1-norm.

We formulate the following convex optimization problem:

$$\begin{aligned} \min_{\boldsymbol{X}} \; & \frac{1}{2} \|\boldsymbol{y} - \mathscr{A}(\boldsymbol{X}_1)\|_2^2 \\ & + \lambda_1 \|\boldsymbol{X}_2\|_* + \lambda_2 \mathrm{TV}_{\mathrm{col}}(\boldsymbol{X}_1; \boldsymbol{X}_3) + \lambda_3 \|\boldsymbol{X}_4\|_{2,1} \\ \text{s.t.} \; & \boldsymbol{X}_2 = \boldsymbol{X}_1, \; \boldsymbol{X}_3 = (\boldsymbol{X}_1)_{2:m,:} - (\boldsymbol{X}_1)_{1:m-1,:} \,, \; \boldsymbol{X}_4 = \boldsymbol{X}_1, \end{aligned}$$

where $\|\cdot\|_*$ denotes the nuclear norm, $\mathrm{TV}_{\mathrm{col}}$ represents row-wise total variation

$$\mathrm{TV}_{\mathrm{col}}(\boldsymbol{X}_1; \boldsymbol{X}_3) = \sum_m \|(\boldsymbol{X}_1)_{m+1,:} - (\boldsymbol{X}_1)_{m,:}\|_2 = \|\boldsymbol{X}_3^\top\|_{2,1} \,,$$

and $\|\cdot\|_{2,1}$ stands for the $l_{2,1}$-norm

$$\|\boldsymbol{X}\|_{2,1} = \sum_n \|\boldsymbol{X}_{:,n}\|_2 \,.$$

We shall solve the above problem using ADMM.

1. Write the explicit form of the augmented Lagrangian (similar to (8.1)). [1]
2. Write a function to update $\boldsymbol{X}_1$ when all other variables are fixed. Save the value of $\boldsymbol{X}_1^1$ into the data file. [4]
3. Write a function to update $\boldsymbol{X}_{2:4}$ when all other variables are fixed. Save the values of $\boldsymbol{X}_{2:4}^1$ into the data file. [9]
4. Write a function to update $\boldsymbol{U}_{2:4}$ when all other variables are fixed. Save the values of $\boldsymbol{U}_{2:4}^1$ into the data file. [3]
5. Now write a function to implement the ADMM algorithm including the stopping criteria mentioned in the lecture notes. Here we use the error tolerance constants $\epsilon_{\mathrm{abs}} = \epsilon_{\mathrm{rel}} = 10^{-4}$. Save your final result $\hat{\boldsymbol{X}}$ into the data file. [4]
   Note that the hyperparameters affect the critical point and convergence rate. In this particular part, you are allowed to change the hyperparameters in order to get good numerical results.

8.4. Blind Deconvolution

Similar to the last section, we observe the output of the cyclic convolution

$$y = x \circledast_n h$$

and try to find the signal $\boldsymbol{x}$ and the channel $\boldsymbol{h}$. The assumption is that $\boldsymbol{x}$ has small total variation

$$\mathrm{TV}(\boldsymbol{x}) = \sum_m |x_{m+1} - x_m| ,$$

and $\boldsymbol{h}$ is sparse and hence small l_1-norm. We also assume that $\|\boldsymbol{x}\|_2 = 1$.

We formulate the following optimization problem:

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{h}} \ & \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{x} \circledast \boldsymbol{h}\|_2^2 + \delta\left(\|\boldsymbol{x}\|_2 = 1\right) \\ & + \lambda_1 \|\mathrm{TV}\left(\boldsymbol{x}\right)\|_0 + \lambda_2 \|\boldsymbol{h}\|_0 . \end{aligned}$$

To proceed, we relax it into

$$\begin{aligned} \min_{\boldsymbol{x}_{1:3},\boldsymbol{h}_{1:2}} \ & \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{x}_1 \circledast \boldsymbol{h}_1\|_2^2 + \delta\left(\|\boldsymbol{x}_2\|_2 = 1\right) \\ & + \lambda_1 \|\boldsymbol{x}_3\|_0 + \lambda_2 \|\boldsymbol{h}_2\|_0 \\ & + \frac{\alpha}{2} \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|_2^2 \\ & + \frac{\alpha}{2} \|\boldsymbol{x}_3 - ((\boldsymbol{x}_1)_{2:m} - (\boldsymbol{x}_1)_{1:m-1})\|_2^2 \\ & + \frac{\alpha}{2} \|\boldsymbol{h}_2 - \boldsymbol{h}_1\|_2^2 , \end{aligned}$$

where $\alpha > 0$.

We solve the above optimization problem using alternating minimization.

1. With fixed $\boldsymbol{h}_{1:2}$, we update $\boldsymbol{x}_{1:3}$ using proximal gradient method (with only one iteration).
   a) Find the Hessian matrix of

   $$\frac{\alpha}{2} \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|_2^2 + \frac{\alpha}{2} \|\boldsymbol{x}_3 - ((\boldsymbol{x}_1)_{2:m} - (\boldsymbol{x}_1)_{1:m-1})\|_2^2 .$$

   [3]
   b) Find the upper bound (τ_{ub}) of the step size in the proximal gradient step to update $\boldsymbol{x}_{1:3}$. [1]
   c) Choose the step size as $0.8\tau_{\mathrm{ub}}$ and update $\boldsymbol{x}_{1:3}$ using one iteration of the proximal gradient method. Save your result into the data file. [2+2+2=6]
   d) Evaluate the subgradient with respect to $\boldsymbol{x}_{1:3}$ [3]
2. With fixed $\boldsymbol{x}_{1:3}$, we update $\boldsymbol{h}_{1:2}$ using proximal gradient method (with only one iteration).

a) Find the Hessian matrix of

$$\frac{\alpha}{2} \|\boldsymbol{h}_2 - \boldsymbol{h}_1\|_2^2$$

[1]

b) Find the upper bound (τ_{ub}) of the step size in the proximal gradient step to update $\boldsymbol{h}_{1:2}$. [1]

c) Choose the step size as $0.8\tau_{\text{ub}}$ and update $\boldsymbol{h}_{1:2}$ using one iteration of the proximal gradient method. Save your result into the data file. [2+2=4]

d) Evaluate the subgradient with respect to $\boldsymbol{h}_{1:2}$ [2]

3. Evaluate the subgradient with respect to $\boldsymbol{x}_{1:3}$ using the updated $\boldsymbol{x}_{1:3}$ and $\boldsymbol{h}_{1:2}$. [2]
4. Complete the alternating minimization method where each iteration involves one proximal gradient step to update $\boldsymbol{x}_{1:3}$ and that to update $\boldsymbol{h}_{1:2}$.
   Include that the number of alternating minimization is at most 500 into your stopping criteria.
   Save your results into the data file.
   Note that the hyperparameters affect the critical point and convergence rate. In this particular part, you are allowed to change the hyperparameters in order to get good numerical results. [4]