SDS 385: Exercises 8 - Spacial smoothing at scale

November 7, 2016

Professor Scott

Spencer Woody

Problem 1

Laplacian smoothing

Normalsize text

$$
\begin{aligned}
\hat{x} &= \arg\min_x \frac{1}{2} \|y - x\|_2^2 + \frac{\lambda}{2} \|Dx\|_2^2 && (1) \\
&= \arg\min_x \frac{1}{2} \|x - y\|_2^2 + \frac{\lambda}{2} \|Dx\|_2^2 && (2) \\
&= \arg\min_x \frac{1}{2} \|x - y\|_2^2 + \frac{\lambda}{2} x^T D^T D x && (3)
\end{aligned}
$$

Simply taking the gradient of this objective function yields

$$
\begin{aligned}
(\hat{x} - y) + \lambda D^T D \hat{x} = 0 && (4) \\
(I + \lambda D^T D)\hat{x} = y && (5) \\
\hat{x}(I + \lambda D^T D)^{-1} y && (6)
\end{aligned}
$$

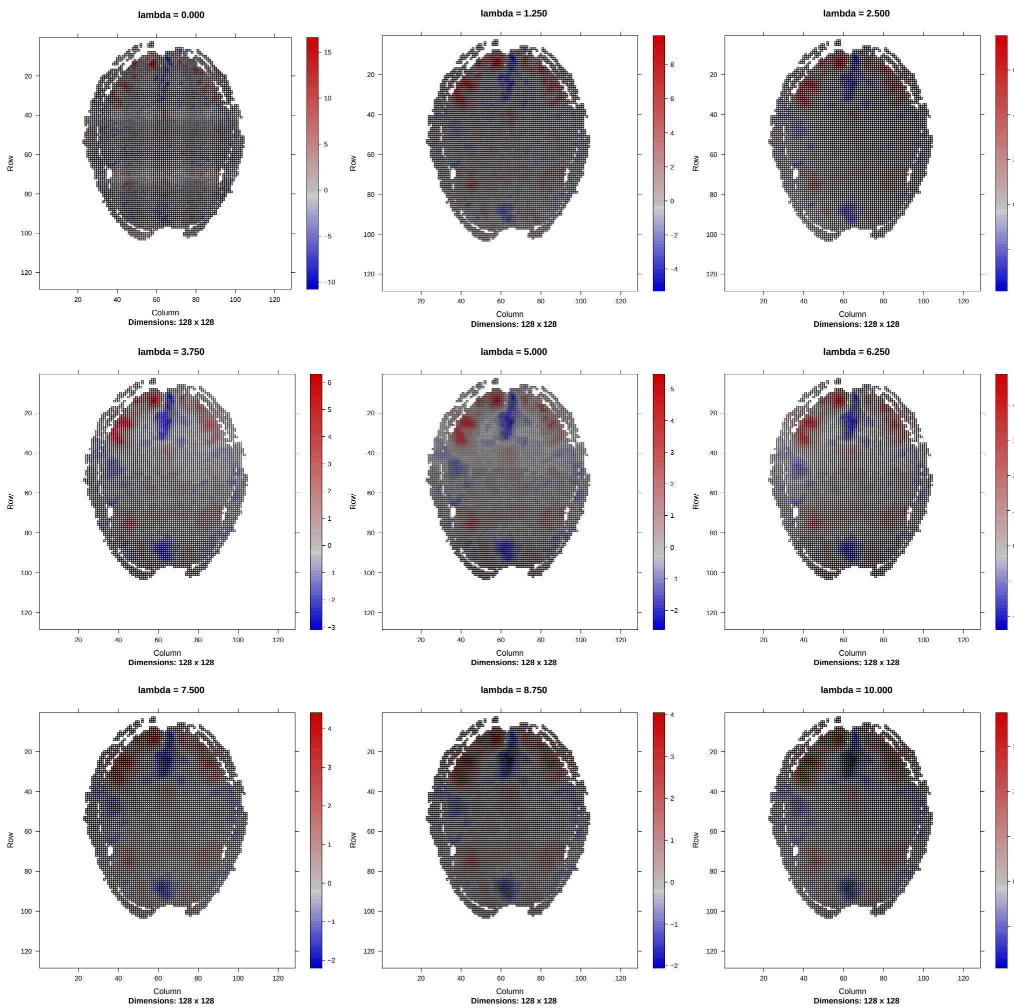

Figure 1: Comparison of smoothing for various λ values

Problem 2

Graph fused lasso

Now the minimization problem becomes

$$\min \frac{1}{2} \|y - x\|_2^2 + \lambda \|Dx\|_1 \tag{7}$$

and then we set up the ADMM problem

$$\begin{aligned} &\min \frac{1}{2} \|y - x\|_2^2 + \lambda \|r\|_1 &(8)\\ &\text{s.t. } Dx - r = 0. &(9)\end{aligned}$$

The augmented Lagrangian is

$$L_p(x, r, v) = \frac{1}{2} \|y - x\|_2^2 + \lambda \|r\|_1 + v^T(Dx - r) + \frac{\rho}{2} \|Dx - r\|_2^2. \tag{10}$$

We define our scaled dual variable $u^k = \frac{v^k}{\rho}$. Finally, our updates are

$$\begin{aligned} x^{k+1} &= \arg\min_x L_p(x, r^k, v^k) &(11)\\ &= \arg\min_x \frac{1}{2} \|y - x\|_2^2 + \lambda \left\|r^k\right\|_1 + (v^k)^T(Dx^k - r) + \frac{\rho}{2} \left\|Dx - r^k\right\|_2^2 &(12)\\ &= \arg\min_x \frac{1}{2} \|y - x\|_2^2 + (Dx - r^k)^T v^k + \frac{\rho}{2} \left\|Dx - r^k\right\|_2^2 &(13)\\ &= \arg\min_x \frac{1}{2} \|x - y\|_2^2 + (Dx - r^k)^T v^k + \frac{\rho}{2} \left\|Dx - r^k\right\|_2^2 &(14)\end{aligned}$$

Taking the gradient of this objective function and setting it to zero yields

$$\begin{aligned} x^{k+1} - y + D^T v^k + \rho(D^T D x^{k+1} - D^T r^k) &= 0 &(15)\\ (I + \rho D^T D) x^{k+1} &= y - D^T v^k + \rho D^T r^k &(16)\\ &= y + \rho D^T (r^k - u^k). &(17)\end{aligned}$$

$$\begin{aligned} r^{k+1} &= \arg\min_r L_p(x^{k+1}, r, v^k) &(18)\\ &= \arg\min_r \frac{1}{2} \left\|y - x^{k+1}\right\|_2^2 + \lambda \|r\|_1 + (v^k)^T(Dx^{k+1} - r) + \frac{\rho}{2} \left\|Dx^{k+1} - r\right\|_2^2 &(19)\\ &= \arg\min_r \lambda \|r\|_1 + (Dx^{k+1} - r)^T v^k + \frac{\rho}{2} \left\|r - Dx^{k+1}\right\|_2^2 &(20)\\ &= \arg\min_r \frac{\lambda}{\rho} \|r\|_1 - \frac{1}{\rho}(r - Dx^{k+1})^T v^k + \frac{1}{2} \left\|r - Dx^{k+1}\right\|_2^2 &(21)\\ &= \arg\min_r \frac{\lambda}{\rho} \|r\|_1 - (r - Dx^{k+1})^T u^k + \frac{1}{2} \left\|r - Dx^{k+1}\right\|_2^2 &(22)\\ &= S_{\lambda/\rho}(Dx^{k+1} + u^k) &(23)\end{aligned}$$

$$\begin{aligned} v^{k+1} &= v^k + \rho(Dx^{k+1} - r^{k+1}) &(24)\\ u^{k+1} &= u^k + Dx^{k+1} - r^{k+1} &(25)\end{aligned}$$

Now we need to define the stopping criteria. First define the primal and dual residuals

$$\begin{aligned} q^{k+1} &= Dx^{k+1} - r^{k+1} &(26)\\ s^{k+1} &= -\rho D^T (r^{k+1} - r^k) &(27)\end{aligned}$$

respectively. We reach convergence when both conditions

$$\left\| q^{k+1} \right\|_2 \leq \epsilon^{\text{prim}} = \sqrt{m}\epsilon^{\text{abs}} + \epsilon^{\text{rel}} \max\left\{ \left\| Dx^{k+1} \right\|_2, \left\| r^{k+1} \right\|_2 \right\} \tag{28}$$

$$\left\| s^{k+1} \right\|_2 \leq \epsilon^{\text{dual}} = \sqrt{n}\epsilon^{\text{abs}} + \epsilon^{\text{rel}} \left\| \rho D^T u^{k+1} \right\|_2 \tag{29}$$

hold for some positive values of ϵ^{abs} and ϵ^{rel}, somewhere on the order of 10^{-3}.

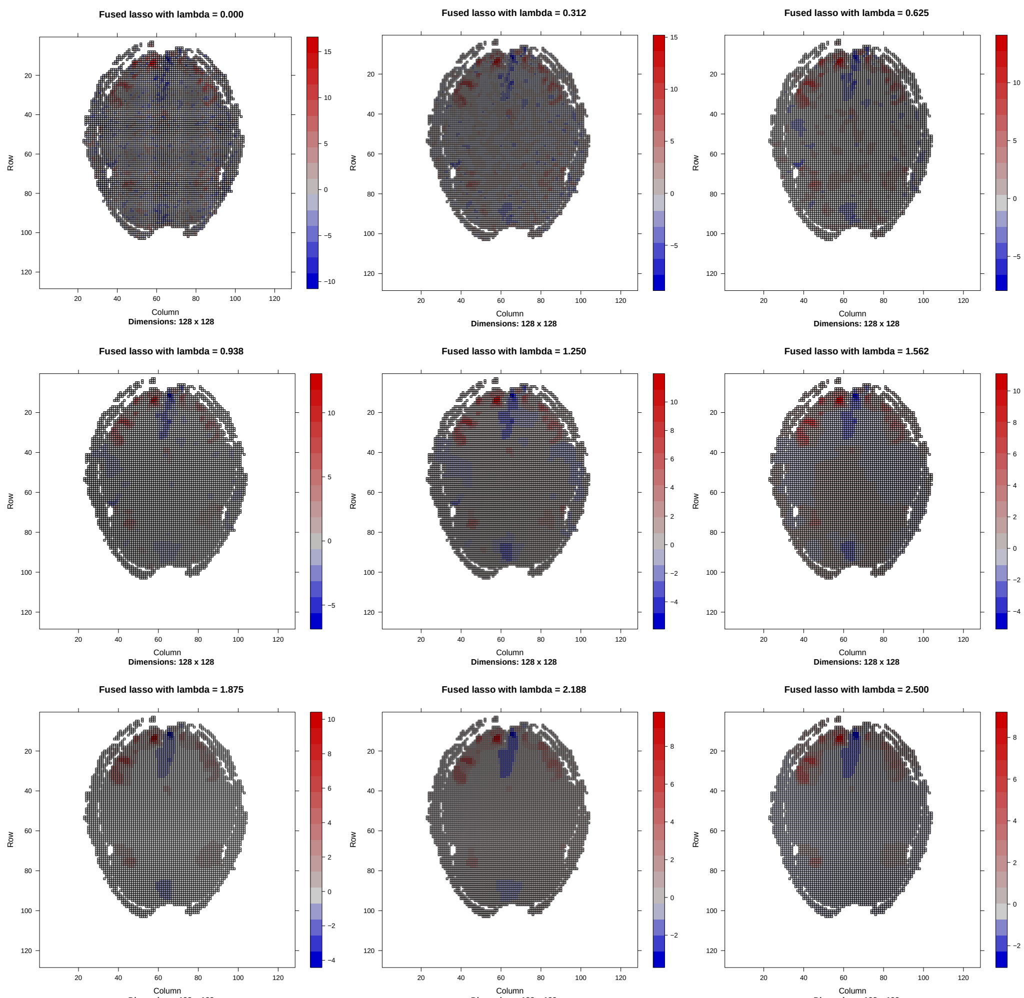

Figure 2: Comparison of fused lasso for various λ values

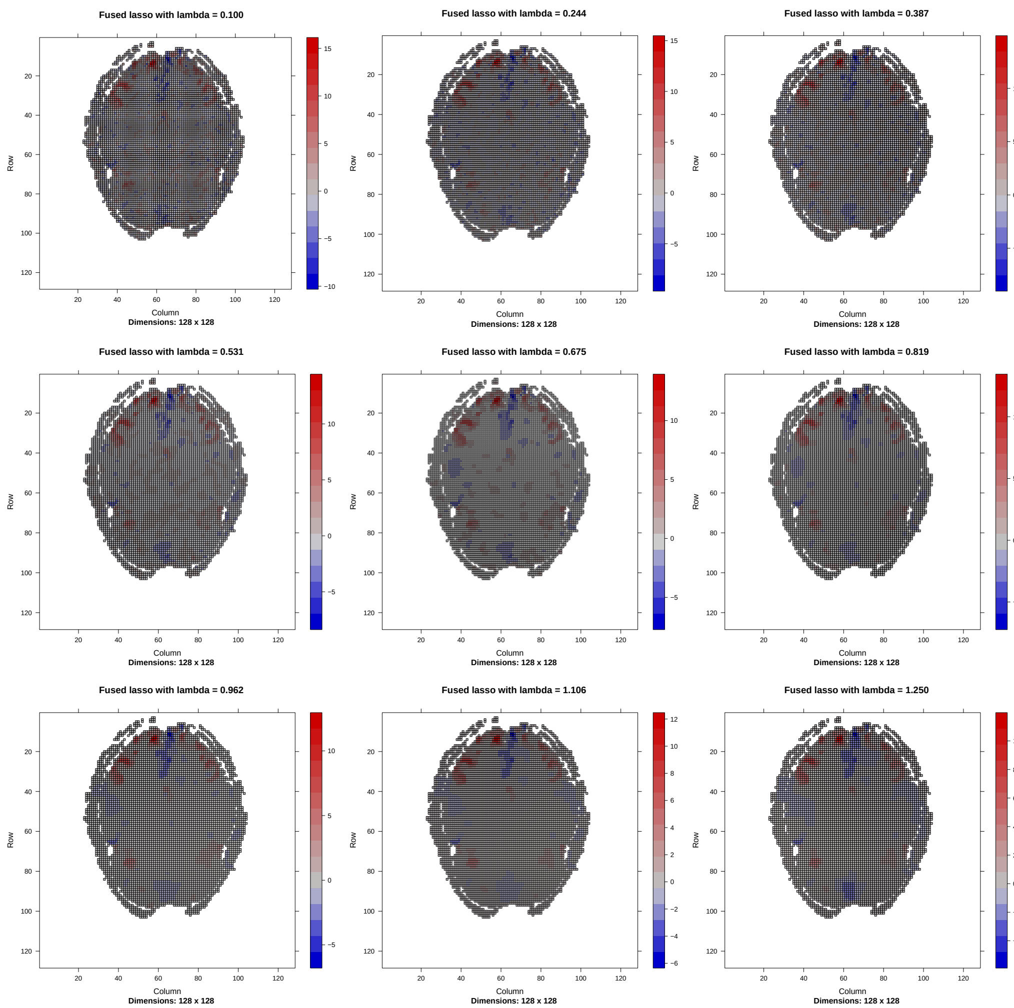

Figure 3: Comparison of fused lasso for smaller λ values

R script for `myfuns.R`

R script for `e6.R`